Music and Language Comprehension in the Brain

Cover design: Jorrit Kiel

ISBN 978-90-76203-79-9

Printed and bound by Ipskamp Drukkers b.v.

The research reported in this thesis was supported by a PhD grant by the Max Planck society (Max-Planck-Gesellschaft) awarded to Richard Kunert and a Spinoza Prize by the Dutch science organization (Nederlandse Organisatie voor Wetenschappelijk Onderzoek) awarded to Peter Hagoort.

Music and Language Comprehension in the Brain

Proefschrift

ter verkrijging van de graad van doctor

aan de Radboud Universiteit Nijmegen

op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,

volgens besluit van het college van decanen

in het openbaar te verdedigen op vrijdag 10 februari 2017

om 12:30 uur precies

door

Richard Kunert

geboren op 3 december 1985

te Halberstadt (Duitsland)

Promotor:

Prof. Peter Hagoort

Copromotor:

Dr. Roel M. Willems

Manuscriptcommissie:

Prof. James McQueen (Radboud Universiteit)

Prof. Sonja A. Kotz (Universiteit Maastricht)

Prof. Henkjan Honing (Universiteit van Amsterdam)

Table of Contents

Foreword

Many a days during my PhD I asked myself why on earth I am doing a PhD on music and language. You, dear reader, might ask yourself the same question about this PhD thesis. Why read it? I describe experiments which *grosso modo* suggest that music and language comprehension partly rely on shared syntactic processing circuitry in the brain. So what? Why spend time on such research? I hope that the following four reasons can convince you to read this thesis as much as they convinced me to keep writing it despite all the obstacles on the way.

1) The idea that music and language share syntactic processing resources has behavioural implications that are hard to believe. My own father would not dare to predict that reading a syntactically difficult sentence changes music perception, yet it does (see chapter 2). This is not just fascinating but has implications for the principles guiding brain organization (Patel, 2013).

2) Music and language are uniquely human communication systems. Few people realize that their pet, for example their dog or cat, has absolutely no idea what humans do when they dance (Schachner, Brady, Pepperberg, & Hauser, 2009). The precise, abstract, and structured communication possible via language is equally entirely beyond them. Thus, investigating the music and language faculties also means investigating what makes us human.

3) Shared processing resources for music and language offer treatment targets for people suffering from music or language problems. For example, patients with brain damage resulting in aphasia are known to have concurrent music perception problems (Patel, Iversen, Wassenaar, & Hagoort, 2008). Moreover, children with specific-language impairment display impaired music-structural processing when hearing chords (Jentschke, Koelsch, Sallat, & Friederici, 2008). Thus, this thesis could be of use as inspiration for treating language problems via a musical route.

4) Shared music-language resources also provide targets for improving education, given the more mature language processing in children following musical training (Jentschke & Koelsch, 2009). Given that language is a basic skill needed to acquire many other faculties taught via written text, music training might offer a way to optimize educational practice.

I hope that with these thoughts in the back of your mind, you, dear reader, will find the research reported in this thesis as fascinating as I do.

Nijmegen, 19/7/2016

Chapter 1

Introduction

1|1 Music and language in daily life

During the more than three years of research which this thesis summarizes, I, just like many people reading this book, have unknowingly performed the complex acts of brain processing which form the basis of this thesis. After getting up I would typically shower and sing a tune still in my head from the day before. Luckily, singing is one of the few kinds of music production that do not require any expertise. Moreover, it is a powerful example of how music and language can be combined, both in terms of the production mechanism (the speech apparatus), the medium (sound), and the reception mechanism (the ear and auditory system). It will come as a surprise to many readers that this thesis about how music and language comprehension relate to each other in the brain is not concerned with everyday examples of music-language combinations like singing. No study is concerned with participants singing songs. The main topic of this PhD thesis is better exemplified by what happens after the shower.

Having left the shower, I sit down and eat breakfast. The radio is switched on, playing music. At the same time, I read the newspaper. While I had just actively combined music and language in song, I now strictly keep them apart. Many readers will think that under such circumstances the brain can process language and music perfectly separately. The two cognitive domains do not share a production mechanism (newspaper versus loud speakers), a medium (light versus sound), or a reception mechanism (the eyes versus the ears). It is under these circumstances that a surprising commonality between music and language, which forms the basis of this thesis, emerges: both are structured sequences.

1|2 A key commonality between music and language: structured sequences

In order to understand the key commonality between music and language, one has to understand the limitations of the most common medium which they both occupy: sound. Both the production (e.g., speaking) and the reception (hearing) of sound limits the medium to one piece of information at a time. Plainly said, the ear drum cannot react in two different ways at the same time. As a result, music and language can be perfectly translated into two-dimensional representations of time and amplitude, as in Figure 1|1 showing increasingly smaller excerpts of Beethoven's *9th* symphony (soprano voice of Ode to Joy played by piano, op. 125). Such a representation is ideal for conveying the simplicity of the medium but very poor for conveying the complexity of the message. One needs to turn from the medium (sound) to the reception system (ears and auditory processing regions in the brain) in order to understand how complexity as a result of structure emerges in music and language.

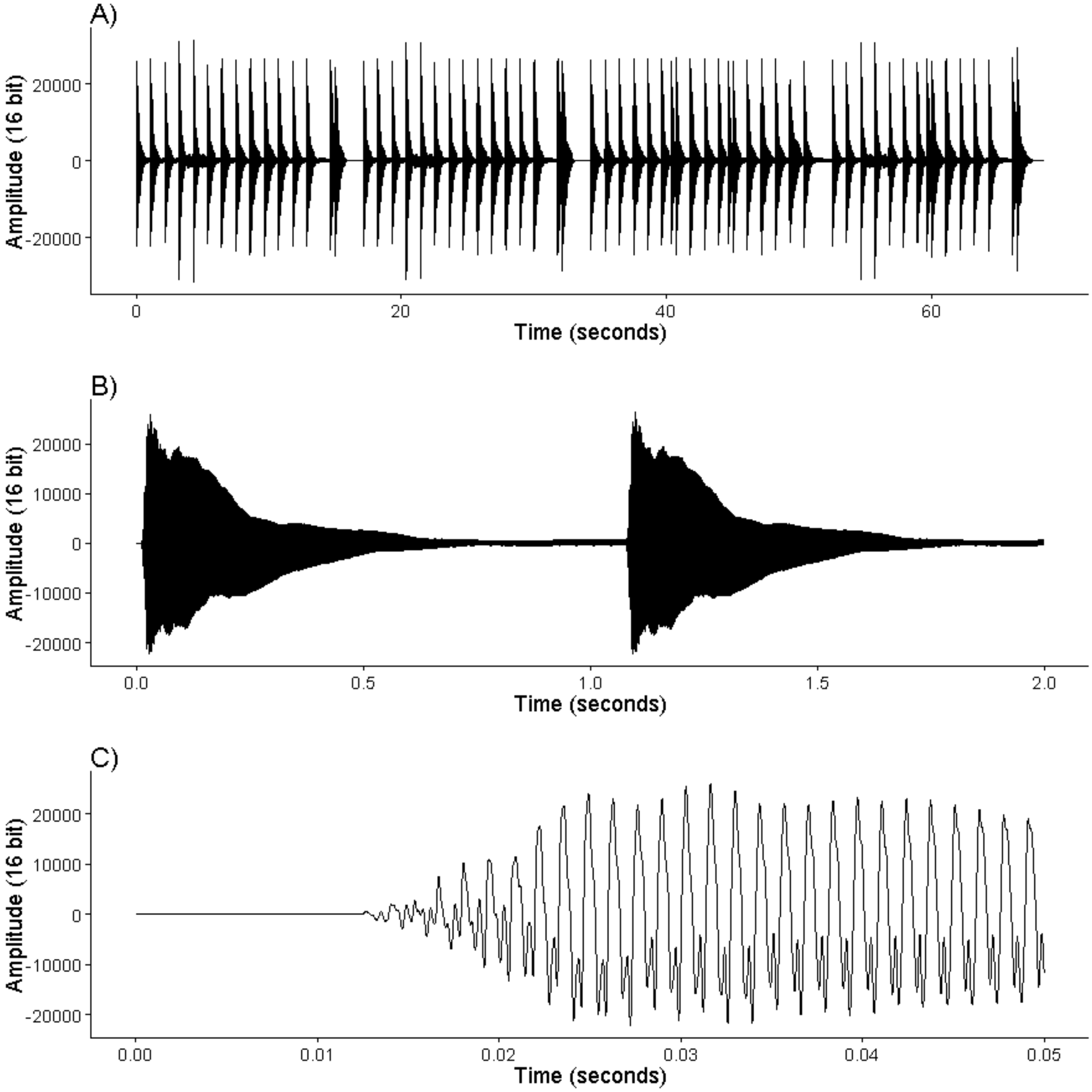


Figure 1|1. Sound wave of Beethoven's Ode to Joy (op. 125, soprano voice) played with a piano timbre, shown in increasingly finer temporal detail. Note that panels **A)** and **B)** include the same level of spectral detail as panel **C)**.

What happens when sound arrives at the ear? The ear-drum vibrates, setting the ossicles in the middle ear in motion which, in turn, move the fluid in the cochlea. There, hair cells in the Organ of Corti react to the fluid set in motion and change an adjacent nerve's firing. This way, sound is translated into the electro-chemical information currency of the brain. The nerve's firing pattern travels through various neural structures, arriving eventually at the primary auditory cortex situated in the temporal lobe (BA 41 and 42). It is generally thought that up until this point auditory stimuli are treated the same whether they are linguistic, musical or something else.

After basic auditory processing, the brain categorizes the stimulus. The details of this music-language categorization remain obscure. Of course, an instrumental piece of music, as represented in Figure 1|1, is easily categorized as musical given the musical timbres. However, human vocalizations are more ambiguous. It is likely that categorizing a stimulus as speech rather than song is facilitated by shorter relative timing changes (phonetic categories depend on temporal differences of tens of milliseconds), and less prominent spectral (pitch) cues (Zatorre, Belin, & Penhune, 2002). Still, the speech-to-song illusion, in which a speech stimulus is perceived as sung as a result of repetition, suggests that the music-language categorization is, at least for some stimuli, unstable (Deutsch, Henthorn, & Lapidis, 2011; Falk, Rathcke, & Dalla Bella, 2014; Vanden Bosch der Nederlanden, Hannon, & Snyder, 2015).

Once a stimulus is categorized as being linguistic or musical it is interpreted domain-specifically. This opens the way for using implicitly learned rules for how elements (words or chords) relate to each other in order to understand the stimulus's overall structure. In language, the rules I will focus on are syntactic in nature. These rules can be used once the words have been identified from the speech stream. They govern the structural relation between words.

Figure 1|2 shows a syntactic interpretation of a nonsense sentence taken from Chomsky (1956). The first thing to note is that not all words are interpreted equally in syntactic terms. Once the listener has identified the language of an utterance, s/he assigns word classes to words. Words belonging to the same word class play a similar role in a sentence. For example, all adjectives qualify nouns, while adverbs qualify other words such as verbs. Nonsense sentences with grammatical structure, so-called Jabberwocky sentences, only include word class information without any meaningful word, suggesting that the process of word class interpretation can be performed in the absence of semantic information. An example would be this extract from Carroll's (1872/1999) poem Jabberwocky:

(1|1)
`Twas brillig, and the slithy toves
Did gyre and gimble in the wabe;
All mimsy were the borogoves,
And the mome raths outgrabe.

Figure 1|2 further shows that word classes are not simply interpreted as a linear string. Instead, there is a rich syntactic structure governing the relation between the words, as shown in the syntactic tree of Figure 1|2. For example, the speech stream does not easily reveal that the words *ideas* and *sleep* have greater structural importance than the other three words. *Ideas* and *sleep* are syntactic heads, i.e. they determine the type of phrase (the noun *ideas* determines the noun phrase status of *colorless green ideas*). Notice that nothing in the speech stream itself points to this importance. For example, the serial position of *ideas* is at the end of its phrase while the serial position of *sleep* is at the beginning. In sum, human listeners use syntactic rules in order to enrich the speech stream and comprehend a sentence’s structure. I shall reserve the term *syntactic integration* for the rule-based combination of elements like words into an overarching structure such as a sentence.

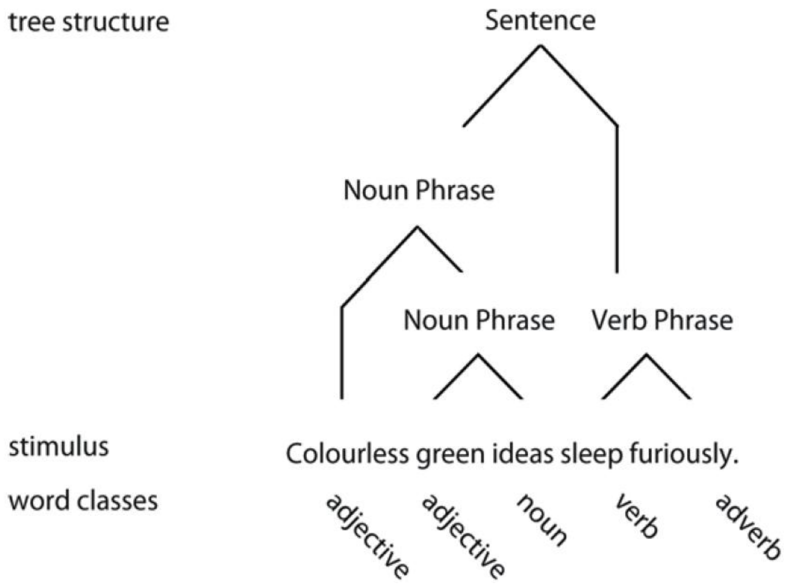


Figure 1|2. A phrase structure representation of an example sentence.

In music, similar structural rules exist. In this thesis, I will focus on structural rules based on harmony in the Western tonal tradition, without meaning to imply that rhythm or timbre do not follow similar structural rules (Herdener et al., 2014; Lerdahl & Jackendoff, 1983) or that other music cultures are without similar rules (Brown & Jordania, 2013). The basic elements of musical structures are not words, as in language, but pitches. The Western musical tradition defines 12 pitches per octave. The minimal distance between pitches is one semitone (approximately 6% difference in pitch). This organization is repeated in every octave leading to octave-equivalent pitch classes (e.g., all the C-notes on a piano keyboard), i.e. the Western tradition knows only 12 unique pitch classes (C, C#/Db, D, D#/Eb, E/Fb, E#/F, F#/Gb, G, G#/Ab, A, A#/Bb, B/Cb).

Some listeners, possessors of absolute pitch (AP), have access to long term memory representations of musical pitches which are so detailed as to include the exact frequency and label of tones (Miyazaki, 1988; Sacks, 1995). Listeners without AP are usually only able to identify mistunings in the context of other tones. For example, listeners to the König-organ in the St. Stevenskerk in Nijmegen are usually unaware that all the tones they hear are mistuned by modern standards (the modern pitch standard is 440 Hz while the organ is tuned substantially lower at 425 Hz), but the context of organ tones following the same pitch standard does not ‘reveal’ the mistuning.

The pitches arriving at the ear are not an unorganized sequence of pitch classes. They can be sounded simultaneously and form a chord. I shall stay with tones but what follows is applicable to chords as well. Tones typically cluster in scales. The most widely heard major and minor scales include seven pitch classes (e.g., C-major includes C, D, E, F, G, A, B). A music listener usually extracts the harmonic key quickly and thereafter expects the upcoming tones to be members of the scale that the key corresponds to. Out-of-key tones (e.g., C-major does not include C#, D#, F#, G#, A#) are relatively unexpected. Even among the expected tones, there is a hierarchy of expectation with the tonic (C in C-major) usually being the most expected tone, i.e. the tonal center(Krumhansl, 1979; Krumhansl & Kessler, 1982). Proximity in pitch is irrelevant for this organization. In C-major, C is cognitively distant from C# while it is very close in terms of pitch. The opposite is true for C and G. This example illustrates one way of interpreting the incoming pitches in terms of harmonic rules.

Harmonic structure theorizing goes even further. The different levels of expectation of a tone in a harmonic key lead to different diatonic functions. For example, when the tonic (C in C -major) is preceded by the second most stable chord, the dominant (G in C -major), they form an authentic cadence, usually signifying a moment of closure often found at the end of a musical piece. If the same dominant is followed by the submediant (A in C -major), the two tones form a deceptive cadence, which can be used to lead to the expectation that a phrase gets extended. An extension of these principles can be used to derive musical phrase structure representations (Figure 1|3) which visually resemble the linguistic phrase structures such as the one in Figure 1|2 (Koelsch, Rohrmeier, Torrecuso, & Jentschke, 2013; Lerdahl & Jackendoff, 1983). Notice that in Figure 1|3 the serial position of a tone in the sequence does not easily translate into its harmonic role. The second tone (subdominant) is more closely structurally related to the third tone (supertonic) than the first tone (tonic). Again, this illustrates that music listeners can infer a rich structure behind the stream of pitches they hear.

Please note that the visual resemblance between the syntactic trees in Figure 1|2 and Figure 1|3 should not be taken to mean that the same structures are involved (but see Van de Cavey & Hartsuiker, 2016). For example, it is commonly agreed that syntactic structures in language require hierarchical processing while in music this is a matter of debate (Koelsch et al., 2013; Tillmann & Bigand, 2004). Moreover, linguistic phrase structures tend to be left branching (e.g., the syntactic head introduces the phrase as in the verb phrase of Figure 1|2) while musical phrase structures tend to be right branching (tones lead up to the most stable pitch as in Figure 1|3). Finally, the elements making up the syntactic trees are fundamentally different between phrase structures in language and music, e.g., there is no musical equivalent to a verb or a verb phrase.

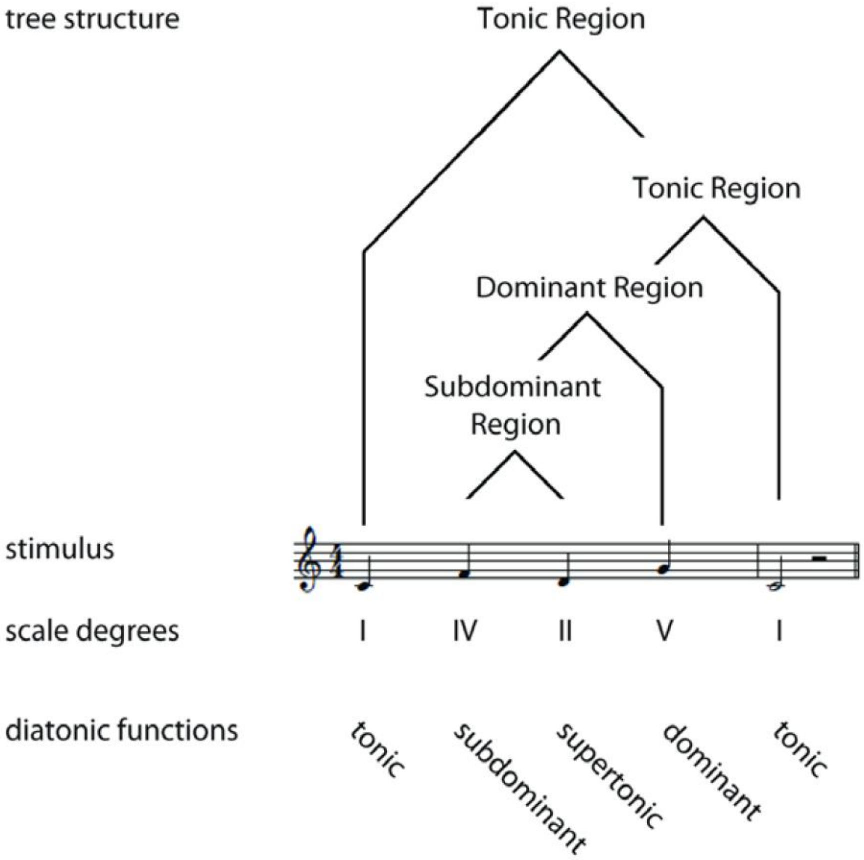


Figure 1|3. Structural representation of a simple cadence. Note that the tree structure conforms to the generative syntax model by Rohrmeier (2011).

Despite clear formal differences between structure in music and language, the key commonality remains. Once the sounds which arrived at the ear have been classified as music or language, they are interpreted according to implicitly learned rules for how sounds combine.

This way, basic elements (words or tones/chords) can form higher order structures (sentences or melodies/harmonic sequences) which are not obvious to an outsider who is unaware of the combinatorial rules governing the stimulus. Formal instruction is not necessary for this ability, even in the case of music (Bigand & Poulin-Charronnat, 2006; Tillmann, Bharucha, & Bigand, 2000).

1|3 The relation between structural processing in music and language

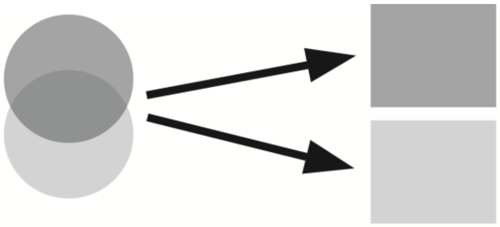
Given the formal similarities between language and music as two examples of structured sequences, many researchers wondered whether the brain exploits this similarity in order to use the same neural circuitry for the structural processing of both music and language stimuli. Early investigations in the 1990s suggested that this is not the case. They focused on music and semantic processing in language (Besson, Faïta, Peretz, Bonnel, & Requin, 1998) or more globally on the effect of brain damage on music and language understanding in general (Griffiths et al., 1997; Peretz, 1993; Peretz et al., 1994; Peretz, Belleville, & Fontaine, 1997). In each case, they suggested a clear differentiation between the processing of stimuli from both cognitive domains from the moment a stimulus had been classified as music or language.

Peretz & Coltheart (2003) summarized the cases of brain damage leading to musical impairments without language problems or language impairments without music problems. Their highly influential article suggests three music-specific processing modules, i.e. neural circuitry specific to music which is not shared with language: pitch organization, the musical lexicon, and the vocal plan formation. It is worth noticing that the pitch organization module includes a tonal encoding submodule which processes the aforementioned structural relations between pitches. According to this modular view of music-language processing, structural processing of stimuli is wholly independent between music and language.

However, evidence conflicting with this strictly modular account quickly emerged. Patel, Gibson, Ratner, Besson, & Holcomb (1998) compared the electrical brain response to syntactically easy versus difficult sentences (the critical word was either expected, an unexpected word class triggering the need to syntactically re-evaluate the sentence context, or a false word class) and the brain response to processing easy versus difficult harmonic sequences (a critical chord was in-key in easy sequences, or out-of-key in difficult sequences).

Surprisingly, structural language and music manipulations resulted in a very similar brain response, a positive deflection in the electro-encephalogram (EEG) around 600 ms after the critical stimulus, a so-called P600 event-related potential (ERP). This result suggests that the same neural circuitry, which results in a characteristic electrical brain response, is involved in structural processing in music and language.

A few years later, functional magnetic resonance imaging (fMRI) and magneto-encephalography (MEG) studies found brain regions typically associated with language to also be sensitive to structural relations in music (Koelsch et al., 2002; Maess, Koelsch, Gunter, & Friederici, 2001; Tillmann, Janata, & Bharucha, 2003), suggesting overlapping or shared music-language neural circuitry. In a highly influential article, Patel (2003) attempted to reconcile 1) the aforementioned evidence for distinct music and language circuitry found by investigating language semantics (Besson et al., 1998) and brain damaged patients (Peretz & Coltheart, 2003) and 2) the evidence for shared processing using brain imaging studies of structural processing in healthy participants (e.g., Patel et al., 1998). He suggested a dual-system model distinguishing domain-general neural circuitry sensitive to both structural relations in music and in language, and domain-specific neural circuitry sensitive to either music or language, see Figure 1|4. This model, the Shared syntactic integration resource hypothesis (SSIRH) is the dominant theoretical account of music and language processing in the brain which I evaluate in this thesis.



Role	Syntactic integration	Storage of structural representations
Location	Frontal brain regions	Posterior brain regions
Shared between music and language?	Partly	No

Figure 1|4. Summary of the Shared syntactic integration resource hypothesis (SSIRH; Patel, 2003)

Shared neural circuitry, hypothesized in frontal regions, is thought to be responsible for syntactic integration. Specifically, when an unexpected word or tone/chord is encountered its unexpectedness means that it is not immediately available for structural processing. In this case, syntactic integration resources rapidly and selectively increase the activation level of the required representation in posterior regions up to threshold after which integration with the structural context can take place. The processing cost of this operation is operationalized in terms of language as distance in words between the currently processed representation and the integration site, in line with dependency locality theory (Gibson, 1998, 2000). For example, the sentences below are difficult (1|2a) or easy (1|2b) to process under this account.

(1|2)

a) *After | the trial | the attorney | advised | **the defendant** | <u>was</u> | likely | to commit | more crimes.*

b) *After | the trial | the attorney | advised that | **the defendant** | <u>was</u> | likely | to commit | more crimes.*

The critical word (underlined) is either expected (1|2b) and, thus, does not need to be rapidly activated by integration resources shared between music and language, or not (1|2a). In the latter case, encountering the unexpected verb *<u>was</u>* signals that the preceding noun phrase ***the defendant*** should be regarded as a subject in a new clause, instead of a direct object interpretation of ***the defendant***. Thus, in (1|2a) but not in (1|2b) syntactic integration resources need to change the activation level of the subject interpretation of ***the defendant*** and the activation level of a currently encountered verb (*<u>was</u>*). According to the SSIRH this taxes shared syntactic integration resources.

The integration cost in terms of music is operationalised as tonal distance according to tonal pitch space theory (Lerdahl, 2001). This refers to the expectation of encountering a tone/chord in a particular harmonic key (independent of the tree-like structures as shown in Figure 1|3). For example, an in-key chord is more expected than an out-of-key chord. Encountering the latter likely taxes shared syntactic integration resources.

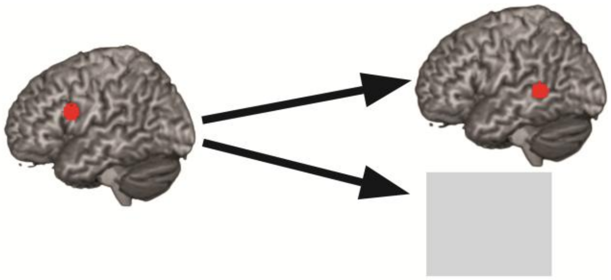
The SSIRH makes a number of interesting predictions. First of all, for patients suffering from Broca's aphasia which is thought to result from damage to the frontal syntactic integration resources 'SSIRH predicts that syntactic comprehension deficits in language will be related to harmonic processing deficits in music' (Patel, 2003, p. 679). In two experiments, Broca's aphasics displayed problems with linguistic, structural processing (identifying a number mismatch between verb and subject), and with harmonic, structural processing (identifying a harmonic mismatch between an out-of-key chord and the preceding context, and reacting faster to harmonically expected versus unexpected chords) (Patel et al., 2008). Similar results were obtained by Sammler, Koelsch, & Friederici (2011), who did not focus on patients with Broca's aphasia but instead on patients with a lesion in Broca's area, and who also included EEG evidence for unusual harmonic processing in this patient group. Apparently, in contrast to the global statements of neuropsychologists in the past (Peretz & Coltheart, 2003), a targeted investigation of harmonic processing in brain lesioned participants reveals a specific deficit in musical, structural processing which correlates with a linguistic, structural processing deficit. Thus, lesioning syntactic integration resources has consequences for both music and language abilities, as predicted by the SSIRH.

A second prediction concerns concurrently taxing structural integration resources through music and language. The SSIRH predicts that 'tasks which combine linguistic and musical syntactic integration will show interference between the two' (Patel, 2003, p. 679). A number of behavioural, EEG, and MEG studies found support for this prediction (Carrus, Koelsch, & Bhattacharya, 2011; Carrus, Pearce, & Bhattacharya, 2013; Fedorenko, Patel, Casasanto, Winawer, & Gibson, 2009; Hoch, Poulin-Charronnat, & Tillmann, 2011; Koelsch, Gunter, Wittfoth, & Sammler, 2005; Kunert & Slevc, 2015; Maidhof & Koelsch, 2011; Slevc, Rosenberg, & Patel, 2009; Steinbeis & Koelsch, 2008). For example, Slevc et al. (2009) asked participants to read sentences as in (1|2) word by word while their reading time was recorded. When each word was shown on the screen they also heard a chord. The reading time during the reading of a disambiguating word (e.g., *<u>was</u>* in 1|2) is longer in a garden-path sentences which encourage an initial syntactic misinterpretation (1|2a) than in a non-garden-path sentence (1|2b). This so called garden-path effect in reading time was greater if the accompanying chord was out-of-key compared to in-key. Apparently, hearing an out-of-key chord indeed taxes shared music-language syntactic integration resources which, as a consequence, can no longer process linguistic syntax optimally, as revealed by reading speed. The research in this thesis employed similar interference designs in chapters 2, 3, and 4.

In order to provide precise neural predictions from the SSIRH, the research in this thesis combines the SSIRH with another, very similar, dual-system model which is language-specific, the Memory-Unification-Control model (MUC model; Hagoort, 2005, 2013). The MUC model proposes that language processing requires three components: a memory component (akin to the storage component in the SSIRH), a unification component (akin to the syntactic integration component in the SSIRH), and a control component (relates language to action, no equivalent in the SSIRH). The memory component is thought to be mainly stored in temporal lobe structures while the left inferior frontal gyrus (IFG), where Broca's area is situated, is thought of as the unification space.

Support for a distinction between syntactic integration in the left IFG and memory operations in the left temporal lobe comes from a study by Snijders et al. (2009). Participants lay in the fMRI scanner and read sentences including a word-class ambiguous word which could be interpreted as a verb or a noun (e.g., *watch* in English), or an unambiguous sentence including only unambiguous words. Moreover, they also read word lists without syntactic structure including ambiguous and matched unambiguous words. As expected, word lists induced less activity in the left IFG than sentences, suggesting that syntactic integration indeed only takes place in the presence of the syntactic structure of sentences. Moreover, sentences with word-class ambiguous words induced more activity than sentences without, suggesting that the additional syntactic operations required to disambiguate the critical word indeed tax integration resources in the left IFG. Memory operations appear to engage the left posterior middle temporal gyrus (MTG) which was more active for ambiguous words than unambiguous words (marginally so in the case of word lists), suggesting that recruiting two syntactic frames for a word (noun interpretation and verb interpretation) taxes memory operations more than recruiting just one frame.

A neurologically more explicit version of the SSIRH augmented with predictions from the MUC model is shown in Figure 1|5. The evaluation of this combined model requires a range of methods which will be introduced next.



Role	Syntactic integration	Storage of structural representations
Location	Left IFG	Left MTG or unknown
Shared between music and language?	Partly	No

Figure 1|5. The shared syntactic integration resource hypothesis (SSIRH) augmented with neuronal predictions from a very similar language-specific model (the Memory-Unification-Control, MUC model). Regions of interest derived from chapter 5. IFG = inferior frontal gyrus; MTG = middle temporal gyrus

1|4 Methods for investigating the brain basis of structural processing in music and language

In the present thesis, three methods will be employed in order to investigate the predictions by the SSIRH: behavioural responses, fMRI, and MEG. All come with advantages and disadvantages which are important to understand before turning to the experimental chapters. Behavioural methods are perhaps the most straight-forward as well as the most neglected approach to answering cognitive neuroscientific questions (e.g., Kunert & Slevc, 2015). There are two main advantages to using behaviour to measure brain responses: 1) the measurement is cheap allowing for many participants and, thus, great statistical power, 2) the measurement is behaviourally relevant (by definition). The latter point simply stresses that people usually care about the behavioural consequences of brain activity rather than subtle changes in the brain which do not have any consequence outside of that organ. As mentioned before, the SSIRH makes behavioural predictions concerning interference effects between music and language.

However, behavioural measurements also come with limitations. The source of a behavioural response is neural activity in the primary motor cortex. Any more detailed claims about the neurological origin of behavioural differences between experimental conditions are very difficult to make. In principal, differences in behaviour could reflect activity differences anywhere in the brain resulting in changed motor commands emanating from the primary motor cortex. This lack of neural precision allows for only the broadest neural claims. For example, while it is very difficult to claim shared music-language resources from fMRI results (Peretz, Vuvan, Lagrois, & Armony, 2015a), behavioural methods have shown that shared music-language resources must exist *in some way* in the brain (Kunert & Slevc, 2015). Chapter 2 and chapter 7 focus on behavioural responses.

Perhaps the most popular method employed by cognitive neuroscientists is fMRI. With this method, the participant's head is placed inside a strong magnetic field (1.5 and 3 Tesla in the research reported in this thesis) and radio-frequency pulses are used to disturb the magnetic field. Through signals gathered by sensors around the head, a three-dimensional image of the brain can be re-created from the MRI signal. Oxygenated and deoxygenated blood have different magnetic properties, allowing for the identification of more active, i.e. more oxygenated, brain areas related to an experimental condition of a task performed by the participant during MRI scanning, see Figure 1|6.

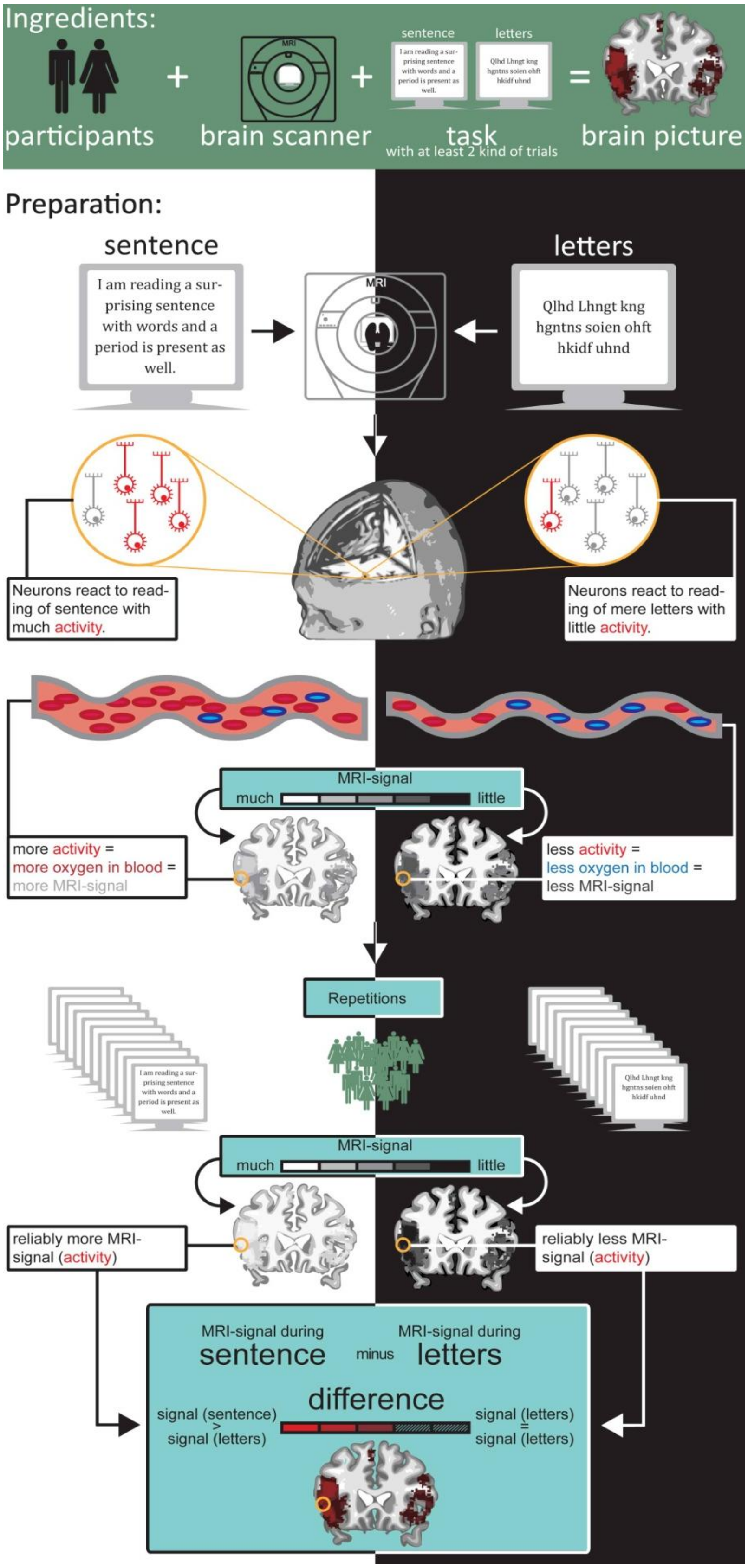

Ingredients:
participants
+
brain scanner
+
sentence
letters
task
with at least 2 kind of trials
=
brain picture
Preparation:
sentence
I am reading a surprising sentence with words and a period is present as well.
MRI
letters
Qlhd Lhngt kng hgntns soien ohft hkidf uhnd
Neurons react to reading of sentence with much activity.
Neurons react to reading of mere letters with little activity.
MRI-signal
much
little
more activity = more oxygen in blood = more MRI-signal
less activity = less oxygen in blood = less MRI-signal
Repetitions
MRI-signal
much
little
reliably more MRI-signal (activity)
reliably less MRI-signal (activity)
MRI-signal during sentence
minus
MRI-signal during letters
difference
signal (sentence) > signal (letters)
signal (letters) = signal (letters)

Figure 1|6. Infographic of how the brain activity pictures of a typical fMRI experiment are derived. The hypothetical experiment contrasting the reading of sentences and letters is not part of this thesis and only serves illustration purposes.

There are two main advantages to using fMRI: 1) potentially all cells in the brain contribute to the measured changes in blood-oxygenation, suggesting that the location of a cell is irrelevant to its fMRI detectability (though proximity to blood vessels matters), 2) compared to other techniques (behavioural, EEG, MEG), fMRI offers excellent spatial resolutions of about 3 mm^3 for one unit of measurement (a voxel). However, the limitations of fMRI should not be underestimated. First, the temporal resolution of fMRI is limited by the time it takes to take one MRI image (the repetition time, TR) and by the sluggishness of the blood-oxygenation level dependent (BOLD) response which takes about 5 seconds to peak after an event of interest. Moreover, not all neural responses to a stimulus necessarily result in MRI signal changes. For example, if neural activity in a voxel changes in terms of its synchronization between neurons but not in terms of its overall number of spikes, this change will be invisible to fMRI researchers because no change in blood-oxygenation accompanies this neural reaction. Besides, MRI scanners are very noisy, rendering auditory stimulus presentation difficult. On top of that, MRI measurement equipment is very costly, limiting the number of participants and, thus, statistical power. Chapter 3 and chapter 5 use MRI measurements.

The third technique employed in this thesis is magneto-encephalography (MEG). It is based on the electro-magnetic brain response during different conditions of a task the participant performs. Once a neuron gets excited by another neuron, electrical current flows from a neuron's dendrite to the soma and on along the axon. MEG is mainly sensitive to postsynaptic dendritic currents because of the alignment of dendrites into columns in the neocortex. This alignment allows for the summation of electromagnetic signals of thousands of neurons and the detection of this summed signal by MEG sensors. It is important to realize that MEG experiments can only make claims about a subset of brain cells. There are three main reasons for this limitation. First, neurons arranged in a way which does not allow for the summation of their electro-magnetic responses are not detectable by MEG sensors. Second, the magnetic signal drops off very quickly (with cubed distance) meaning that neuronal activity deep inside the brain is equally not noticeable by MEG sensors. Third, the magnetic field is generated at a right angle to the electric current along the dendrite (Fleming’s right hand rule). Therefore, the magnetic field of neurons positioned radially to the cortical surface (e.g., at the crest of gyri) cannot be measured with MEG. Thus, apart from a cost roughly comparable to MRI systems limiting sample sizes, the main disadvantage of MEG is that it only detects about half of the brain activity (Hillebrand & Barnes, 2002).

However, MEG comes with important advantages over other methods. It offers excellent temporal resolution on the order of milliseconds which is on the same temporal order as neuronal action potentials. Moreover, while its spatial resolution is not as good as for fMRI, it is better than for other, comparable techniques such as EEG. Finally, given that MEG scanners are silent they offer an ideal testing environment for investigating brain reactions to auditory stimuli. The research reported in chapter 4 is based on MEG.

It should be apparent now that each technique used by cognitive neuroscientists comes with advantages and disadvantages. Therefore, the research in this thesis employs a variety of research techniques in order to evaluate the SSIRH in multiple ways. This way, the advantages of each technique can be optimally exploited for answering specific questions about how music and language comprehension is implemented in the brain.

1|5 Overview of the thesis

The main goal is to characterize how music and language comprehension relate to each other in the brain. For this purpose the following chapters will evaluate the SSIRH which proposes that shared syntactic integration resources are involved in language and music processing. These resource’s functional characteristics, location, time-course, and relation to domain-specific resources are the central areas of investigation of this thesis.

In chapter 2 the focus lies on the functional characteristics of shared music-language resources involved in syntactic integration. In two behavioural experiments the prediction that challenging syntax in language should have an effect on concurrent music processing will be investigated. In this way, chapter 2 tests the interference prediction usually associated with impaired language syntax processing during challenging harmonic integration in a novel way

(Fedorenko et al., 2009; Hoch et al., 2011; Slevc et al., 2009). Are shared music-language resources really involved in music perception?

In chapter 3 the focus shifts to a spatial prediction by the SSIRH. The model predicts that shared syntactic integration resources are located in frontal brain areas. In combination with the MUC model this prediction can be made more specific to ask whether Broca's area displays a response pattern typical for shared syntactic integration resources. If these resources are located where syntactic unification processes have been claimed in many previous studies, then this is direct support for the SSIRH which claims a syntactic role for shared music-language resources. Chapter 3 reports on an fMRI study using an interference design to test this prediction.

In chapter 4 the focus shifts from the location to the time-course of music-language resources. Different ERPs have been associated with syntactic processing and harmonic processing (Koelsch, Gunter, et al., 2005; Maidhof & Koelsch, 2011; Steinbeis & Koelsch, 2008). Moreover, the oscillatory dynamics of music-language interactions have only once before been investigated (Carrus et al., 2011). Therefore, chapter 4 reports on an MEG study very similar to the fMRI study in chapter 3 which asks what the temporal dynamics of shared music-language resources are.

In contrast to chapters 2, 3, and 4, chapter 5 does not investigate shared music-language resources. Instead, it is an fMRI study evaluating the non-shared component of music and language comprehension. The SSIRH predicts that the memory storage of music and language representations together with their structural properties occurs in distinct posterior brain areas. This prediction is evaluated in chapter 5.

Chapter 6 is thematically somewhat removed from chapters 2 to 5 because it does not evaluate the SSIRH. It presents the evaluation of a musical aptitude measure (Law & Zentner, 2012). Such research is important for possible extensions of the work presented in chapters 2 to 5 which wants to take individual differences in musical skills into account. The methodological question chapter 6 asks is whether one particular musical aptitude measure, the Profile of Music Perception Skills (PROMS), has good psychometric properties, i.e. whether it is a good research tool.

Finally, chapter 7 summarizes the findings and discusses future directions for investigating music and language comprehension.

Chapter 2

Language influences music harmony perception: effects of shared syntactic integration resources beyond attention

Many studies have revealed shared music-language processing resources by finding an influence of music harmony manipulations on concurrent language processing. However, the nature of the shared resources has remained ambiguous. They have been argued to be syntax specific and thus due to shared syntactic integration resources. An alternative view regards them as related to general attention and, thus, not specific to syntax. The present experiments evaluated these accounts by investigating the influence of language on music. Participants were asked to provide closure judgements on harmonic sequences in order to assess the appropriateness of sequence endings. At the same time participants read syntactic garden-path sentences. Closure judgements revealed a change in harmonic processing as the result of reading a syntactically challenging word. We found no influence of an arithmetic control manipulation (experiment 1) or semantic garden-path sentences (experiment 2). Our results provide behavioural evidence for a specific influence of linguistic syntax processing on musical harmony judgements. A closer look reveals that the shared resources appear to be needed to hold a harmonic key online in some form of syntactic working memory or unification workspace related to the integration of chords and words. Overall, our results support the syntax specificity of shared music-language processing resources.

Kunert, R., Willems, R.M., Hagoort, P. (2016). Language influences music perception - effects of shared syntactic integration resources beyond attention. *Royal Society Open Science, 3*, 150685. doi: 10.1098/rsos.150685

2|1 Introduction

2|1.1 Syntax versus attention explanations for music-language interactions

Music and language are universal human faculties (Brown & Jordania, 2013). Moreover, both of these cognitive domains are based on structured auditory sequences, i.e. they contain discrete elements (e.g., words in language, tones/chords in music) which relate to each other in a rule-governed way in order to form higher order structures (e.g., sentences in language and harmonic sequences in music). This similarity has led researchers to hypothesise that the same processing resources underlie both domains. Supporting evidence comes from a host of experiments which found music harmony manipulations to change language syntax processing (Fedorenko et al., 2009; Hoch et al., 2011; Koelsch, Gunter, et al., 2005; Kunert & Slevc, 2015; Kunert, Willems, Casasanto, Patel, & Hagoort, 2015; Slevc et al., 2009; Steinbeis & Koelsch, 2008). However, the nature of the shared resources which lead to these music-language interactions has remained controversial. On the one hand, they are argued to be specific to music and language syntax, i.e. they are the result of shared syntactic processing resources (Patel, 2008). On the other hand, these effects are thought to be an instance of many other, similar, interactions including those between music harmony and language semantics (Perruchet & Poulin-Charronnat, 2013; Poulin-Charronnat, Bigand, Madurell, & Peereman, 2005). In this view, music-language interactions arise due to the effect of music manipulations on general attention. In this paper, a previously little explored direction of influence will be investigated: the linguistic influence on music harmony perception. By doing so, we directly test predictions from a syntax account versus those from general attention accounts.

2|1.2 Music-language interactions as explained by a syntax account

Patel’s (2008) shared syntactic integration resource hypothesis (SSIRH) argues that music and language share structural processing resources. It was designed to reconcile neuropsychological evidence for domain independence (Peretz & Coltheart, 2003) with neuroimaging studies which found evidence for shared resources in terms of similar activation sites and time-courses for linguistic syntactic processing and tonal harmonic processing (Koelsch et al., 2002; Patel et al., 1998). Both kinds of findings together were thought to support a distinction between representations in long-term memory (domain specific, explaining neuropsychological double-dissociations) and resources for online syntactic integration (domain general). Similar activation effects reflect domain-general syntactic integration resources which in turn draw on domain-specific representations.

A key prediction of Patel's (2008) SSIRH is an interaction between music and language syntax processing when both occur at the same time. This prediction has been supported by two behavioural studies finding impaired syntactic integration abilities in language if a concurrent musical tone or chord is difficult to harmonically integrate (Fedorenko et al., 2009; Slevc et al., 2009; see also Hoch et al., 2011; Kunert & Slevc, 2015; Kunert et al., 2015). For example, Slevc et al. (2009) presented participants with syntactic garden-path sentences of the form below.

a) After | the trial | the attorney | advised | **the defendant** | <u>was</u> | likely | to commit | more crimes.

b) After | the trial | the attorney | advised that | **the defendant** | <u>was</u> | likely | to commit | more crimes.

Participants reading the sentence in a) are likely to, at first, misinterpret '**the defendant**' as being advised. Upon encountering the verb '<u>was</u>', readers realize that someone is being advised about '**the defendant**', a syntactic reanalysis which taxes integration resources. In b) the correct reading is enforced early on when reading 'that'. As a consequence, reading times on disambiguating words such as '<u>was</u>' are typically longer in sentences like a) compared to b). Crucially, Slevc et al. (5) found that this syntactic garden-path effect is intensified by a harmonically unexpected (out-of-key) chord presented concurrently with the disambiguating word ('<u>was</u>'). A timbral unexpectancy (an unusual instrument) was without effect. The authors interpreted this as evidence against acoustic deviancy alone underlying the musical influence on language. The SSIRH explains this pattern by assuming that an out-of-key chord is difficult to integrate into the prevailing key, i.e. it taxes resources involved in syntactic integration. This leaves fewer resources available for concurrently reanalyzing the syntactic structure of the sentence, hence the increased linguistic garden-path effect. An unexpected instrument, on the other hand, does not tax integration resources, and hence has no effect.

2|1.3 Music-language interactions as explained by attention accounts

An alternative explanation for the influence of music on reading times is based on general attention mechanisms. These can take two forms. One proposal is an attentional load account. For example, Perruchet and Poulin-Charonnat (2013) found a semantic garden-path effect when a harmonically unexpected (out-of-key) chord was presented concurrently with the disambiguating word. No semantic garden-path effect was found when the chord was expected (in-key) instead. They found no influence of the music harmony manipulation on semantic error processing. The results were argued to support an attentional load account, i.e. one domain (e.g., music) influences another (e.g., language) only if enough attentional resources are left to process both. Semantic garden-path processing is thought to allow for the distribution of attention across domains, while semantic error processing does not. Crucially, the key prediction of this account for the present investigation is that the nature of the processing difficulty, e.g., language syntax or something else, is irrelevant, placing it in sharp contrast to the syntax account by Patel (2008).

A different attentional account was put forward by Poulin-Charronnat and colleagues (2005) to explain an effect of music harmony on the semantic priming effect seen in lexical decisions. Participants were presented with sung sentences ending either on a word or a non-word and participants were asked to decide on the lexical status of the final item in the sentence. Words could be either semantically expected or not. The semantic expectancy effect was larger when the final word was sung on pitches forming an expected (tonic) rather than a less expected (subdominant) final chord. This result was taken to reflect attentional entrainment, as theorised in Jones' dynamic attending theory (Jones & Boltz, 1989; Large & Jones, 1999). According to this account, attentional fluctuations entrain to harmonic accents such as expected chords. This leads to an attentional peak when a tonic is heard. Greater attention at this point in time facilitates linguistic processing, hence the increased semantic priming effect. The same account has also been suggested to explain the music effect on visual processing (Escoffier & Tillmann, 2008, p. 200) and phoneme monitoring (Bigand, Tillmann, Poulin, D'Adamo, & Madurell, 2001). Crucially, the nature of the stimulus is irrelevant in this account as long as it leads to attentional entrainment. This – again – is in sharp contrast to Patel's (2008) account which suggests that only syntactic manipulations in language will affect music processing.

2|1.4 The present paper: can language influence music perception? If so, how is music perception affected?

All the studies mentioned so far investigated musical influences on language or other tasks. We decided to reverse the direction of influence, i.e. we focus on linguistic influences on music processing. People listened to music while reading sentences or control arithmetic problems. The predictions of the syntax and the attention accounts outlined above are quite clear. Patel's (2008) SSIRH predicts that only syntactic manipulations will affect music harmony processing. Specifically, a challenging language syntax condition [ambiguous S-coordinations (Frazier, 1987; Hoeks, Hendriks, Vonk, Brown, & Hagoort, 2006; Hoeks, Vonk, & Schriefers, 2002) in experiment 1 and object-relative clauses (Fedorenko et al., 2009) in experiment 2] should tax musico-linguistic integration resources, leading to an impaired ability to harmonically integrate chords into an unfolding sequence. Non-syntactic processing challenges [difficult arithmetic operations in experiment 1 and semantic garden-path sentences (Perruchet & Poulin-Charronnat, 2013) in experiment 2] should be without effect. The aforementioned attentional accounts, on the other hand, predict a non-specific effect, i.e. an effect for both the syntactic manipulations and the control manipulations.

Either way, the influence of a concurrent task on music perception could take two different forms since there are two aspects to harmony processing. First, the listener has to integrate chords in order to establish a harmonic key. Second he/she must integrate chords with an already established key which is held online in some form of unification space (see section 2|5.3 for details). We investigated these processes by presenting the challenging moment in concurrent tasks while the music piece modulated from an established harmonic key to a new key, i.e. when an old key has to be kept online without being reinforced by chords and a new key has to be established from the incoming chords. Chord sequences ended either on an authentic cadence typical for the first established key (probing whether the old key could be held online after it was no longer reinforced) or the second established key (probing whether a new key could be established after a critical moment in the concurrent task). The measure of interest which provides us a window into these processes is found in closure ratings. Participants are simply asked to rate their feeling of completeness (closure) regarding the music (see Bigand & Pineau, 1997). Closure is high if the last chords can be integrated into an available key. Reduced closure ratings indicate difficulty with harmonic integration - as expected in the case of syntactically challenging sentences compared to easier sentences.

2|2 Experiment 1 (exploratory): language syntax versus arithmetic difficulty

2|2.1 Method

2|2.1.1 Participants

Fifty-eight participants were invited to take part in the experiment. One participant did not advance to the experiment as she did not understand the musical task after repeated chance-level practice trial performance. Three participants were rejected due to an error in counter-balancing[1]. This leads to a final sample of 54 participants (19 men, 44 right handed). They were all Dutch native speakers, aged 23 on average (*SD* = 6.3), with 3.9 years of musical training on average (*SD* = 3.7). Thirty were self-described non-musicians, nineteen amateur musicians, five semi-professional musicians. They were paid 16 € or undergraduate course credits for their participation and were naive as to the purpose of the experiment.

2|2.1.2 Design and Material

We employed a 2 (Key: first key ending or second key ending) × 3 (Difficulty: ambiguous S-coordination, unambiguous S-coordination or NP-coordination) design for the musico-linguistic part of the experiment and a 2 (Key: first key ending or second key ending) × 2 (Difficulty: hard, easy) design for the musico-arithmetical part. All factors were manipulated within-subjects. All critical stimuli can be found in the appendices. Sound files are available as supplementary materials. The task paired musical stimuli (harmonic sequences requiring a

---

[1] Including these three participants does not alter the findings of the study. Notably, the crucial Difficulty main effect in the musico-linguistic part of the experiment was still significant [$F_{(1,56)} = 12.59$, $p = .001$, ${}_{p}\eta^2 = .184$], as was the Task × Difficulty interaction when including Task as a factor [$F_{(1, 56)} = 5.50$, $p = .023$, ${}_{p}\eta^2 = .089$].

closure rating) with linguistic stimuli (visually presented sentences requiring a comprehension answer) or arithmetic stimuli (visually presented arithmetic formulae requiring a solution judgment). Auditory music stimuli were presented concurrently with visual language or arithmetic. The critical point in time was always the ninth position (highlighted or underlined in the examples below: Figure 2|1, and examples 2|1 and 2|2), corresponding to the start of a new key in the musical material, the garden-path disambiguation in the ambiguous S-coordination sentences of the language stimuli and a difficult operation in the hard arithmetic trials.

2|2.1.2.1 Music stimuli

The music stimuli consisted of chord sequences specifically composed for this experiment by the first author. As shown in Figure 2|1, 10 items were composed beginning in C-major (seven chords), followed by a pivot chord which is part of both the C-major and B-flat-major keys (F-major chord or d-minor chord), followed by four chords in the B-flat-major key, ending with two chords forming an authentic cadence in C-major (first key ending) or B-flat-major (second key ending). These two versions of the ten items were transposed twice (first key = D-major & second key = C-major; first key = B-flat major & second key = A-flat major). This resulted in 60 critical chord sequences (10 items × 2 endings × 3 transpositions). Next to the critical items, filler items in only one key (e.g., C-major) were constructed. For example, the second-key-region was transposed into C-major and a first-key-typical ending (C-major cadence) was chosen. Half of the filler sequences ended in an authentic cadence (requiring high closure ratings), half did not (dominant followed by supertonic or subdominant, requiring low closure ratings). This resulted in 60 filler chord sequences which allowed us to check whether participants paid attention to the musical task. Overall, participants were exposed to as many modulating (critical) sequences as non-modulating (filler) sequences. All chord sequences were played by a piano at a tempo of 96 bpm, consisting only of crotchets (quarter notes) except for the last chord which was made up of dotted minims (three-quarter notes).

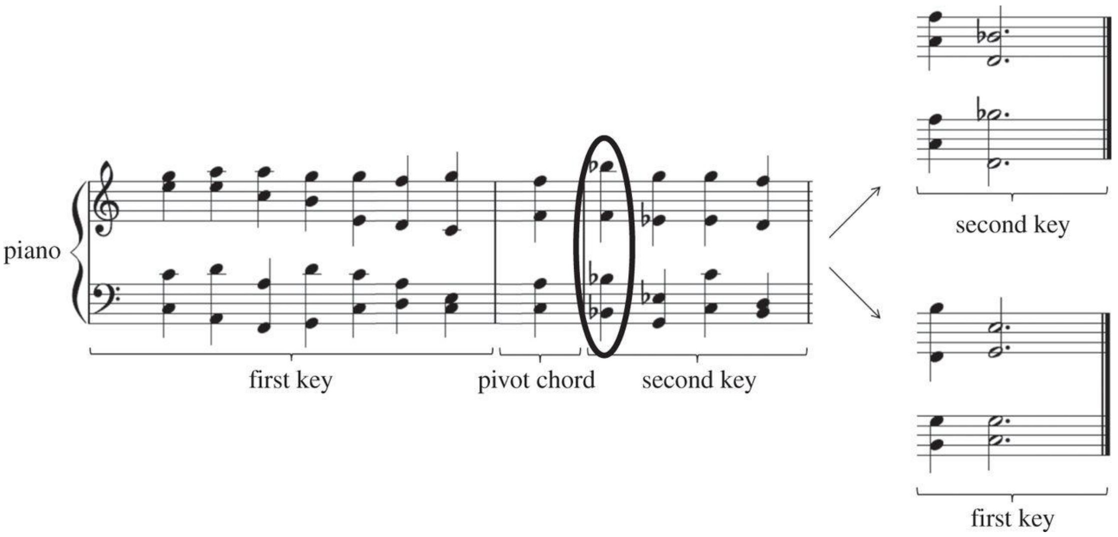


Figure 2|1. A sample music item in two versions. The top version ends in the second key (B-flat-major), the bottom version ends in the first key (C-major). The pivot chord (F-major chord) is part of both keys. The ninth chord (coinciding with the critical word in the concurrent language task and a manipulated operation in the concurrent arithmetic task) is encircled. From this chord onwards the first key has to be kept online without being reinforced by incoming chords (as tested by ratings of first key endings), and the second key has to be built up from new chords (as tested by ratings of second key endings). Bar lines denote boundaries between sections belonging to different harmonic keys.

2|2.1.2.2 Language syntax stimuli

The critical language stimuli were a modified version of a stimulus set developed by Kerkhofs, Vonk, Schriefers, & Chwilla (2008). As can be seen in example 1, 60 critical items were constructed in three versions each: an ambiguous S-coordination (garden-path version), an unambiguous S-coordination (intermediate difficulty version with a disambiguating comma before '**en**', ***and***) and an NP-coordination (non-garden-path version). They were 14 words long. Pre-test results (see section 2|1.3) revealed that the unambiguous S-coordination sentences were not perceived as significantly easier than the ambiguous S-coordination sentences. Therefore, we will concentrate on the contrast between the challenging syntax condition (ambiguous S-coordination) and the easier syntax condition (NP-coordination) here, see supplementary materials for an analysis of critical trials involving unambiguous S-coordination sentences.

(2|1) Language example

a) ambiguous sentence-coordination (S-coordination, garden-path condition)

De | chirurg | troostte | de | man | **en** | de | vrouw | <u>legde</u> | haar | hand | op | zijn | voorhoofd.

Translation: *The surgeon consoled the man **and** the woman <u>put</u> her hand on his forehead.*

b) noun-phrase-coordination (NP-coordination, non-garden-path condition)

De | chirurg | troostte | de | man | **en** | de | vrouw | <u>omdat</u> | de | operatie | niet | gelukt | was.

Translation: *The surgeon consoled the man **and** the woman <u>because</u> the operation had not been successful.*

The beginning of each sentence (‘De chirurg troostte de man …’, *The surgeon consoled the man…*) was followed by ‘**en**’ (***and***) and a two-word noun phrase (‘de vrouw’, *the woman*), followed by a six-word ending of the sentence. This ending began either with a verb (ambiguous S-coordinations, ‘…<u>legde</u> haar hand op zijn voorhoofd.’, …*<u>put</u> her hand on his forehead.*) or without (NP-coordination, ‘…<u>omdat</u> de operatie niet gelukt was.’, …*<u>because</u> the operation had not been successful.*). Thus, these two stimulus versions only differed by the sentence ending.

The 60 filler items were also 14 words long and were made up of diverse syntactic constructions. 20 of the 60 filler stimuli included a NP-coordination, ensuring that overall participants read as many S-coordination sentences as NP-coordination sentences. Comprehension prompts of critical trials targeted the S- or NP-coordination ambiguity (e.g., for example 2|1): ‘De chirurg troostte alleen de man.’ *The surgeon only consoled the man.* - True for S-coordination, false for NP-coordination.). Filler comprehension prompts targeted various aspects of the filler sentence. Half the prompts required a ‘matches the sentence’ response, half required a ‘no match’ response.

2|2.1.2.3 Arithmetic stimuli

The arithmetic stimuli were made up of seven numbers and seven operators each. Only additions and subtractions involving numbers equal or below 10 were used after the first number. Interim solutions never exceeded 21 and were always positive integers. Arithmetic operations were classified as easy, hard, or intermediate. Easy operations were additions or subtractions of the numbers one, two or ten. Hard operations did not respect a ten-boundary, e.g. the addition of two numbers smaller than ten whose sum is greater than ten. All other operations were seen as of intermediate difficulty. Note that the time pressure of rapid serial visual presentation meant that the arithmetic task was not too easy as shown by an accuracy of 84% on average.

There were 40 critical items, thus ensuring an equal number of arithmetic and linguistic critical trials in each Difficulty × Key-ending design cell (10 trials). As shown in example 2|2, their hard and easy versions were identical with regards to the first eight positions as well as the final two and the solution. The two stimulus versions differed according to the operations required at positions nine (underlined in example 2|2) to twelve. At the ninth position, in the hard version, the task required a difficult arithmetic operation, while in the easy version the operation was easy. In both hard and easy stimuli the operations after position twelve were easy. The 40 filler items were of the same length and only included easy and intermediate operations. Akin to comprehension prompts in the language task, prompts in the arithmetic task either matched the true solution (50% of trials) or differed from it by 1 (50% of trials).

(2|2) Arithmetic example

a) hard

20 | - | 1 | - | 2 | - | 1 | - | 7 | - | 2 | + | 1 | =

b) easy

20 | - | 1 | - | 2 | - | 1 | - | <u>10</u> | + | 1 | + | 1 | =

2|2.1.3 Pre-test: the strength of the language syntax and arithmetic manipulations

A pre-test ($N = 24$) showed that the difficulty manipulations in the language syntax task and the arithmetic task were both noticeable as revealed by trial difficulty ratings (Difficulty main effect; language: $p < .001$; arithmetic: $p < .001$; see supplementary material). As mentioned before, ambiguous and unambiguous S-coordination sentences were not perceived as significantly different [$t_{(23)} < 1$], suggesting the contrast ambiguous S-coordination versus NP-coordination will be more powerful in revealing linguistic influences on music processing. Therefore, we will focus on this contrast. Overall, the arithmetic manipulation was more salient (Manipulation × Difficulty interaction, $p = .028$). This suggests that effects based on general task difficulty should be more easily visible in the arithmetic task. Thus, if there is an effect of language on music and it is due to general task difficulty, then the arithmetic control task should show an effect as well.

2|2.1.4 Procedure

Participants were instructed to perform two tasks simultaneously: an auditory music task presented concurrently with a language task or an arithmetic task; see Figure 2|2. The type of concurrent task (language or arithmetic) was blocked and the order counter-balanced. An experimental session was organized as follows. Participants first completed an experimental block. Thereafter, they completed a musical background questionnaire and a working memory test (digit span, Groth-Marnat, 2001), followed by another experimental block. A testing session took approximately two hours.

In the music task participants had to judge the closure, i.e. feeling of completeness, of a chord sequence on a seven-point Likert scale (1 = low closure, 7 = high closure). Closure judgments were required after each chord sequence. Participants were asked to answer a language/arithmetic comprehension prompt which followed one third of the closure judgments (randomly chosen). Given the limited interest in the language/arithmetic task behaviour itself and in order to save time, we only included comprehension prompts on catch trials in order to ensure task adherence. Their unpredictable occurrence meant that the linguistic/arithmetic task had to be carried out throughout the experiment. Participants received feedback on their language/arithmetic accuracy every 20 trials.

In the musico-linguistic part of the experiment participants completed 120 trials, 40 of which were critical trials containing a garden-path (ambiguous S-coordination) or non-garden-path sentence (NP-coordination) in the language task and a harmonic sequence with modulation in the music task. Critical linguistic items were randomly paired with critical musical sequences.

Each language item was only shown in one version to each participant. The choice of item version was counterbalanced. Specifically, language item 1 might be presented in language condition A to participant 1, condition B to participant 2 and condition C to participant 3. For these three participants, the position of the item in the experiment was the same. Furthermore, the language item was paired with the same music stimulus for these three participants. For the next three participants, this item had a different position (determined by chance) and was matched with a different music stimulus (determined by chance). This way, a particular music-language item match could have no systematic influence on the results and the position in the experiment is counter-balanced for different language conditions. Each music item was presented once in every version to each participant. Trials were pseudo-randomized with the following constraints: 1) at least 3 trials between different versions of a music item, 2) at most 3 filler or critical trials after each other, 3) at most three same answer conditions (prompt matching or not matching) after each other. The music stimuli were presented auditorily (608 ms per chord). The language stimuli were presented word by word (500 ms per word followed by 108 ms ISI) at the centre of the screen. The onset of a word presentation coincided with the onset of a chord.

In the musico-arithmetic part of the experiment participants completed 80 trials, 40 of which were critical trials containing a hard or an easy arithmetic problem in the arithmetic task and a harmonic sequence with modulation in the music task. The other half were filler trials. The combination of musical and arithmetic items, counter-balancing and trial randomization was done in the same way as described for the musico-linguistic task, except that only 80 musical sequences were used (original items and transposition down by two semitones, otherwise all versions of an item) because of the reduced trial count. The arithmetic stimuli were presented in a similar way as the language material (500 ms per number or operator followed by 108 ms ISI).

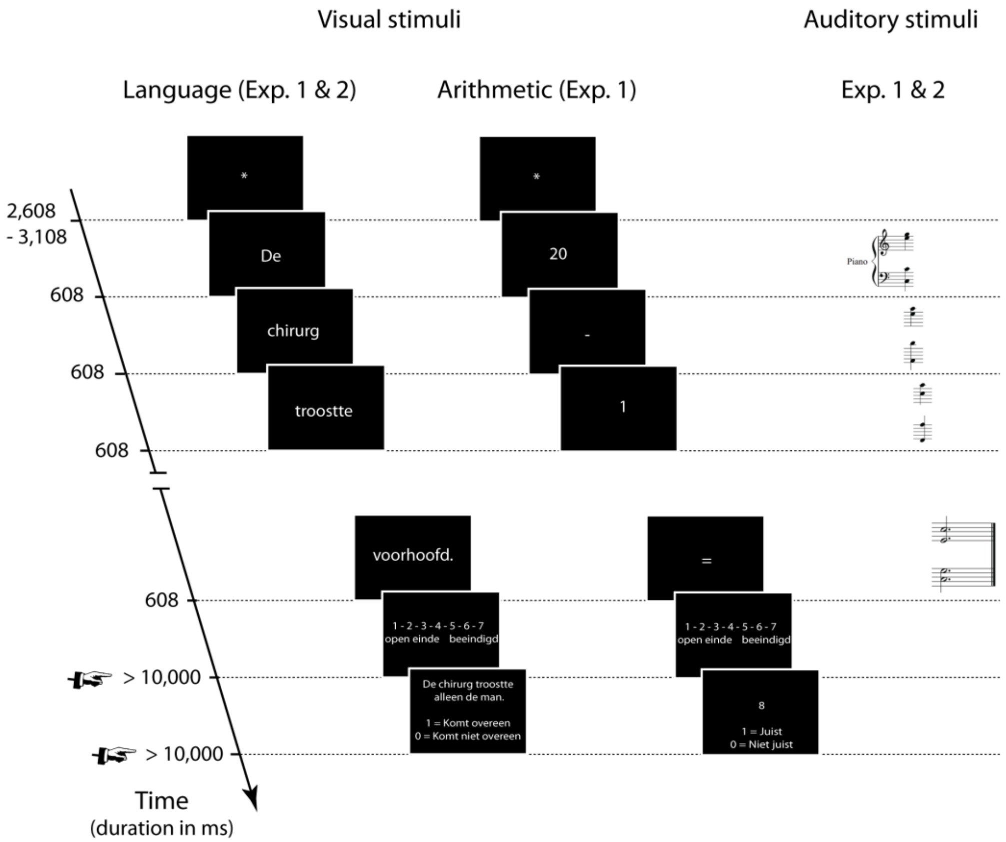


Figure 2|2. Stimulus sequence and timing in the visual and auditory modalities. On the left, a linguistic sequence (words from experiment 1) as used in experiments 1 and 2. In the middle, an arithmetic sequence as used in experiment 1. Note that comprehension prompts were presented only in one third of the trials. On the right, an example of the chord sequences which were presented concurrently with the visual stimuli. Note that closure judgements on the chord sequence were part of each trial.

2|2.1.5 Analysis

The closure ratings of critical trials were negatively skewed with a median rating of six (out of seven), i.e. our data displayed a ceiling effect likely because all critical sequences ended on an authentic cadence. Therefore, an analysis based on mean ratings appeared inappropriate as the mean would not be a good representation of the *typical* closure rating behaviour. Instead, we based our analysis on the median. Specifically, we applied a subject-specific median-split analysis on ratings. For each participant we coded ratings as either indicating high closure (subject-specific median of critical trial ratings or higher) or not. The analyses were performed on the proportion of trials in each condition which indicated high closure. This analysis strategy amplifies small differences in ratings between different concurrent task conditions while acknowledging that most critical trials are perceived as highly complete (high closure).

2|2.2 Results

2|2.2.1 General task performance

Before presenting the crucial results of the critical trials, we first present evidence for good task adherence. Concerning the music task, using the filler trials which ended either in an authentic cadence (requiring high closure ratings) or not (requiring low closure ratings) we found that all participants had an equal or greater proportion of high-closure ratings for the authentic cadence endings than the no-cadence endings ($M_{\text{authentic cadence}} = .81 > M_{\text{no cadence}} = .12$, difference $SD = .21$). Therefore, all participants appear to have rated filler trial music as expected by music theory.

Language task accuracy (including filler and critical trials) was high in general ($M = 85\%$, $SD = 8\%$). As expected, the critical trial accuracy showed a clear difference between challenging ambiguous S-coordination ($M = 83\%$, $SD = 16\%$) and less challenging NP-coordination trials ($M = 90\%$, $SD = 15\%$) [$t_{(53)} = 2.28$, $p = .027$, ${}_{p}\eta^2 = .089$]. Arithmetic task accuracy was similar in general ($M = 84\%$, $SD = 10\%$). As expected, the critical trial accuracy showed a clear difference between the hard ($M = 75\%$, $SD = 23\%$) and the easy arithmetic trials ($M = 92\%$, $SD = 10\%$) [$t_{(53)} = 5.25$, $p < .001$, ${}_{p}\eta^2 = .342$].

2|2.2 Critical task performance

2|2.2.1 Critical language results

The critical closure data of the musico-linguistic part of the experiment were analyzed in a 2 (Difficulty: ambiguous S-coordination, NP-coordination) × 2 (Key: first key ending, second key ending) within-subjects ANOVA, see Figure 2|3A. There was a Difficulty main effect [$F_{(1,53)}$ = 11.99, p = .001, $_p\eta^2$ = .184]. Otherwise, first key endings were less likely (M = .60, SD = .21) to receive a high closure rating than second-key endings (M = .77, SD = .14) [Key, $F_{(1,53)}$ = 51.74, p < .001, $_p\eta^2$ = .494]. These two factors did not interact [Key × Difficulty, $F_{(1,53)}$ = 1.80, p = .185, $_p\eta^2$ = .033]. Still, as shown in Figure 2|3A, the simple main effects revealed a Difficulty effect only for first key endings [$t_{(53)}$ = 3.28, p = .002, $_p\eta^2$ = .169], not for second key endings [$t_{(53)}$ = 1.66, p = .104, $_p\eta^2$ = .049].

2|2.2.2 Critical arithmetic results

The critical closure data of the musico-arithmetic part of the experiment were analyzed in a 2 (Difficulty: hard, easy) × 2 (Key: first key ending, second key ending) within-subjects ANOVA, see Figure 2|3B. There was no effect of arithmetical difficulty [$F_{(1,53)}$ < 1]. Otherwise, first key endings were less likely (M = .59, SD = .20) to receive a high closure rating than second-key endings (M = .68, SD = .18) [Key, $F_{(1,53)}$ = 13.69, p = .001, $_p\eta^2$ = .205]. These two factors did not interact [Key × Difficulty, $F_{(1,53)}$ < 1].

2|2.2.3 Language syntax vs. arithmetic

In order to see whether the difficulty effect in the musico-linguistic task was significantly different from the one in the musico-arithmetic task we performed a three-way within-subjects ANOVA on the harmonic closure ratings with the factors Task (language, arithmetic), Difficulty (ambiguous S-coordination/hard, NP-coordination/easy) and Key (first key ending, second key ending). Indeed, the difficulty effect was greater in the language task than in the arithmetic task [Task × Difficulty, $F_{(1,53)}$ = 5.20, p = .027, $_p\eta^2$ = .089]. Otherwise, all three factors showed a main effect [Task, $F_{(1,53)} = 5.52$, $p = .023$, ${}_{p}\eta^2 = .094$] [Difficulty, $F_{(1,53)} = 8.16$, $p = .006$, ${}_{p}\eta^2 = .133$] [Key, $F_{(1,53)} = 41.33$, $p < .001$, ${}_{p}\eta^2 = .438$]. The Key and Task factors interacted [$F_{(1,53)} = 8.83$, $p = .004$, ${}_{p}\eta^2 = .143$] while Key and Difficulty did not [$F_{(1,53)} < 1$]. The three-way interaction was marginally significant [Task × Difficulty × Key, $F_{(1,53)} = 3.47$, $p = .068$, ${}_{p}\eta^2 = .062$].

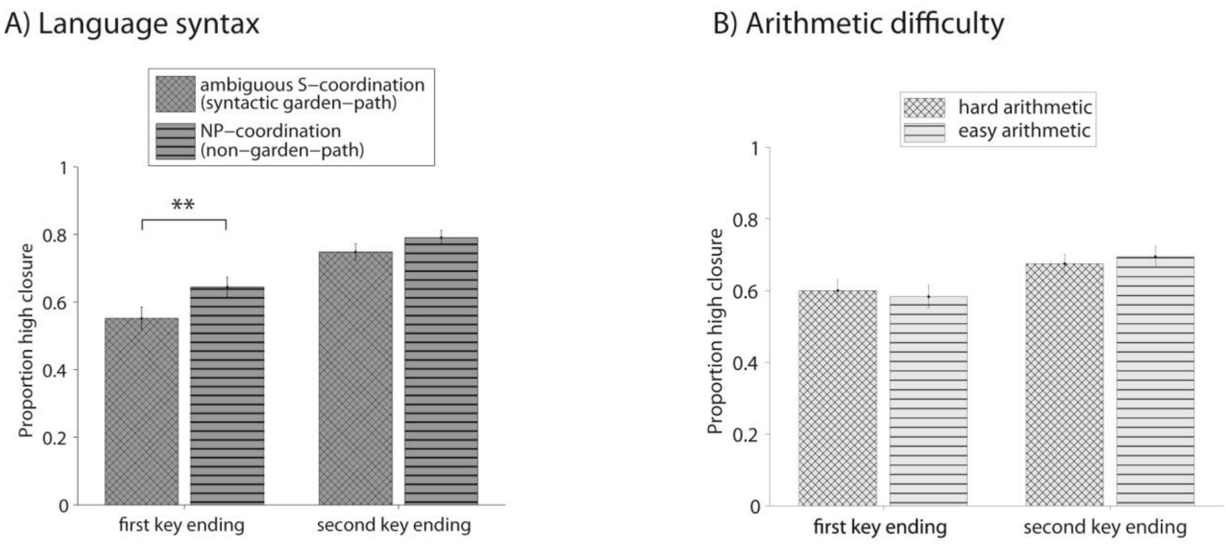


Figure 2|3. Experiment 1: Closure ratings of critical trials. Participants were asked to rate their feeling of closure, i.e. completeness (y-axis), of harmonic sequences ending either in a way typical for a first established key or in a second-key-typical way (x-axis). Different bars represent concurrent task conditions. **A)** In one block people solved the auditory music task while they were also asked to read sentences. We found an influence of language syntax on music harmony ratings. **B)** In a different block, people solved the auditory task while they also performed an arithmetic task. The arithmetic manipulation was without effect. Significance levels represented as stars refer to simple main effect *t*-tests and do not imply significant interaction effects. Error bars = SEM. ** p < .01

2|2.3 Discussion: language syntax affects harmony judgements, arithmetic does not

The results of experiment 1 show that a syntactic garden-path effect can indeed influence harmony perception by reducing the ability to integrate chords, as shown by lower closure ratings at the end of a harmonic sequence. An arithmetic control condition, without syntactic demands but with a more salient difficulty manipulation, did not affect harmony perception. This shows that general trial difficulty is not involved in the syntax effect. Instead, as predicted by Patel's (2008) SSIRH, a syntactic challenge in language probably taxed shared musico-linguistic syntactic integration resources which in turn can then only sub-optimally process chords. A non-specific effect of linguistic or arithmetic difficulty on music ratings, as predicted by attentional accounts (Jones & Boltz, 1989; Perruchet & Poulin-Charronnat, 2013), was not observed. Note that the syntax effect was not specific to either type of music ending, suggesting that harmonic processing in general was affected, rather than holding a key online or establishing a new key.

The observed main effect of key is not of interest in this study. The fact that second-key endings more often receive high closure ratings than first-key-endings likely reflects the relatively fast decline of key availability if a harmonic key is not reinforced by heard chords. Whether this is also the case without a concurrent task was investigated in a post-test which is part of the supplementary materials. Briefly, this post-test showed that no-cadence, first-key, and authentic cadence endings were rated very similarly without a concurrent language or arithmetic task. Second-key endings, however, less often received a high closure rating without a concurrent task compared to with, leading to a smaller difference between first-key and second-key ending ratings in this post-test. Still, the numerical trends were in the same direction, i.e. there is no indication in the post-test data that second-key endings are heard as less well closed compared to first-key endings. In sum, participants who only did the music task performed this task similarly to participants who did it while engaging in a second task at the same time.

However, a series of open questions remain concerning the effect of language on music ratings. Firstly, the difficulty in the language domain consisted in the syntactic re-analysis of a conjunction (**'en'**, *and*) held in memory. This sixth word either linked two sentences (S-coordination) or two noun-phrases (NP-coordination) depending on the ninth word in the sentence. The arithmetic control task required no such retrieval of an item from memory.

Therefore, it could be hypothesised that the different effects of linguistic and arithmetic difficulty are based on a working memory mechanism: one either has to retrieve a temporally distant element or not. Re-analysing a word held in memory because of an unexpected new word could be thought to impair the ability to simultaneously hold chords in working memory. This in turn might impair the ability to use them for harmonic integration. In order to evaluate this alternative scenario we included digit span as a co-variate in the analysis of the musico-linguistic task data. If working memory, rather than common syntactic integration resources, was responsible for the language effect, then the greater availability of working memory resources should lead to a reduction in the influence of language on music. An ANCOVA analysis with *z*-scored digit span as a measure of working memory did not support this scenario. Neither forward digit span, nor backward digit span, nor overall digit span significantly modulated any of the main effects or interactions (*p*s > .1; see supplementary material). This supports a functional distinction between syntactic processing and general working memory (Fiebach, Schlesewsky, Lohmann, von Cramon, & Friederici, 2005; Makuuchi, Bahlmann, Anwander, & Friederici, 2009; Martin, 1993; Waters, Caplan, Alpert, & Stanczak, 2003).

A second open question relates to the specificity of the language manipulation. The linguistic contrast chosen in experiment 1 did not control for semantic or lexical differences after the 8th word of the sentence. Experiment 2 was designed to directly evaluate the influence of language syntax and language semantics while tightly controlling lexical items. We included a new syntactic contrast which matches lexical items between conditions (object- and subject extracted relative clauses). The semantic manipulation consisted of semantic garden-path sentences (also called lexical garden-path sentences) in which a disambiguating word leads to the semantic re-interpretation of an ambiguous word held in memory. Note that semantic garden-path effects have previously been shown to be modulated by a music harmony manipulation (Perruchet & Poulin-Charronnat, 2013). This renders this particular semantic manipulation an ideal testing ground for the specificity of the syntactic effect encountered in experiment 1.

Finally, an additional reason for including a second experiment here is the post-hoc nature of the analysis strategy we chose in experiment 1, i.e. the decision to focus on the proportion of high-closure ratings rather than raw ratings due to an unexpected ceiling effect. As opposed to the exploratory nature of experiment 1, experiment 2 was purely confirmatory and used the same analysis strategy.

2|3 Experiment 2 (confirmatory): language syntax versus language semantics

2|3.1 Method

2|3.1.1 Participants

Sixty-two participants were invited to take part in the experiment (11 men, 54 right handed). None had participated in experiment 1. They were all Dutch native speakers, aged 21 on average (*SD* = 3.0), with 3.3 years of musical training on average (*SD* = 3.9). Forty-six were self-described non-musicians, sixteen amateur musicians. They were paid 12 € or undergraduate course credits for their participation and were naive as to the purpose of the experiment.

2|3.1.2 Design and Material

We employed a 2 (Key: first key ending or second key ending) × 2 (Difficulty: object-relative clause, subject-relative clause) design for the language syntax part of the experiment and a 2 (Key: first key ending or second key ending) × 2 (Difficulty: semantic garden-path, non-garden-path) design for the semantic part of the experiment. All factors were manipulated within-subjects. As in experiment 1, auditory musical stimuli and visual linguistic stimuli were presented concurrently. The critical point in time was always the ninth position (underlined in examples 3 and 4 below), corresponding to the start of a new key in the musical material and the disambiguating word in the sentences.

2|3.1.2.1 Music stimuli

We used the same music stimuli as in experiment 1.

2|3.1.2.2 Language syntax stimuli

As can be seen in example 2|3, 40 critical items were constructed in two versions each: an object-extracted relative clause and a subject-extracted relative clause. Sentences were 14 words long and were identical apart from the relative clause verb ('hielp'/'hielpen', *helped*$_{\text{singular}}$/*helped*$_{\text{plural}}$) which either agreed in number with the second noun-phrase ('de zoon', *the son*) in the object-extracted relative clause condition or the first noun phrase ('de vrienden', *the friends*) in the subject extracted relative clause condition.

(2|3) Language example (syntactic manipulation)

a) object-extracted relative clause (OR)

De | vrienden | **die** | de | zoon | op | de | been | hielp | lieten | hem | het | gebouw | zien.

Translation: *The friends* ***who*** *the son helped to get back on their feet let him see the building.*

b) subject-extracted relative clause (SR)

De | vrienden | **die** | de | zoon | op | de | been | hielpen | lieten | hem | het | gebouw | zien.

Translation: *The friends* ***who*** *helped the son get back on his feet let him see the building.*

The 40 filler items included various syntactic constructions. Comprehension prompts of critical trials targeted the object- or subject relative clause ambiguity (e.g., for example 2|3: 'De vrienden hielpen de zoon.' *The friends helped the son*. - false for object-relative clause, true for subject-relative clause). Half the prompts required a 'matches sentence' response, half the opposite response.

2|3.1.2.3 Language semantics stimuli

As can be seen in example 2|4, 40 critical items were constructed in two versions each: a semantic garden-path version and a non-garden-path version. Sentences were 14 words long and were identical apart from one manipulated word which occurred in positions two to seven (‘**muis**’/’**veldmuis**’, ***mouse*/*field vole***) which was either ambiguous or not. The disambiguating word (‘<u>rondlopen’</u>, <u>*run around*</u>) was always the ninth word of the sentence.

(2|4) Language example (semantic manipulation)

a) semantic garden-path (GP)

De | programmeur | liet | zijn | **muis** | op | de | tafel | <u>rondlopen</u> | nadat | hij | hem | had | gevoerd.

Translation: *The programmer let his **mouse** <u>run around</u> on the table after he had fed it.*

b) non-garden-path (non-GP)

De | programmeur | liet | zijn | **veldmuis** | op | de | tafel | <u>rondlopen</u> | nadat | hij | hem | had | gevoerd.

Translation: *The programmer let his **field vole** <u>run around</u> on the table after he had fed it.*

Additionally, 40 filler items were included. They used various syntactic constructions and avoided semantic ambiguities. Comprehension prompts of critical trials targeted various parts of the sentence (e.g., for example 2|4: ‘De muis was gevoerd.’ *The mouse had been fed.* - true). Half the prompts matched the sentence.

2|3.1.3. Pre-test 1: sentence completions show the intended misinterpretation of semantic material

Before starting the main experiment we conducted two pre-tests with two purposes. Firstly, we wanted to choose the 40 best semantic items out of an original set of 60. Items were included if they exhibited the typical semantic garden-path pattern of a semantically ambiguous word being initially misinterpreted until a disambiguating word changes the interpretation. Below we only report the analysis involving the 40 items which were used in experiment 2. Secondly, as in experiment 1, we wanted to establish the strength of the manipulations in the syntactic and the semantic parts of the experiment.

In a first pre-test (*N* = 54) participants were asked to complete sentence beginnings which did not include a disambiguating word (words one to eight, e.g., ‘De programmeur liet zijn **muis** op de tafel …’, *The programmer let his **mouse** …*). Sentence completions revealed that the intended word interpretation was overwhelmingly adopted in the non-garden-path condition (93.6% of trials) while this hardly happened in the semantic garden-path condition (13.3%). This is exactly the pattern expected of semantic garden-path sentences (see supplementary materials).

2|3.1.4 Pre-test 2: the strength of the syntactic and semantic manipulations

The aim of the second pre-test was to establish the strength of the difficulty manipulations of the syntactic and the semantic items. Participants (*N* = 24) read sentences word by word while the timing of word presentation was under their control (self-paced reading). Afterwards they rated trials for difficulty. The semantic manipulation led to higher reading times from the disambiguating word onwards in the semantic garden-path condition (compared to non-garden-path condition). The difference was 44 ms ($p = .005$) for the critical disambiguating word and 85 ms ($p < .001$) for the post-critical word. Overall, this resulted in a greater perceived difficulty with these sentences ($p = .003$). Moreover, the syntactic manipulation was equally salient as seen in reading times (reading time difference on critical ninth word: 64 ms [$p = .009$], on post-critical word: 58 ms [$p = .008$]) and difficulty ratings ($p = .001$; see supplementary materials). Thus, if we observe differences between the syntactic and the semantic manipulations these cannot be attributed to quantitatively different processing requirements of the two manipulations as they were matched in this respect.

2|3.1.5 Procedure

The task was the same as in experiment 1. The type of manipulation (syntax or semantics) was blocked and the order counter-balanced. Experimental sessions did not include a working memory test. A testing session took approximately 90 minutes.

2|3.1.6 Analysis

The analysis was the same as in experiment 1.

2|3.2 Results

2|3.2.1 General task performance

Before presenting the crucial critical trial data, we first show evidence of good task-adherence. Concerning the music task, using the filler trials we found, as expected, that all participants had a greater proportion of high-closure ratings for the authentic cadence endings than the no-cadence endings ($M_{\text{authentic cadence}} = .76 > M_{\text{no cadence}} = .09$, difference $SD = .18$).

Accuracy during the syntax part was high in general ($M = 70\%$, $SD = 9\%$). As expected, prompts after object-relative clauses were answered less accurately ($M = 37\%$, $SD = 24\%$) than those after subject relative clauses ($M = 76\%$, $SD = 19\%$) [$t_{(61)} = 9.80$, $p < .001$, ${}_{p}\eta^2 = .661$]. Accuracy during the semantics part was higher in general ($M = 84\%$, $SD = 7\%$). There was no significant accuracy difference between garden-path ($M = 86\%$, $SD = 14\%$) and non-garden-path trials ($M = 85\%$, $SD = 13\%$) [$t_{(61)} < 1$]. However, this null effect should not be taken to mean that the semantic garden-path effect was somehow absent. Instead, it reflects the fact that the comprehension prompts in the semantics part often did not target the semantic ambiguity of the sentences. As pre-test 2 showed, using the same comprehension prompts, critical semantics trials still exhibited a garden-path effect both in reading time and difficulty ratings, showing that it is not necessary to focus comprehension prompts on the garden-path manipulation in order for it to have an effect.

2|3.2.2 Critical task performance

2|3.2.2.1 Critical results of the syntax part

The critical closure data of the syntax part were analyzed in a 2 (Difficulty: object-relative clause, subject-relative clause) × 2 (Key: first key ending, second key ending) within-subjects ANOVA, see Figure 2|4A. Object-relative clauses led to significantly fewer (M = .64, SD = .15) high closure ratings than subject-relative clauses (M = .68, SD = .14) [Difficulty, $F_{(1,61)} = 5.45$, $p = .023$, ${}_{p}\eta^2 = .082$]. Otherwise, first key endings were less likely (M = .59, SD = .17) to receive a high closure rating than second-key endings (M = .73, SD = .15) [Key, $F_{(1,61)} = 38.53$, $p < .001$, ${}_{p}\eta^2 = .387$]. These two factors did not interact [Key × Difficulty, $F_{(1,61)} = 2.48$, $p = .121$, ${}_{p}\eta^2 = .039$]. Still, as shown in Figure 2|4A, the simple main effects revealed a Difficulty effect only for first key endings [$t_{(61)} = 2.51$, $p = .015$, ${}_{p}\eta^2 = .094$], not for second key endings [$t_{(61)} < 1$].

2|3.2.2.2 Critical results of the semantics part

The critical closure data of the semantics part were analyzed in a 2 (Difficulty: semantic garden-path, non-garden-path) × 2 (Key: first key ending, second key ending) within-subjects ANOVA, see Figure 2|4B. Closure ratings did not differ between semantic garden-path (M = .63, SD = .16) and non-garden-path trials (M = .62, SD = .15) [$F_{(1,61)} < 1$]. Otherwise, first key endings were less likely (M = .56, SD = .16) to receive a high closure rating than second-key endings (M = .69, SD = .15) [Key, $F_{(1,61)} = 27.00$, $p < .001$, ${}_{p}\eta^2 = .307$]. These two factors did not interact [Key × Difficulty, $F_{(1,61)} < 1$].

2|3.2.2.3 Language syntax vs. Language semantics

In order to see whether the influence of the syntax manipulation on music ratings was significantly different from the influence of the semantics manipulation we performed a three-way within-subjects ANOVA with the factors Manipulation (syntax, semantics), Difficulty (object-relative clause/semantic garden-path, subject-relative clause/non-garden-path) and Key (first key ending, second key ending). Indeed, the difficulty effect was different in the syntax part and the semantics part [Manipulation × Difficulty, $F_{(1,61)} = 4.10$, $p = .047$, ${}_p\eta^2 = .063$]. Otherwise, only one main effect was significant [Key, $F_{(1,61)} = 44.56$, $p < .001$, ${}_p\eta^2 = .422$] [Manipulation, $F_{(1,61)} = 3.85$, $p = .054$, ${}_p\eta^2 = .059$] [Difficulty, $F_{(1,61)} < 1$] [Manipulation × Key, $F_{(1,61)} < 1$] [Difficulty × Key, $F_{(1,61)} = 1.39$, $p = .242$, ${}_p\eta^2 = .022$] [Manipulation × Difficulty × Key, $F_{(1,61)} < 1$].

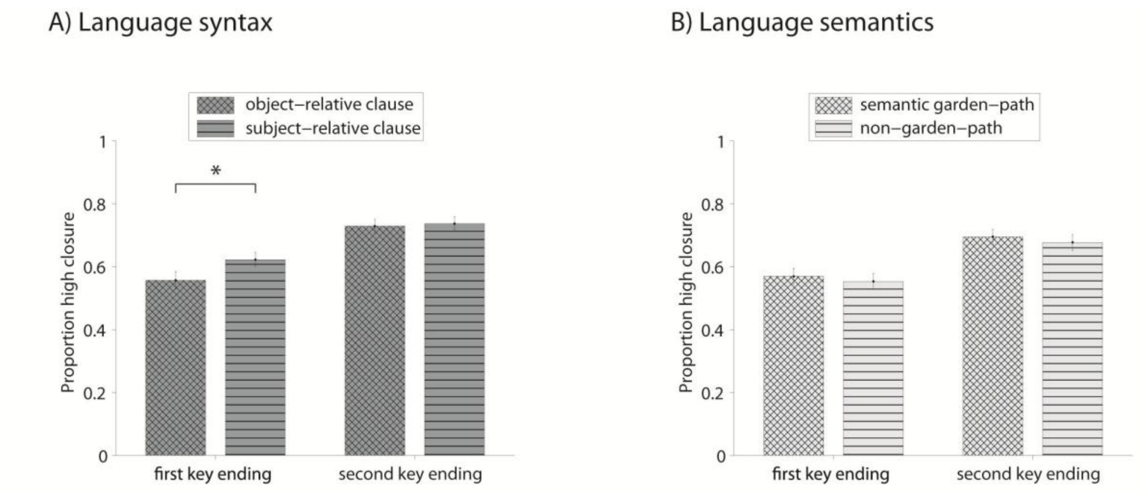


Figure 2|4 Experiment 2: Closure ratings of critical trials. Participants were asked to rate their feeling of closure (y-axis) of first-key-typical endings or second-key-typical endings (x-axis). While people solved the auditory music task they were also asked to read sentences. **A)** We found an influence of a syntactic manipulation on music harmony ratings. **B)** A semantic garden-path manipulation was without effect. Significance levels represented as stars refer to simple main effect *t*-tests and do not imply significant interaction effects. Error bars = SEM. * p < .05

2|3.3 Discussion: language syntax affects harmony judgements, semantics does not

In line with the exploratory first experiment, experiment 2 has shown that a syntactic challenge in language can influence harmony perception. Moreover, it was shown that a non-syntactic language manipulation, namely a semantic garden-path manipulation, has no such effect. This pattern of results is indicative of a shared syntactic resource account (2008), not a general attention mechanism (Jones & Boltz, 1989; Perruchet & Poulin-Charronnat, 2013). Interestingly, our findings suggest that semantic processing itself does not influence music harmony perception in contrast to the influence of music itself on semantic processing found by Perruchet & Poulin-Charronnat (2013). Potential reasons for this will be discussed.

The difference between the syntactic effect and the semantic effect is remarkable. These two manipulations were rated as equally salient and led to very similar reading time changes. Nonetheless, the nature of the manipulation had different effects on the concurrent music perception. As discussed in section 2|2.3, one could also propose a memory-based explanation for the difference between syntactic and semantic effects on music. Differences in the distance between the ambiguous word and the disambiguating ninth word could be responsible. However, when looking at these distances it becomes clear that the semantic manipulation (distance ambiguous - disambiguating word: *M* = 4.08, *SD* = 1.65, 95% Confidence Interval based on Bootstrapping with 50,000 sample = [3.58; 4.60]) was intermediate between the two syntactic manipulations (experiment 1: always 3 words between ‘**en**’, ***and***, and the ninth word; experiment 2: distance between ‘**die’**, ***who*** and disambiguating ninth word, *M* = 4.98 words, *SD* = 1.03, 95% CI = [4.65; 5.28]). This makes a memory based explanation for the difference between the semantic and the syntactic effects unlikely.

Having investigated what kind of manipulations do exert an influence on music harmony perception, we will now turn towards a more detailed analysis of the syntax effect itself. So far, we were unable to distinguish an effect on either of two harmonic sub-processes: holding a key online (reflected in first-key ending ratings) and establishing a new key (reflected in second-key ending ratings). The greater statistical power afforded by combining the data of the syntactic manipulations of both experiments might illuminate this issue.

2|4 Experiments 1 & 2: combined analysis of the syntax effect

2|4.1 Results

We combined the data of the syntax manipulations of experiment 1 (ambiguous S-coordination, NP-coordination) and experiment 2 (object-relative clause, subject-relative clause) to form a single Difficulty factor, see Figure 2|5. A 2 (Experiment: one, two) × 2 (Difficulty) × 2 (Key: first key ending, second key ending) mixed between- and within-subjects ANOVA exhibited the main effects of Difficulty and Key which we observed before [Difficulty, $F_{(1,114)} = 17.63$, $p < .001$, ${}_p\eta^2 = .134$] [Key, $F_{(1,114)} = 89.47$, $p < .001$, ${}_p\eta^2 = .440$]. The factor Experiment was without effect [Experiment, $F_{(1,114)} < 1$] [Difficulty × Experiment, $F_{(1,114)} = 1.50$, $p = .224$, ${}_p\eta^2 = .013$] [Key × Experiment, $F_{(1,114)} < 1$] [Difficulty × Key × Experiment, $F_{(1,114)} < 1$]. However, the greater power revealed a previously non-significant interaction between Difficulty and Key [$F_{(1,114)} = 4.21$, $p = .042$, ${}_p\eta^2 = .036$]. Follow-up *t*-tests showed that the difficulty effect was specific to first-key-typical endings [$t_{(115)} = 4.09$, $p < .001$, ${}_p\eta^2 = .127$] while it was non-significant for second-key-typical endings [$t_{(115)} = 1.43$, $p = .155$, ${}_p\eta^2 = .018$]. Note that including musical training as a covariate does not change the pattern of these results. In the overall ANCOVA *z*-scored musical training years did not modulate any of the main effects or interactions [*p*s > .2], see supplementary materials. In the supplementary materials we also provide analyses by items.

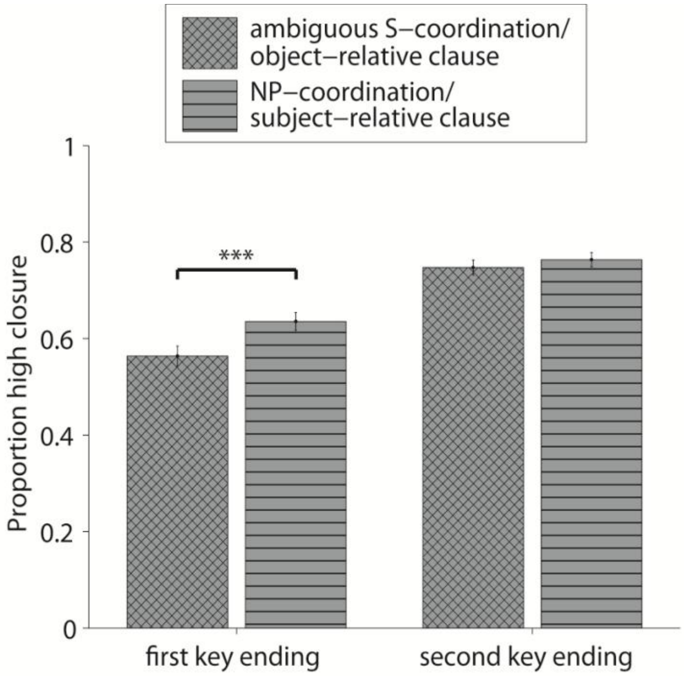


Figure 2|5. Experiments 1 & 2: Closure ratings of critical trials. Participants were asked to rate their feeling of closure (y-axis) of first-key-typical endings or second-key-typical endings (x-axis). While people solved the auditory music task they were also asked to read sentences. The language syntax effect was specific to harmonic sequences ending in a first-key-typical way. Error bars = SEM. *** p < .001

2|4.2 Discussion: a first-key specific syntax effect

Using the greater power afforded by the combination of the data of experiments 1 and 2, we have shown that the syntax effect is specific to harmonic sequences ending in the first key. This suggests that during music listening shared music-language resources are involved in holding a key online in order to interpret incoming chords in its context. Building up a new key seems to be processed by different resources. We will discuss the implications of these findings in the general discussion. One should bear in mind that experiment 1 was exploratory in nature, and thus the result of the combined analysis should be replicated.

The comparison also showed that the syntax manipulation in experiment 1 appears to be numerically stronger ($_p\eta^2$ = .169 for the difficulty effect related to first key endings) than in experiment 2 ($_p\eta^2$ = .094). Looking at the comprehension accuracy data reveals that this small difference might be related to participants sometimes giving up a full syntactic analysis of the very challenging object-relative clauses (accuracy = 37%) but do so less in the challenging ambiguous S-coordination trials (accuracy = 83%). This might indicate that syntactic integration resources can also be taxed too much, leading to incomplete parsing of the sentence (Ferreira, Bailey, & Ferraro, 2002). The details of such non-linear effects should be worked out in future studies. For the present purposes we can conclude that challenging syntax processing, even when it leads to incomplete syntactic interpretations, draws resources away from the harmonic integration of chords.

2|5 General discussion

2|5.1 Summary

We present two experiments which investigated whether music harmony processing can be influenced by a concurrent language or arithmetic task. Patel's (2008) SSIRH hypothesises that previously observed musical influences on the processing of language syntax are due to shared musico-linguistic resources involved in syntactic integration. Therefore, under this account music harmony should be influenced only if the concurrent task involves a syntactic manipulation. Our results support this prediction. While two different language syntax contrasts (ambiguous S-coordination versus NP-coordination and object-relative clause versus subject-relative clause) modulated harmonic closure ratings, arithmetical difficulty as well as a semantic garden-path manipulation had no such effect. This contradicts accounts relating previous music-language interactions solely to general attention mechanisms which would have predicted an influence of the arithmetic and semantic control manipulations as well.

The functional role of shared syntax resources in terms of harmonic processing was further investigated by looking at the syntax effect on ratings of different kinds of musical endings. This analysis revealed that a syntactic challenge in language reduces the music listener's ability to maintain an already established key online for the purpose of harmonic integration of chords. Building up a new key was unaffected. In what follows we will discuss these results.

2|5.2 Novel support for shared syntactic integration resources

The present findings offer novel support for Patel's (2008) SSIRH. The novelty stems from the use of a musical measure of interest, as opposed to previously used linguistic measures such as language comprehension accuracy (Fedorenko et al., 2009), reading times (Slevc et al., 2009) or word judgment times (Hoch et al., 2011). Our findings suggest that the direction of influence is not crucial for interference effects. Music can influence language and vice versa. This is in line with studies which measured brain activity during music and language processing in similar paradigms. For example, in an EEG event-related potential (ERP) study, Steinbeis & Koelsch (2008) found language processing in the form of the left anterior negativity (LAN) to be altered by a concurrent harmonic violation (see also Koelsch, Gunter, et al., 2005). They also found the early right anterior negativity (ERAN) elicited by an unexpected chord to be influenced by language syntax errors (but see Koelsch, Gunter, et al., 2005). This was taken as evidence for a change in musical processing as a result of a linguistic syntactic violation. Clearly, both offline measures as used here or by Fedorenko et al. (2009) and on-line measures of processing (Hoch et al., 2011; Koelsch, Gunter, et al., 2005; Kunert et al., 2015; Slevc et al., 2009; Steinbeis & Koelsch, 2008) reveal a mutual influence between music and language.

However, there is also a negative deflection in the EEG signal which peaks around 650ms after stimulus onset and which is elicited by an unexpected chord. This deflection is often called the N5 and it can be modulated by a language semantics manipulation (Steinbeis & Koelsch, 2008), suggesting an effect of language semantics on music harmony in contradiction to our results in experiment 2. However, the N5 is not elicited in some experiments (Koelsch, Gunter, et al., 2005) or participant groups (Featherstone, Morrison, Waterman, & MacGregor, 2013). Moreover, other EEG studies working with similar interference paradigms failed to find an interaction between music harmonics and language semantics (Besson et al., 1998; Koelsch, Gunter, et al., 2005; see also Carrus et al., 2013). Therefore, it appears difficult to interpret the apparent contradiction between Steinbeis and Koelsch's (2008) ERP result and our semantics null effect in experiment 2. More research is needed in order to better characterize the N5. Until then we are inclined to rely on the straightforward interpretation of our behavioural measure to indicate an absence of semantic influences on harmonic processing. Only language syntax appears to reliably influence music harmony behaviour.

2|5.3 The role of shared resources in the music network

In the current experiment only the first harmonic key which was held online was affected by a concurrent language syntax manipulation. This could be taken to reflect shared resource's role as a specifically *syntactic* working memory in the brain's music network (Fiveash & Pammer, 2014). In terms of language, this aspect of working memory makes 'syntactic information actively available over sustained periods of the sentence while new information is being processed continuously' (Fiebach et al., 2005, p. 88). An additional role of syntactic working memory could be to hold a *harmonic key* online while new chords are integrated. This process might be impaired if concurrently processing complex sentences restricts the syntactic working memory resources left available for music.

Alternatively, music and language might share syntactic unification resources (Hagoort, 2005, p. 416/417). According to this account, lexical items are retrieved from memory into a 'unification workspace' in which 'constituent structures spanning the whole utterance are formed' by combining items according to syntactic principles, i.e. by unifying them. The more demanding unification operation associated with disambiguating words of garden-path sentences might impair the ability of shared resources to keep a harmonic key in the unification workspace. Therefore the key is no longer a potential site with which incoming chords can be unified. The present study cannot distinguish between these two accounts.

2|5.4 Non-syntactic overlap between music and language?

We found no evidence for a general attentional mechanism driving the linguistic influence on music, as both control conditions were without effect. Thus, the nature of the manipulation (syntactic rather than something else) appears important. However, according to Perruchet and Poulin-Charronat's (2013) attentional load account influences from one domain on the other only emerge if enough attentional resources are left to process both. In the present experiments, it could be argued that critical language syntax trials might be overall less demanding than control task trials, leaving more attention to music perception. However, there is no evidence for such an account here. As the difficulty ratings of the pre-tests revealed, critical language syntax trials were sometimes seen as *overall* easier (experiment 1) or harder (experiment 2) than their control task counterparts (ps ≤ .001; see supplementary material). Thus, in this study, there is no linear relationship between the effect of a task on music perception and its overall difficulty. General attention as well as shared syntactic resources might be involved in musical influences on language (Perruchet & Poulin-Charronnat, 2013; Slevc et al., 2009), while language influencing music relies solely on syntactic resources. This proposal should be tested in the future.

Next to attention based explanations, an error based explanation of previous studies' results (Hoch et al., 2011; Koelsch, Gunter, et al., 2005; Steinbeis & Koelsch, 2008) was put forward by Rogalsky, Rong, Saberi, & Hickok (2011). However, our study only used well-formed sentences and musical sequences without errors, as done in previous behavioural studies investigating musical influences on language (Fedorenko et al., 2009; Slevc et al., 2009). Similarly, our results could also be said to be due to a working memory based mechanism (not to be confused with the notion of syntactic working memory (Fiebach et al., 2005; Fiveash & Pammer, 2014)). In order to hold a key online listeners might need working memory resources which the syntactic/semantic reinterpretation of a previously encountered word taxes. However, phonological working memory was not associated with the syntax effect found in experiment 1, and working memory effects would lead to similar semantic and syntactic garden-path influences on harmonic processing. Instead, we found that only the latter affects music ratings. This leads us to favour the SSIRH put forward by Patel (2008) as the likely mechanism behind our results, rather than attentional accounts (Jones & Boltz, 1989; Large & Jones, 1999; Perruchet & Poulin-Charronnat, 2013), an error processing explanation (Rogalsky et al., 2011), or a working memory account.

2|5.5 Conclusion

The present study found evidence that music and language share resources for syntactic integration (Patel, 2008). Our results suggest that challenging language syntax impacts on the ability to integrate chords into an already established key. It appears that music and language are processed in a similar way by the human brain due to a key commonality: they are both structured sequences, i.e. their constituent elements relate in a rule-governed way to each other. Shared syntax resources are responsive to structured sequences in general - whether they are linguistic or musical in nature.

2|6 Ethics statement

All participants provided informed, written consent and the study was approved by the local ethics committee ‘‘CMO Arnhem Nijmegen’’ (CMO no 2001/095 and amendment ‘‘Imaging Human Cognition’’ 2006, 2008), in accordance with the Research involving human subjects Act, following the principles of the Declaration of Helsinki.

2|7 Data accessibility statement

The datasets supporting this article have been uploaded as part of the Supplementary Material at http://rsos.royalsocietypublishing.org/content/3/2/150685.figures-only. Matlab code for reproducing Figures 2|3, 2|4, and 2|5 is provided as part of the Supplementary Material under the same internet address as well.

2|8 Competing interest statement

We report no competing interests.

2|9 Author’s contribution statement

RK conceived and designed the experiments, acquired the data, analysed them and wrote the paper. RMW conceived the experiments and revised the paper. PH conceived the experiments and revised the paper. All authors gave final approval for publication.

2|10 Acknowledgements

We would like to thank L. Robert Slevc, Aniruddh D. Patel, Sean J. Fallon, and Susanne Vogel for helpful discussions on the background and analysis of the study. We are grateful to Gwilym Lockwood for providing useful feedback on an earlier version of this draft.

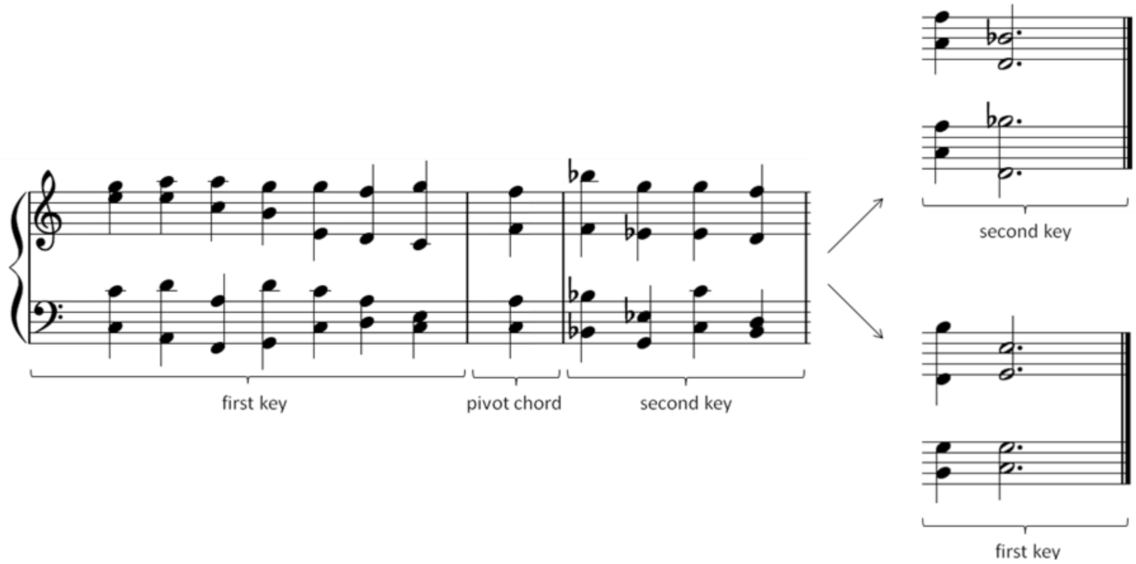
2|11 Funding statement

This work was supported by an NWO Spinoza Prize awarded to PH and a PhD grant from the Max Planck Society to RK.

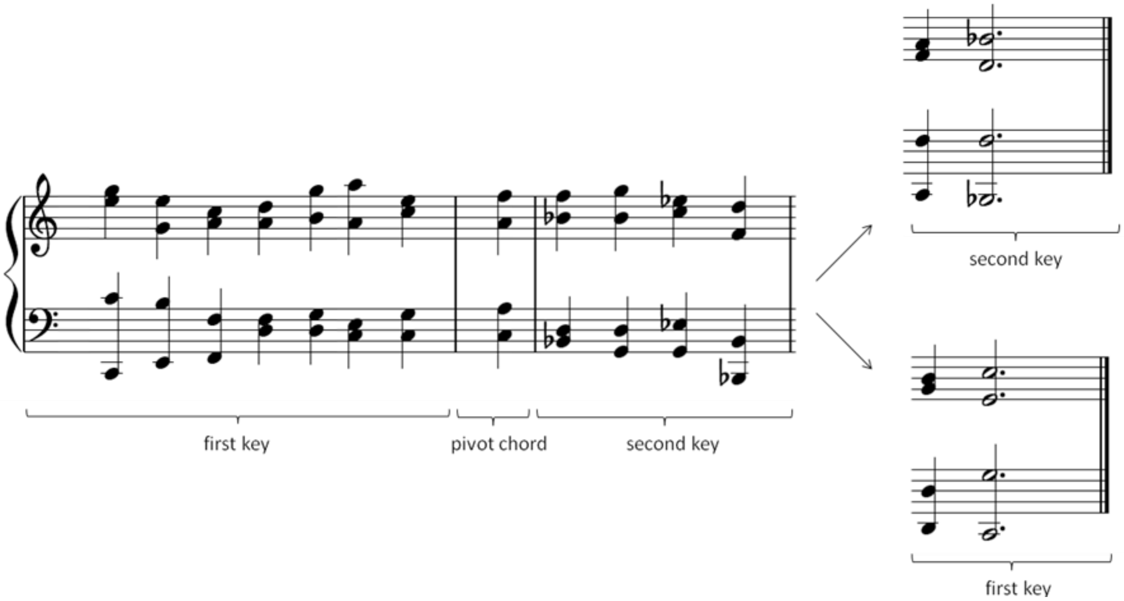
2|12 Appendix

Experiments 1 & 2: Music

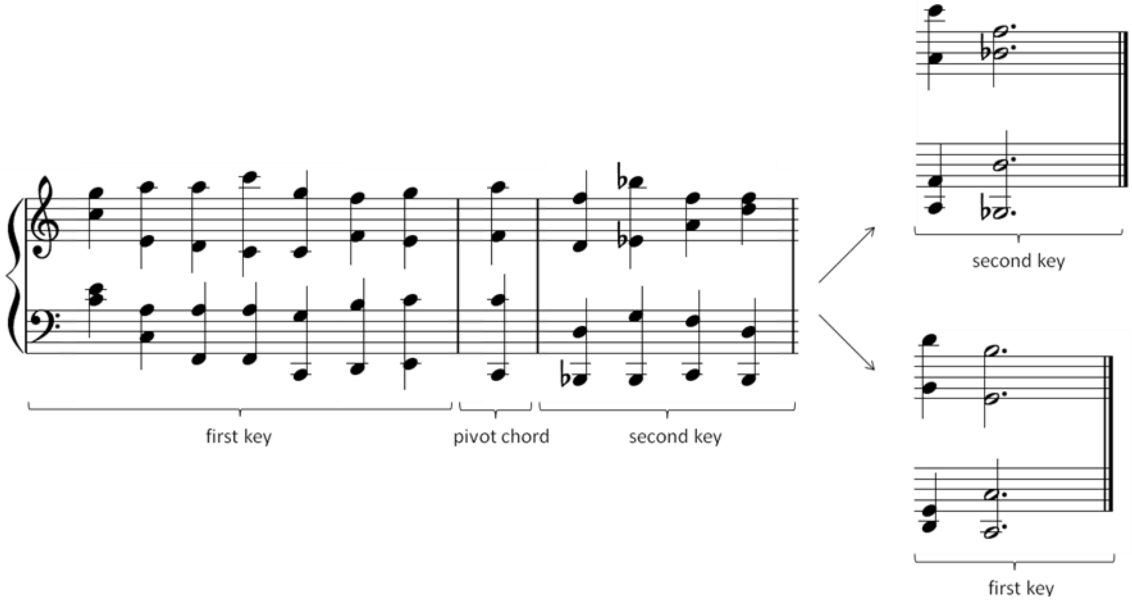
Only the C-major/B-flat major version is shown. Two transpositions were made of these items.
See section 2|2.1.2.1 for details.

#	item
1	

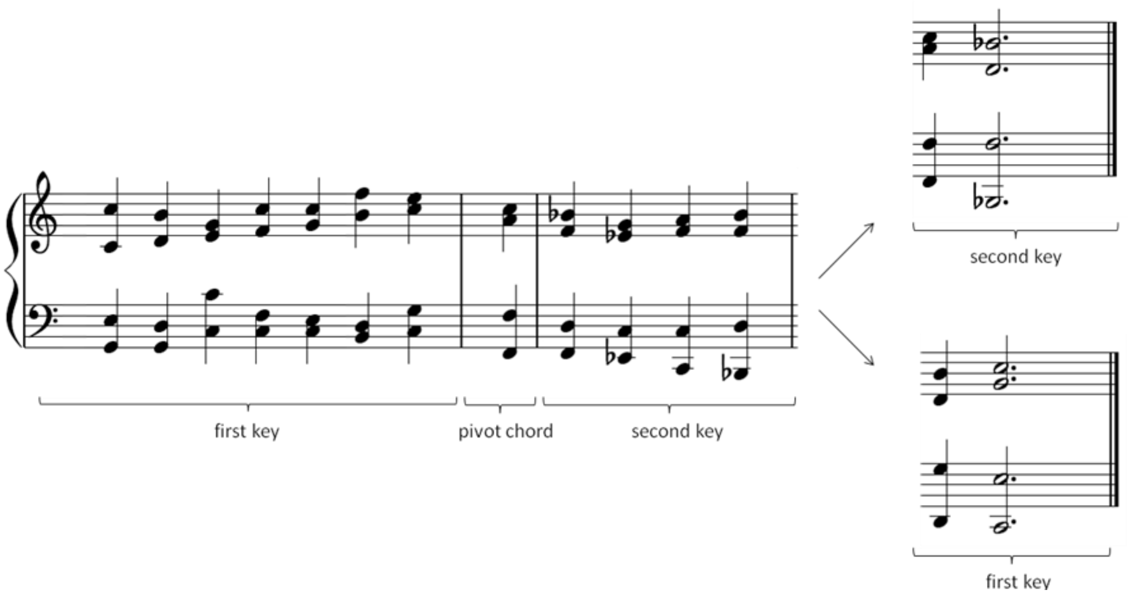
2


first key
pivot chord
second key
second key
first key

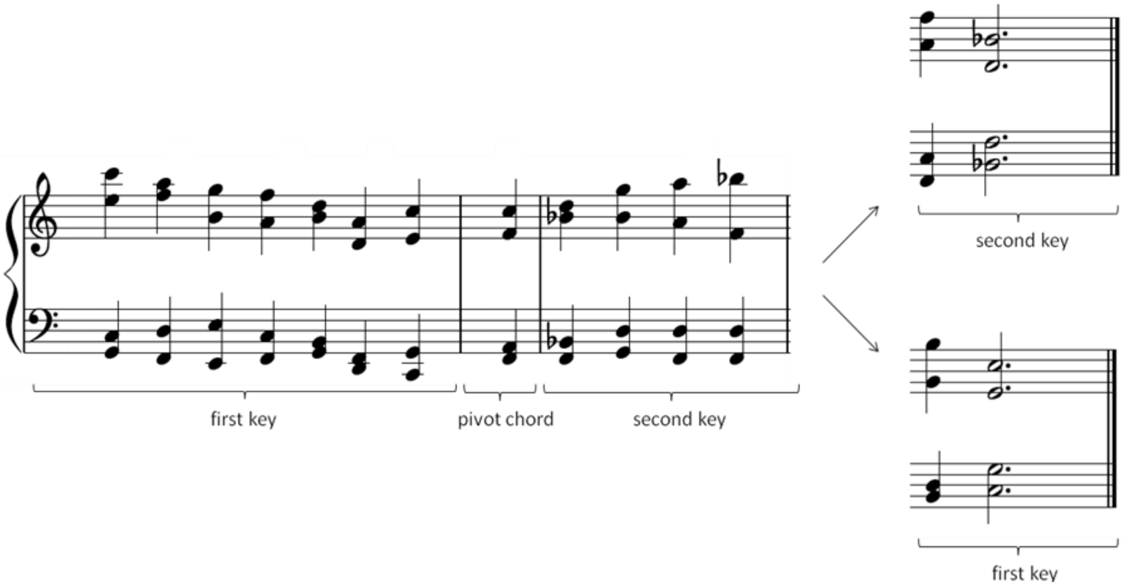
3


first key
pivot chord
second key
second key
first key

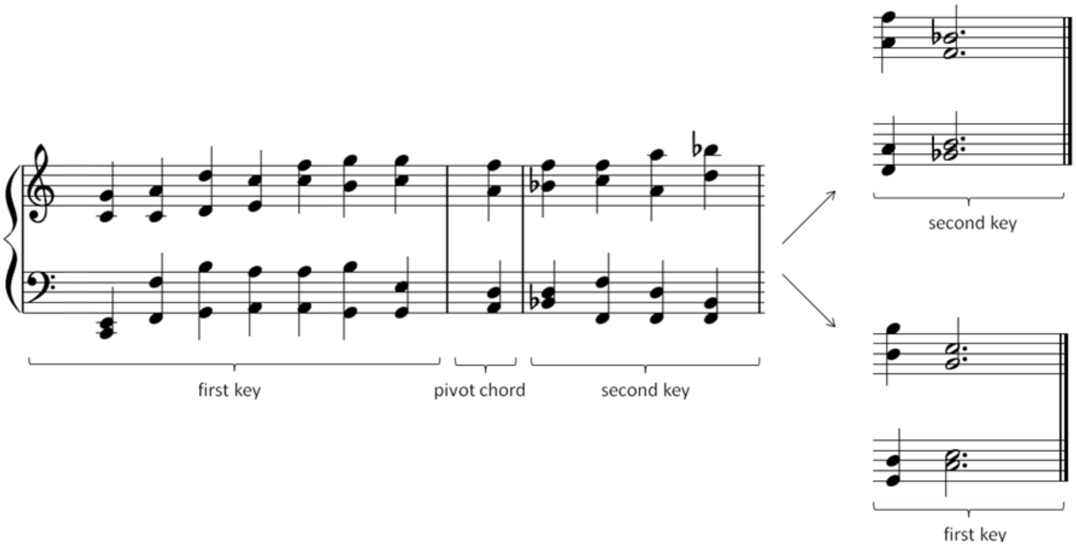
4


second key
first key
pivot chord
second key
first key

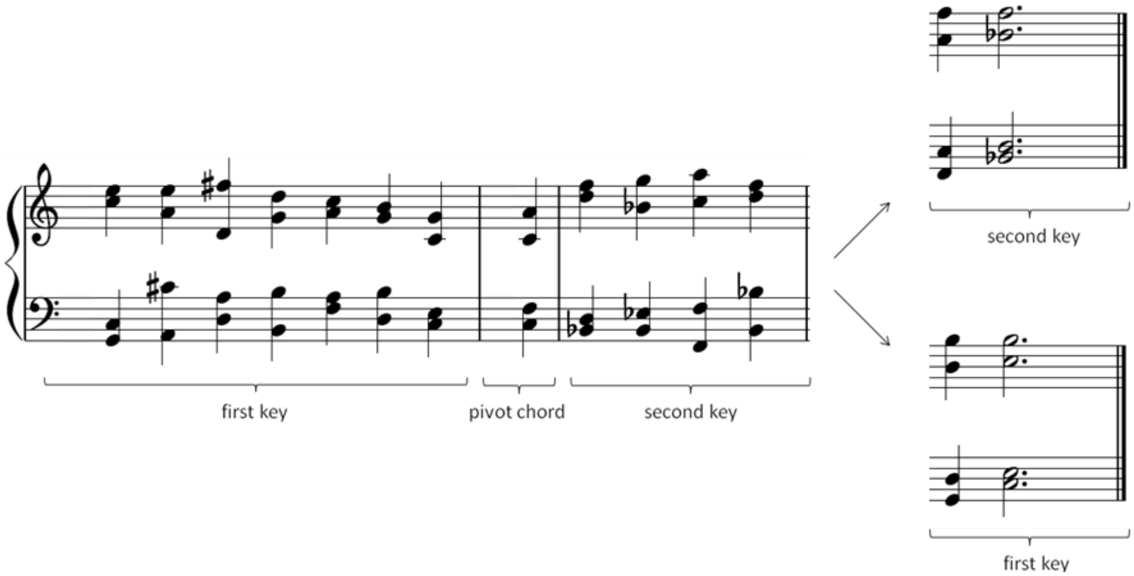
5


second key
first key
pivot chord
second key
first key

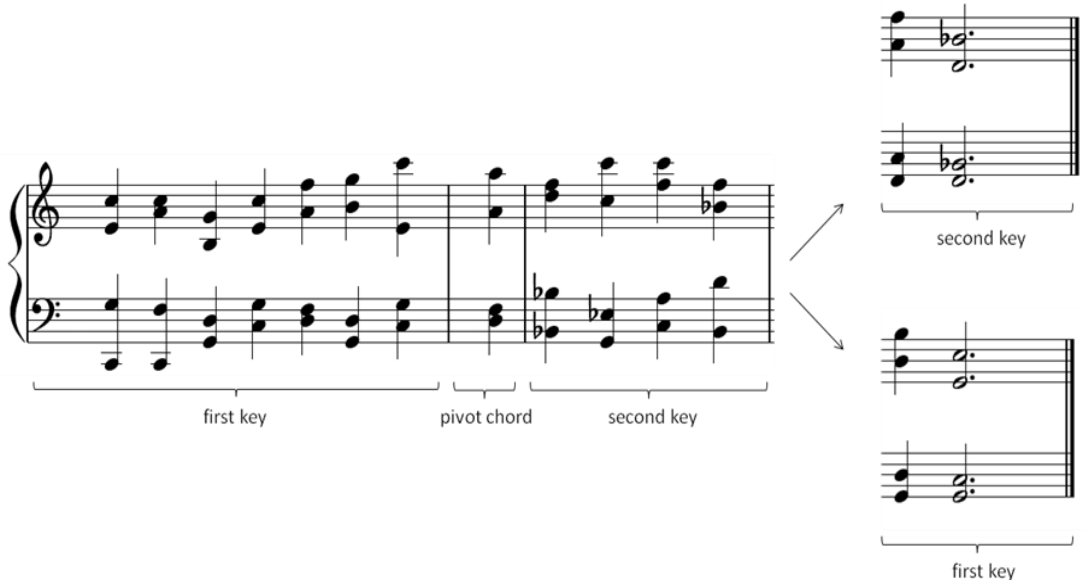
6


first key
pivot chord
second key
second key
first key

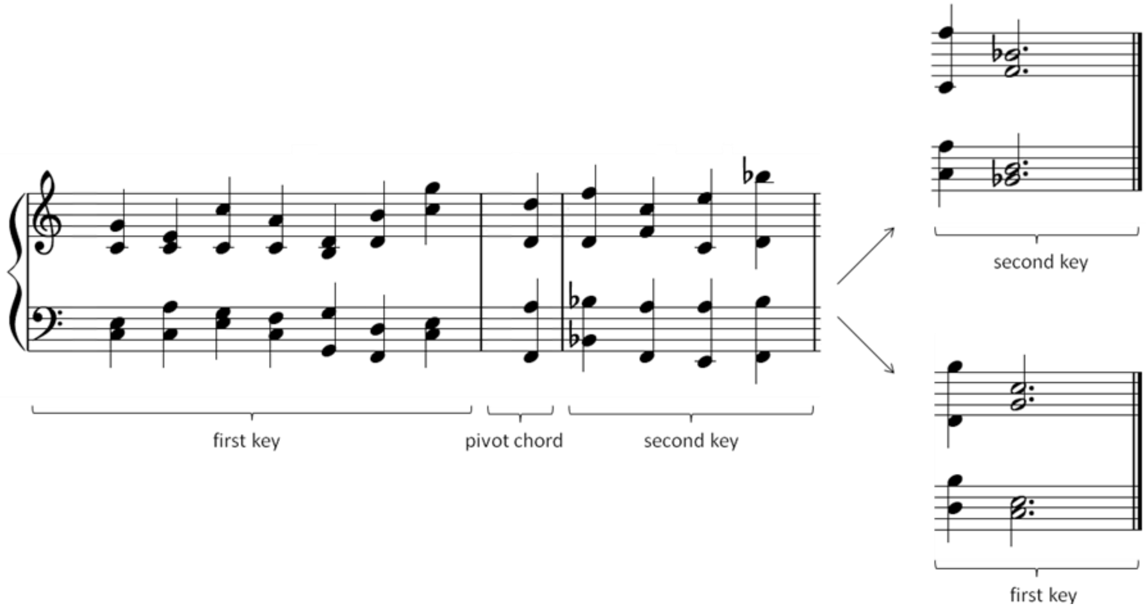
7


first key
pivot chord
second key
second key
first key

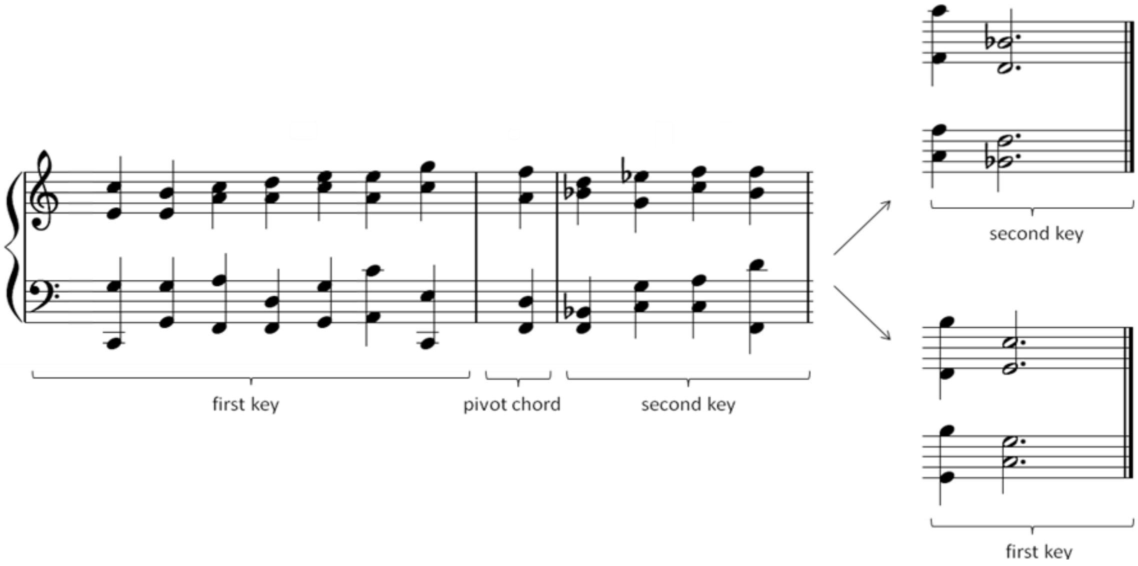
8


first key
pivot chord
second key
second key
first key

9


first key
pivot chord
second key
second key
first key

10



Experiment 1: Language syntax

#	ambiguous S-coordination	unambiguous S-coordination	NP-coordination
1	De chirurg troostte de man en de vrouw legde haar hand op zijn voorhoofd.	De chirurg troostte de man, en de vrouw legde haar hand op zijn voorhoofd.	De chirurg troostte de man en de vrouw omdat de operatie niet gelukt was.
2	De voorzitter bedankte de sponsor en de trainer bestelde lachend een biertje voor iedereen.	De voorzitter bedankte de sponsor, en de trainer bestelde lachend een biertje voor iedereen.	De voorzitter bedankte de sponsor en de trainer zonder de goede spelers te noemen.
3	De mannequin kuste de ontwerper en de fotograaf pakte een champagne en wat kaviaar.	De mannequin kuste de ontwerper, en de fotograaf pakte een champagne en wat kaviaar.	De mannequin kuste de ontwerper en de fotograaf hoewel zij hen niet leuk vond.
4	De rector ondervroeg de leraar en de leerling volgde het gesprek vanaf de gang.	De rector ondervroeg de leraar, en de leerling volgde het gesprek vanaf de gang.	De rector ondervroeg de leraar en de leerling over het ongeval op het schoolplein.
5	De gevangene gijzelde de priester en de bewaker riep geschrokken om hulp te krijgen.	De gevangene gijzelde de priester, en de bewaker riep geschrokken om hulp te krijgen.	De gevangene gijzelde de priester en de bewaker in de omsingelde gevangenis van Breda.
6	De weduwe bedankte de organist en de predikant	De weduwe bedankte de organist, en de predikant	De weduwe bedankte de organist en de predikant voor

	bekeek de mensen die waren gekomen.	bekeek de mensen die waren gekomen.	de geschikte dienst op Allerzielen.
7	De bedrijfsleider kalmeerde de gast en de ober bracht het bord naar de keuken.	De bedrijfsleider kalmeerde de gast, en de ober bracht het bord naar de keuken.	De bedrijfsleider kalmeerde de gast en de ober nadat het licht ineens weer aanging.
8	8 De redacteur prees de fotograaf en de journalist bekeek bewonderend de foto's van vluchtelingen.	De redacteur prees de fotograaf, en de journalist bekeek bewonderend de foto's van vluchtelingen.	De redacteur prees de fotograaf en de journalist voor het artikel over de vluchtelingen.
9	De sheriff beschermde de boer en de knecht verdedigde de boerderij tegen Jonsons bende.	De sheriff beschermde de boer, en de knecht verdedigde de boerderij tegen Jonsons bende.	De sheriff beschermde de boer en de knecht tegen de aanval van Johnson's bende.
10	De grimeur schminkte de schrijver en de interviewer besprak de vragen die zouden komen.	De grimeur schminkte de schrijver, en de interviewer besprak de vragen die zouden komen.	De grimeur schminkte de schrijver en de interviewer voordat de camera's begonnen te filmen.
11	De verdachte beledigde de rechter en de advocaat belde ontstemd het kantoor zonder succes.	De verdachte beledigde de rechter, en de advocaat belde ontstemd het kantoor zonder succes.	De verdachte beledigde de rechter en de advocaat in de rechtszaal voor het publiek.
12	De eigenaar prees de kok en de ober floot zachtjes een heel bekend liedje.	De eigenaar prees de kok, en de ober floot zachtjes een heel bekend liedje.	De eigenaar prees de kok en de ober op het jubileumfeest van zijn restaurant.
13	De dirigent bekritiseerde de cellist en de pianist smeet zijn partituur op de grond.	De dirigent bekritiseerde de cellist, en de pianist smeet zijn partituur op de grond.	De dirigent bekritiseerde de cellist en de pianist omdat zij niet harmonieus samen speelden.
14	De portier bespioneerde de chef en de secretaresse belde heimelijk vrienden bij de politie.	De portier bespioneerde de chef, en de secretaresse belde heimelijk vrienden bij de politie.	De portier bespioneerde de chef en de secretaresse op order van de rijke baron.
15	De dief beschoot de juwelier en de agent riskeerde zijn leven door te intervenieren.	De dief beschoot de juwelier, en de agent riskeerde zijn leven door te intervenieren.	De dief beschoot de juwelier en de agent met een gestolen vuurwapen uit Amerika.
16	De regisseur bespotte de nieuwslezer en de weerman vervloekte de opzet van het programma.	De regisseur bespotte de nieuwslezer, en de weerman vervloekte de opzet van het programma.	De regisseur bespotte de nieuwslezer en de weerman nadat hij veel wijn had gedronken.

17	De winnares omhelsde de sponsor en de trainer groette het publiek op de tribune.	De winnares omhelsde de sponsor, en de trainer groette het publiek op de tribune.	De winnares omhelsde de sponsor en de trainer met een glimlach op haar lippen.
18	De rechter berispte de verdachte en de advocaat bedacht snel redenen voor een verdaging.	De rechter berispte de verdachte, en de advocaat bedacht snel redenen voor een verdaging.	De rechter berispte de verdachte en de advocaat omdat zij onafgebroken met elkaar praatten.
19	De presentator introduceerde de schrijver en de criticus maakte grijnzend buigingen naar het publiek	De presentator introduceerde de schrijver, en de criticus maakte grijnzend buigingen naar het publiek	De presentator introduceerde de schrijver en de criticus aan het publiek in de zaal.
20	De stalker achtervolgde de danseres en de manager opende vlug de deur van metaal.	De stalker achtervolgde de danseres, en de manager opende vlug de deur van metaal.	De stalker achtervolgde de danseres en de manager hoewel hij eigenlijk naar huis moest.
21	De politieman ondervroeg de koerier en de infiltrant achterhaalde later zonder problemen zijn naam.	De politieman ondervroeg de koerier, en de infiltrant achterhaalde later zonder problemen zijn naam.	De politieman ondervroeg de koerier en de infiltrant in een onderzoek naar internationaal terrorisme.
22	De gravin wenkte de koetsier en de lakei droeg zuchtend de koffers naar huis.	De gravin wenkte de koetsier, en de lakei droeg zuchtend de koffers naar huis.	De gravin wenkte de koetsier en de lakei vanuit de koets die geparkeerd stond.
23	De presentator omarmde de zanger en de zangeres zong huilend hun trieste eerste hit.	De presentator omarmde de zanger, en de zangeres zong huilend hun trieste eerste hit.	De presentator omarmde de zanger en de zangeres terwijl hun trieste eerste hit weerklonk.
24	De hulpverlener informeerde de arts en de brandweerman bracht gehaast het slachtoffer buiten gevaar.	De hulpverlener informeerde de arts, en de brandweerman bracht gehaast het slachtoffer buiten gevaar.	De hulpverlener informeerde de arts en de brandweerman over de gevaren van witte asbest.
25	De tovenaar bewaakte de koningin en de prinses haalde het toverboek om te helpen.	De tovenaar bewaakte de koningin, en de prinses haalde het toverboek om te helpen.	De tovenaar bewaakte de koningin en de prinses met zijn toverstafje voor de draak.
26	De voorbijganger bevrijdde het kind en de vrouw schreeuwde de longen uit haar lijf.	De voorbijganger bevrijdde het kind, en de vrouw schreeuwde de longen uit haar lijf.	De voorbijganger bevrijdde het kind en de vrouw uit de auto die gekanteld was.
27	De boswachter berispte de padvinder en de hopman doofde gauw het vuurtje met	De boswachter berispte de padvinder, en de hopman doofde gauw het vuurtje met	De boswachter berispte de padvinder en de hopman nadat hij hun vuur had gezien.

	zand.	zand.	
28	De toerist fotografeerde de visser en de reisleider vertelde een verhaal over de visserij.	De toerist fotografeerde de visser, en de reisleider vertelde een verhaal over de visserij.	De toerist fotografeerde de visser en de reisleider zonder hun om toestemming te vragen.
29	De dichter bezong de zwerver en de dronkaard prees luidkeels de schoonheid van Amsterdam.	De dichter bezong de zwerver, en de dronkaard prees luidkeels de schoonheid van Amsterdam.	De dichter bezong de zwerver en de dronkaard met een melodie uit zijn kinderjaren.
30	De professor belde de aannemer en de architect eiste direct een onderzoek naar woningcorporaties.	De professor belde de aannemer, en de architect eiste direct een onderzoek naar woningcorporaties.	De professor belde de aannemer en de architect om over de villa te praten.
31	De klant bedankte de bedrijfsleider en de bediende vroeg de kassabon voor de trui.	De klant bedankte de bedrijfsleider, en de bediende vroeg de kassabon voor de trui.	De klant bedankte de bedrijfsleider en de bediende voor de ruil van haar trui.
32	De lerares begroette de leerling en de moeder beschreef uitvoerig de thuissituatie zonder vader.	De lerares begroette de leerling, en de moeder beschreef uitvoerig de thuissituatie zonder vader.	De lerares begroette de leerling en de moeder in het klaslokaal van haar klas.
33	De pastoor zegende de stuurman en de kapitein bedankte de geestelijke voor zijn zorgen.	De pastoor zegende de stuurman, en de kapitein bedankte de geestelijke voor zijn zorgen.	De pastoor zegende de stuurman en de kapitein voordat het schip in zee voer.
34	De chauffeur vervoerde de baron en de butler bracht de bagage naar het kasteel.	De chauffeur vervoerde de baron, en de butler bracht de bagage naar het kasteel.	De chauffeur vervoerde de baron en de butler na de bruiloft van de prins.
35	De actrice vervloekte de stuntman en de producent gooide woedend zijn dikke sigaar weg.	De actrice vervloekte de stuntman, en de producent gooide woedend zijn dikke sigaar weg.	De actrice vervloekte de stuntman en de producent toen zij op haar kamer zat.
36	De burgemeester ondervroeg de leraar en de onderzoeker onderkende de voordelen van het onderwijsplan.	De burgemeester ondervroeg de leraar, en de onderzoeker onderkende de voordelen van het onderwijsplan.	De burgemeester ondervroeg de leraar en de onderzoeker over de ontwikkeling van jonge schoolkinderen.
37	Het Kamerlid bespotte de interviewer en de minister herhaalde minachtend de vragen van journalisten.	Het Kamerlid bespotte de interviewer, en de minister herhaalde minachtend de vragen van journalisten.	Het Kamerlid bespotte de interviewer en de minister net nadat het interview uitgezonden was.

38	De lijfwacht beschermde de president en de generaal beval direct de omgeving te onderzoeken.	De lijfwacht beschermde de president, en de generaal beval direct de omgeving te onderzoeken.	De lijfwacht beschermde de president en de generaal tegen de bedreiging door de demonstranten.
39	De automobilist raakte de voetganger en de fietser viel geschrokken op de natte straat.	De automobilist raakte de voetganger, en de fietser viel geschrokken op de natte straat.	De automobilist raakte de voetganger en de fietser met zijn onachtzaam geopende deur aan.
40	De tuinman bespiedde het dienstmeisje en de butler pakte verrekijkers om mee te doen.	De tuinman bespiedde het dienstmeisje, en de butler pakte verrekijkers om mee te doen.	De tuinman bespiedde het dienstmeisje en de butler met een verrekijker vanuit de tuin.
41	De clown ontvluchtte de goochelaar en de acrobaat beklom de ladder naar de trapeze.	De clown ontvluchtte de goochelaar, en de acrobaat beklom de ladder naar de trapeze.	De clown ontvluchtte de goochelaar en de acrobaat terwijl het hele publiek hard lachte.
42	De suppoost waarschuwde de student en de studente stopte snel drugs in haar jas.	De suppoost waarschuwde de student, en de studente stopte snel drugs in haar jas.	De suppoost waarschuwde de student en de studente voor de drugssmokkelaar die gezocht werd.
43	De psychiater observeerde de patient en de assistent schreef zorgvuldig de medische gegevens op.	De psychiater observeerde de patient, en de assistent schreef zorgvuldig de medische gegevens op.	De psychiater observeerde de patient en de assistent met bewakingscamera's in de grote behandelkamer.
44	De huisvrouw zoende de kennis en het kind bekeek nieuwsgierig de mensen die voorbijliepen.	De huisvrouw zoende de kennis, en het kind bekeek nieuwsgierig de mensen die voorbijliepen.	De huisvrouw zoende de kennis en het kind howel zij hen niet goed kende.
45	De directeur ontsloeg de werknemer en de afdelingsbaas riskeerde zijn baan door het ontslag.	De directeur ontsloeg de werknemer, en de afdelingsbaas riskeerde zijn baan door het ontslag.	De directeur ontsloeg de werknemer en de afdelingsbaas omdat de hele afdeling gesloten werd.
46	De burgemeester loofde de wethouder en de ondernemer liet meteen een fles cognac bezorgen.	De burgemeester loofde de wethouder, en de ondernemer liet meteen een fles cognac bezorgen.	De burgemeester loofde de wethouder en de ondernemer voor hun inzet voor het weeshuis.
47	De koningin beloonde de lakei en de hofdame kreeg onmiddellijk een kleur van opwinding.	De koningin beloonde de lakei, en de hofdame kreeg onmiddellijk een kleur van opwinding.	De koningin beloonde de lakei en de hofdame allebei met twee zware gouden daalders.

48	De reiziger vervloekte de piloot en de stewardess verzorgde keurig een passagier met hoofdpijn.	De reiziger vervloekte de piloot, en de stewardess verzorgde keurig een passagier met hoofdpijn.	De reiziger vervloekte de piloot en de stewardess na de noodlanding in de zee.
49	De astronaut groette de technicus en de monteur opende behoedzaam de sluis naar binnen.	De astronaut groette de technicus, en de monteur opende behoedzaam de sluis naar binnen.	De astronaut groette de technicus en de monteur voordat hij in de capsule klom.
50	De fan belaagde de drummer en de gitarist riep ontzet naar de ontbrekende beveiliging.	De fan belaagde de drummer, en de gitarist riep ontzet naar de ontbrekende beveiliging.	De fan belaagde de drummer en de gitarist met een mes in de hand.
51	De commissaris bedreigde de parkeerwacht en de rechercheur vertrok waarbij hij de deur dichtsmeet.	De commissaris bedreigde de parkeerwacht, en de rechercheur vertrok waarbij hij de deur dichtsmeet.	De commissaris bedreigde de parkeerwacht en de rechercheur met een aangifte voor illegaal gokken.
52	De archeoloog betaalde de indiaan en de graver stopte alle spullen in een koffer.	De archeoloog betaalde de indiaan, en de graver stopte alle spullen in een koffer.	De archeoloog betaalde de indiaan en de graver uit zijn eigen fonds zonder steun.
53	De dichter belaagde de criticus en de redacteur besloot meteen een rectificatie te plaatsen.	De dichter belaagde de criticus, en de redacteur besloot meteen een rectificatie te plaatsen.	De dichter belaagde de criticus en de redacteur nadat de negatieve kritiek was gepubliceerd.
54	De verpleger verschoonde de junk en de zwerfster waste mopperend haar gezicht met zeep.	De verpleger verschoonde de junk, en de zwerfster waste mopperend haar gezicht met zeep.	De verpleger verschoonde de junk en de zwerfster in het tehuis voor arme mensen.
55	De kapelaan vermaande de koorknaap en het hulpje wist nauwelijks zijn lachen te bedwingen.	De kapelaan vermaande de koorknaap, en het hulpje wist nauwelijks zijn lachen te bedwingen.	De kapelaan vermaande de koorknaap en het hulpje omdat zij 's nachts veel kletsten.
56	De medicijnman besprenkelde de bezetene en het opperhoofd goot voorzichtig olie over het masker.	De medicijnman besprenkelde de bezetene, en het opperhoofd goot voorzichtig olie over het masker.	De medicijnman besprenkelde de bezetene en het opperhoofd met sacraal water uit de bron.
57	De priester offerde de slavin en de slaaf bewierookte dromerig het mysterieuze stenen beeld.	De priester offerde de slavin, en de slaaf bewierookte dromerig het mysterieuze stenen beeld.	De priester offerde de slavin en de slaaf aan zijn goden die zoiets vorderden.
58	De volgeling vereerde de goeroe en de ingewijde	De volgeling vereerde de goeroe, en de ingewijde	De volgeling vereerde de goeroe en de ingewijde zonder

	luisterde ademloos naar zijn gepassioneerde toespraak.	luisterde ademloos naar zijn gepassioneerde toespraak.	hun boek te hebben gelezen.
59	De activist besmeurde de lijfwacht en de officier morste koffie op zijn smetteloze uniform.	De activist besmeurde de lijfwacht, en de officier morste koffie op zijn smetteloze uniform.	De activist besmeurde de lijfwacht en de officier met veel melk van zijn boerderij.
60	De fakir betoverde de toeschouwer en de danseres vertoonde geamuseerd haar tropische sensuele buikdans.	De fakir betoverde de toeschouwer, en de danseres vertoonde geamuseerd haar tropische sensuele buikdans.	De fakir betoverde de toeschouwer en de danseres nadat de olifant eindelijk gekalmeerd was.

Experiment 1: Arithmetic

#	easy	hard
1	1 + 1 + 2 - 1 + 10 + 2 - 1 =	1 + 1 + 2 - 1 + 9 + 1 + 1 =
2	20 - 1- 2 - 1 - 10 + 1 + 1 =	20 - 1 - 2 - 1 - 7 - 2 + 1 =
3	9 + 1 - 1 - 1 + 2 + 10 - 2 =	9 + 1 - 1 - 1 + 6 + 2 + 2 =
4	7 + 2 + 1 + 1 - 1 - 2 - 2 =	7 + 2 + 1 + 1 - 5 + 2 - 2 =
5	10 + 10 - 1 - 1 + 2 - 1 - 1 =	10 + 10 - 1 - 1 + 3 - 2 - 1 =
6	4 - 1 + 1 + 2 + 2 + 2 + 2 =	4 - 1 + 1 + 2 + 7 + 1 - 2 =
7	2 + 2 - 1 + 2 + 1 + 1 + 1 =	2 + 2 - 1 + 2 + 6 - 2 - 1 =
8	18 - 2 - 1 - 1 - 10 + 2 + 1 =	18 - 2 - 1 - 1 - 6 - 2 + 1 =
9	6 + 1 + 2 - 2 + 1 + 1 + 2 =	6 + 1 + 2 - 2 + 5 + 1 - 2 =
10	13 - 2 + 1 + 1 - 1 - 2 - 2 =	13 - 2 + 1 + 1 - 8 + 1 + 2 =
11	1 + 1 + 2 + 1 + 1 + 1 + 2 =	1 + 1 + 2 + 1 + 8 - 2 - 2 =
12	19 - 2 - 1 - 1 - 2 - 2 - 1 =	19 - 2 - 1 - 1 - 7 + 1 + 1 =
13	20 - 1 - 10 - 2 + 1 + 1 + 2 =	20 - 1 - 10 - 2 + 8 - 2 - 2 =
14	7 + 1 + 2 + 2 - 2 - 2 - 2 =	7 + 1 + 2 + 2 - 5 + 1 - 2 =
15	10 + 1 - 2 - 1 + 2 + 1 + 1 =	10 + 1 - 2 - 1 + 5 - 2 + 1 =
16	9 + 1 + 1 + 1 - 10 + 1 + 2 =	9 + 1 + 1 + 1 - 7 - 2 + 2 =

17	11 - 1 - 2 - 2 + 2 + 2 + 1 =	11 - 1 - 2 - 2 + 8 - 2 - 1 =
18	8 + 2 + 2 + 1 - 10 + 2 + 2 =	8 + 2 + 2 + 1 - 7 - 1 + 2 =
19	18 - 10 - 2 - 1 + 1 + 2 + 1 =	18 - 10 - 2 - 1 + 7 - 2 - 1 =
20	6 - 2 - 2 + 10 - 2 - 2 - 2 =	6 - 2 - 2 + 10 - 7 - 1 + 2 =
21	12 - 1 - 10 + 2 + 2 - 2 - 1 =	12 - 1 - 10 + 2 + 8 - 10 + 1 =
22	5 + 10 + 1 + 1 - 1 - 2 - 2 =	5 + 10 + 1 + 1 - 9 + 2 + 2 =
23	3 + 1 + 1 + 1 + 1 + 2 + 1 =	3 + 1 + 1 + 1 + 7 - 2 - 1 =
24	1 + 1 + 2 + 10 - 1 - 2 - 2 =	1 + 1 + 2 + 10 - 8 + 1 + 2 =
25	4 + 1 + 1 + 2 + 1 + 2 + 2 =	4 + 1 + 1 + 2 + 6 + 1 - 2 =
26	7 + 10 + 1 - 2 - 1 - 2 - 1 =	7 + 10 + 1 - 2 - 7 + 2 + 1 =
27	2 + 1 + 1 + 1 + 2 + 2 + 2 =	2 + 1 + 1 + 1 + 7 + 1 - 2 =
28	8 + 1 + 2 + 2 - 1 - 2 - 2 =	8 + 1 + 2 + 2 - 6 - 1 + 2 =
29	10 + 10 - 2 - 10 + 2 + 2 + 2 =	10 + 10 - 2 - 10 + 5 - 1 + 2 =
30	6 - 1 - 1 + 10 - 1 - 2 - 1 =	6 - 1 - 1 + 10 - 6 + 1 + 1 =
31	2 + 2 + 2 + 1 + 2 + 1 + 2 =	2 + 2 + 2 + 1 + 6 + 1 - 2 =
32	7 + 1 + 2 + 2 - 10 + 2 + 2 =	7 + 1 + 2 + 2 - 7 - 1 + 2 =
33	10 + 2 + 2 - 10 + 1 + 1 + 2 =	10 + 2 + 2 - 10 + 8 - 2 - 2 =
34	8 - 1 - 1 + 10 - 2 - 1 - 1 =	8 - 1 - 1 + 10 - 7 + 2 + 1 =
35	1 + 10 - 1 - 2 + 1 + 1 + 2 =	1 + 10 - 1 - 2 + 5 + 1 - 2 =
36	3 + 1 + 10 + 2 - 2 - 1 + 1 =	3 + 1 + 10 + 2 - 9 + 10 - 1 =
37	9 + 1 - 2 - 2 + 1 + 1 + 1 =	9 + 1 - 2 - 2 + 5 - 1 - 1 =
38	4 + 1 + 1 + 10 - 1 - 2 - 1 =	4 + 1 + 1 + 10 - 7 + 2 + 1 =
39	10 - 1 - 1 - 2 + 2 + 1 + 1 =	10 - 1 - 1 - 2 + 7 - 2 - 1 =
40	3 + 1 + 2 + 10 - 1 - 10 + 1 =	3 + 1 + 2 + 10 - 8 - 1 - 1 =

Experiment 2: Language syntax

#	object-relative clause	subject-relative clause
1	Hij weet dat de tantes die de jongen begroette op een eiland zijn geweest.	Hij weet dat de tantes die de jongen begroetten op een eiland zijn geweest.
2	De heel aantrekkelijke vrouwen die de ervaren soldaat omhelsde waren erg geliefd bij iedereen.	De heel aantrekkelijke vrouwen die de ervaren soldaat omhelsden waren erg geliefd bij iedereen.
3	De oude generaals die de kolonel met opzet misleidde genoten bekendheid in de media.	De oude generaals die de kolonel met opzet misleidden genoten bekendheid in de media.
4	De profs die de naieve tegenstander zonder schaamte uitlachte begingen toen een grote vergissing.	De profs die de naieve tegenstander zonder schaamte uitlachten begingen toen een grote vergissing.
5	De vrienden die de zoon op de been hielp lieten hem het gebouw zien.	De vrienden die de zoon op de been hielpen lieten hem het gebouw zien.
6	De twee broers die de iets oudere zuster bijstond stonden bekend om hun succes.	De twee broers die de iets oudere zuster bijstonden stonden bekend om hun succes.
7	Twan hoorde dat de inbrekers die de bandiet doodde zeiden berouwloos te hebben gehandeld.	Twan hoorde dat de inbrekers die de bandiet doodden zeiden berouwloos te hebben gehandeld.
8	De rijke blanken die de eenzame zwarte ongevraagd uitnodigde betaalden voor al het eten.	De rijke blanken die de eenzame zwarte ongevraagd uitnodigden betaalden voor al het eten.
9	De goed betaalde advocaten die de rechter blind vertrouwde namen het meest verstandige besluit.	De goed betaalde advocaten die de rechter blind vertrouwden namen het meest verstandige besluit.
10	De officieren die de luitenant met weinig succes geruststelde werden meegenomen naar het politiebureau.	De officieren die de luitenant met weinig succes geruststelden werden meegenomen naar het politiebureau.
11	De heiligen die de wijze drie jaar lang bewonderde spraken vol lof over hem.	De heiligen die de wijze drie jaar lang bewonderden spraken vol lof over hem.
12	De bazen die de nieuwe werknemer uiteindelijk toch bezocht gingen reizen maken door Europa.	De bazen die de nieuwe werknemer uiteindelijk toch bezochten gingen reizen maken door Europa.
13	De boze bewoners die de agent uit Breda verbaasde schreeuwden kwaad naar de	De boze bewoners die de agent uit Breda verbaasden schreeuwden kwaad naar de

	voorbijgangers.	voorbijgangers.
14	De juffrouwen die de student met veel passie kuste maakten zich uit de voeten.	De juffrouwen die de student met veel passie kusten maakten zich uit de voeten.
15	De meisjes die de zeer oude grootvader graag knuffelde leken zich goed te vermaken.	De meisjes die de zeer oude grootvader graag knuffelden leken zich goed te vermaken.
16	De geestelijken die de huisarts in een restaurant ontving vroegen of alles goed was.	De geestelijken die de huisarts in een restaurant ontvingen vroegen of alles goed was.
17	De jonge meneren uit Amsterdam die de vrouw verleidde kregen ruzie met de politieagent.	De jonge meneren uit Amsterdam die de vrouw verleidden kregen ruzie met de politieagent.
18	De gevaarlijke dieren die de hongerige wilde stiekem besloop kwamen uit een donker bos.	De gevaarlijke dieren die de hongerige wilde stiekem beslopen kwamen uit een donker bos.
19	De dronken gasten die de net getrouwde echtgenoot vermaakte zongen mee met de muziek.	De dronken gasten die de net getrouwde echtgenoot vermaakten zongen mee met de muziek.
20	De artsen die de client zonder veel moeite begreep raadden hem een operatie aan.	De artsen die de client zonder veel moeite begrepen raadden hem een operatie aan.
21	De schattige baby die de ouders bijna niet herkenden groette hen met een glimlach.	De schattige baby die de ouders bijna niet herkende groette hen met een glimlach.
22	De leerling die de leraren en leraressen herhaaldelijk opbelden was lang niet meer gezien.	De leerling die de leraren en leraressen herhaaldelijk opbelde was lang niet meer gezien.
23	De minister die de politici met onverwachte onthullingen choqueerden besloot zijn nevenfunctie niet neer te leggen.	De minister die de politici met onverwachte onthullingen choqueerde besloot zijn nevenfunctie niet neer te leggen.
24	De onbekende vreemde die de omstreden pastoors meermaals benaderden volgde hun en praatte beleefd.	De onbekende vreemde die de omstreden pastoors meermaals benaderde volgde hun en praatte beleefd.
25	De machtige priester die de bisschoppen in Rome toespraken werkte gedreven aan een boek.	De machtige priester die de bisschoppen in Rome toesprak werkte gedreven aan een boek.
26	De man die de directeuren met een brief feliciteerden droeg een nieuw zwart pak.	De man die de directeuren met een brief feliciteerde droeg een nieuw zwart pak.
27	De vriendelijke dame die de kunstenaars	De vriendelijke dame die de kunstenaars

	van harte bedankten klonk erg onder de indruk.	van harte bedankte klonk erg onder de indruk.
28	De vreemde psycholoog die de patientes zonder onderbreking aanstaarden droeg een bril uit Italie.	De vreemde psycholoog die de patientes zonder onderbreking aanstaarde droeg een bril uit Italie.
29	De dichter die de nog onbekende schrijvers publiekelijk aanmoedigden wilde zij beter leren kennen.	De dichter die de nog onbekende schrijvers publiekelijk aanmoedigde wilde zij beter leren kennen.
30	De opscheppende deskundige uit Canada die de Amerikanen corrigeerden wilde de resultaten meteen presenteren.	De opscheppende deskundige uit Canada die de Amerikanen corrigeerde wilde de resultaten meteen presenteren.
31	De uitgeputte chef die de vele klanten kennelijk negeerden maakte een totaal chagrijnige indruk.	De uitgeputte chef die de vele klanten kennelijk negeerde maakte een totaal chagrijnige indruk.
32	De ambtenaar die de kennissen op de straat ontweken leek moe en vooral gestrest.	De ambtenaar die de kennissen op de straat ontweek leek moe en vooral gestrest.
33	De paranoide patiente die de bezoekers uit Groningen uithoorden verkeerde nog in kritieke toestand.	De paranoide patiente die de bezoekers uit Groningen uithoorde verkeerde nog in kritieke toestand.
34	De stoere vent die de jongeren zonder reden aanvielen sloeg snel op de vlucht.	De stoere vent die de jongeren zonder reden aanviel sloeg snel op de vlucht.
35	Journalisten beweerden dat de deelnemer die de medewerkers geloofden niet alle details kon weten.	Journalisten beweerden dat de deelnemer die de medewerkers geloofde niet alle details kon weten.
36	De blonde prins die de grote zwarte paarden vreesden droeg alleen zeer kleurrijke kleding.	De blonde prins die de grote zwarte paarden vreesde droeg alleen zeer kleurrijke kleding.
37	De hoestende jongeman die de dokters in opleiding zagen was eigenlijk helemaal niet ziek.	De hoestende jongeman die de dokters in opleiding zag was eigenlijk helemaal niet ziek.
38	De eigenaar die de arbeiders op het werk stoorden moest zich naar buiten haasten.	De eigenaar die de arbeiders op het werk stoorde moest zich naar buiten haasten.
39	Niemand voorzag dat de dochter die de spelers verbluften graag een nieuwe taal leerde.	Niemand voorzag dat de dochter die de spelers verblufte graag een nieuwe taal leerde.
40	De erg bezige verpleegster die de verwarde zieken zochten had afspraken met de	De erg bezige verpleegster die de verwarde zieken zocht had afspraken met de

	chirurgen.	chirurgen.

Experiment 2: Language semantics

#	semantic garden-path	non-garden-path
1	De kokkin dacht aan het gerecht dat haar aanklaagde voor een zeer ernstig delict.	De kokkin dacht aan het gerechtshof dat haar aanklaagde voor een zeer ernstig delict.
2	De alcoholist had last van zijn kater die miauwde en daar niet mee ophield.	De alcoholist had last van zijn kat die miauwde en daar niet mee ophield.
3	De pianospeler keek naar de toets om hem in te vullen maar hij wist geen antwoorden.	De pianospeler keek naar de Cito-toets om hem in te vullen maar hij wist geen antwoorden.
4	De programmeur liet zijn muis op de tafel rondlopen nadat hij hem had gevoerd.	De programmeur liet zijn veldmuis op de tafel rondlopen nadat hij hem had gevoerd.
5	De generaal vond dat een kwartier niet genoeg plek was voor de groep soldaten.	De generaal vond dat een kamer niet genoeg plek was voor de groep soldaten.
6	De oude prof hield een bal met veel bezoekers, muziek en ook veel eten.	De oude prof hield een feest met veel bezoekers, muziek en ook veel eten.
7	De musicus kende de noten die hij had gegeten omdat hij ook bioloog was.	De musicus kende de kokosnoten die hij had gegeten omdat hij ook bioloog was.
8	De grimmige douanier eiste de tol en ander speelgoed om het te laten checken.	De grimmige douanier eiste de teddybeer en ander speelgoed om het te laten checken.
9	De veer in de vleugel van een grote burcht werd door een schoonmaakster weggeveegd.	De veer in de verdieping van een grote burcht werd door een schoonmaakster weggeveegd.
10	Door verrekijkers konden de jachtopzieners de manen van Jupiter erg duidelijk uit elkaar houden.	Door verrekijkers konden de jachtopzieners de satellieten van Jupiter erg duidelijk uit elkaar houden.
11	Niemand kon vermoeden dat de palm zo erg bloedde dat de hand eraf moest.	Niemand kon vermoeden dat de handpalm zo erg bloedde dat de hand eraf moest.
12	De werklozen zochten een goede baan om te schaatsen en gezellig bier te drinken.	De werklozen zochten een goede ijsbaan om te schaatsen en gezellig bier te drinken.
13	De industrieel wilde een kanaal maken om meer toeschouwers met zijn reclame te	De industrieel wilde een tv-kanaal maken om meer toeschouwers met zijn reclame te

	bereiken.	bereiken.
14	De visser gebruikte een aas om het lange kaartspel te winnen en te beeindigen.	De visser gebruikte een klaverenaas om het lange kaartspel te winnen en te beeindigen.
15	De spin hoorde eigenlijk naar buiten op haar bagagedrager waarop ze vaak boodschappen plaatste.	De snelbinder hoorde eigenlijk naar buiten op haar bagagedrager waarop ze vaak boodschappen plaatste.
16	De bioloog zag het blad, begon het te lezen en vergat zijn dure experiment.	De bioloog zag het weekblad, begon het te lezen en vergat zijn dure experiment.
17	In Hollywood kon je veel sterren in de hemel zien toen de straatverlichting uitviel.	In Hollywood kon je veel sterrenbeelden in de hemel zien toen de straatverlichting uitviel.
18	De kampeerder had veel kolen nodig voor de salade omdat zijn vrouw wilde meeeten.	De kampeerder had veel groenten nodig voor de salade omdat zijn vrouw wilde meeeten.
19	De arts vermoedde dat het kaakje niet goed smaakte want zijn kinderen haatten het.	De arts vermoedde dat het koekje niet goed smaakte want zijn kinderen haatten het.
20	Het vak dat hij moest geven was een erfstuk en deel van een wandkast.	Het schap dat hij moest geven was een erfstuk en deel van een wandkast.
21	De politicus kon de grond voor zijn grote vergissing niet goed aan iedereen uitleggen.	De politicus kon de reden voor zijn grote vergissing niet goed aan iedereen uitleggen.
22	Het rooster was al lang klaar maar de barbecuegasten waren nog steeds niet gearriveerd.	Het grillrooster was al lang klaar maar de barbecuegasten waren nog steeds niet gearriveerd.
23	De cowboy nam de poot van zijn nieuwe bureau en zaagde een stukje eraf.	De cowboy nam de tafelpoot van zijn nieuwe bureau en zaagde een stukje eraf.
24	De automonteur voelde de stroom die rond vijftig volt sterk was en hem blesseerde.	De automonteur voelde de elektriciteit die rond vijftig volt sterk was en hem blesseerde.
25	De dierenoppasser vond de lange slang tussen ander tuingereedschap en begon met het afspuiten.	De dierenoppasser vond de lange tuinslang tussen ander tuingereedschap en begon met het afspuiten.
26	De restaurantbezoeker ging naar het buffet dat van hout was en er duur uitzag.	De restaurantbezoeker ging naar het kastje dat van hout was en er duur uitzag.
27	Het vertrek van zijn overleden moeder was nog kleiner en donkerder dan het zijne.	Het kamertje van zijn overleden moeder was nog kleiner en donkerder dan het zijne.
28	De toneelspeler had een rol nodig voor het	De toneelspeler had een plank nodig voor

	deeg nadat vorige taarten waren mislukt.	het deeg nadat vorige taarten waren mislukt.
29	De beurs die de studente kreeg was een designproduct en was in Italie gemaakt.	De portemonnee die de studente kreeg was een designproduct en was in Italie gemaakt.
30	De matroos keek naar de maten van zijn T-shirts voordat hij op reis ging.	De matroos keek naar de kleuren van zijn T-shirts voordat hij op reis ging.
31	De rijdster koos per ongeluk de vaart die onderstroomde en ging meteen over kop.	De rijdster koos per ongeluk de waterweg die onderstroomde en ging meteen over kop.
32	De wetenschapper had een nieuwe vlam die niet uitblust ontwikkeld en presenteerde hem overal.	De wetenschapper had een nieuw vuur dat niet uitblust ontwikkeld en presenteerde het overal.
33	Het wereldberoemde koor kon helaas niet meer goed gerestaureerd worden zonder miljoenen te investeren.	De wereldberoemde kloostergang kon helaas niet meer goed gerestaureerd worden zonder miljoenen te investeren.
34	De net geboren koe die de boer had gelaten stond op de grote bergweide.	De net geboren koe die de scheet had gelaten stond op de grote bergweide.
35	De zin van zijn leven die hij had gedicht maakte hem in Nederland bekend.	De uitspraak van zijn leven die hij had gedicht maakte hem in Nederland bekend.
36	De jonge patient wilde de pil niet meteen lezen omdat hij over ziektes ging.	De jonge patient wilde de tekst niet meteen lezen omdat hij over ziektes ging.
37	De band die de fietsenmaker met zijn zus voelde werd niet door haar beantwoord.	De bloedband die de fietsenmaker met zijn zus voelde werd niet door haar beantwoord.
38	De politieman zag de aanslag die op zijn tanden zat na het koffie drinken.	De politieman zag de kleur die op zijn tanden zat na het koffie drinken.
39	De Turk constateerde dat de sirene niet meer leefde en Odysseus weer vrij was.	De Turk constateerde dat de toverheks niet meer leefde en Odysseus weer vrij was.
40	Het koetje dat pasgeboren was kon bijna al vliegen hoewel zijn veren kort waren.	Het vogeltje dat pasgeboren was kon bijna al vliegen hoewel zijn veren kort waren.

2|13 Supplementary materials

2|13.2 Exp. 1

2|13.2.1.3 Pre-test: the strength of the language syntax and arithmetic manipulations

Before starting the main experiment we conducted a pre-test with 24 participants who did not take part in experiment 1 (native Dutch speakers, age: *M* = 24.5, *SD* = 7.4, musical training: *M* = 4 years, *SD* = 3.8). The aim was to establish the strength of the difficulty manipulations in the language and the arithmetic tasks. Stimuli were presented as shown for the main experiment with three differences: 1) each trial was followed by a prompt, 2) there was no musical task and music was not presented, 3) after each trial participants had to rate the overall trial difficulty on a seven point Likert scale (1 = very easy, 7 = very difficult).

For the analysis of the language task, a one factor three-level (ambiguous S-coordination, unambiguous S-coordination, NP-coordination) ANOVA of the critical trials' difficulty ratings showed a significant main effect [$F_{(2,46)} = 14.08$, $p < .001$, ${}_{p}\eta^2 = .380$]. Follow-up *t*-tests (Bonferroni corrected) revealed that the ambiguous S-coordination sentences were rated as significantly more difficult (*M* = 2.99, *SD* = 1.02) than the NP-coordination sentences (*M* = 2.46, *SD* = 1.11) [$t_{(23)} = 4.67$, $p < .001$, ${}_{p}\eta^2 = .487$]. The unambiguous S-coordination sentences received intermediate difficulty ratings (*M* = 2.90, *SD* = 1.03) which were not significantly lower than those for ambiguous S-coordinations [$t_{(23)} < 1$] but significantly higher than those for NP-coordinations [$t_{(23)} = 4.51$, $p < .001$, ${}_{p}\eta^2 = .469$]. Thus, the pre-test revealed that the syntactic language manipulation was generally noticeable with an indication that S-coordinations are more difficult in general, even when disambiguated by a comma.

For the analysis of the arithmetic task, a *t*-test revealed a significant difficulty effect (hard: *M* = 3.90, *SD* = 1.35; easy: *M* = 2.99, *SD* = 1.43) [$t_{(23)} = 5.71$, $p < .001$, ${}_{p}\eta^2 = .586$]. In order to compare the effects in the language and the arithmetic tasks we conducted an overall ANOVA with two factors: Task (language, arithmetic) and Difficulty (ambiguous S-coordination/hard, NP-coordination/easy). The arithmetic task was perceived as harder (*M* = 3.45, *SD* = 1.34) than the language task (*M* = 2.72, *SD* = 1.03) [Task, $F_{(1,23)} = 14.48$, $p = .001$, ${}_{p}\eta^2 = .386$]. Furthermore, ambiguous S-coordination/hard trials were generally rated as more difficult (*M* = 3.44, *SD* = 1.10) than NP-coordination/easy trials (*M* = 2.73, *SD* = 1.16) [Difficulty, $F_{(1,23)} = 41.59$, $p < .001$, ${}_{p}\eta^2$

= .644]. However, this difficulty effect was stronger in arithmetic than language, as revealed by a significant interaction effect [$F_{(1,23)} = 5.47$, $p = .028$, ${}_{p}\eta^2 = .192$].

2|13.2.2.2.1 Critical language results including unambiguous S-coordination trials

Unambiguous S-coordination trials, which include a disambiguating comma after the conjunction ‘**en**’ (***and***), were analysed alongside ambiguous S-coordination trials, which are identical to unambiguous S-coordination trials except for the absence of a comma, and NP-coordination trials. We used a 3 (Difficulty: ambiguous S-coordination, unambiguous S-coordination, NP-coordination) × 2 (Key: first key ending, second key ending) within-subjects ANOVA, see Figure 2|S1. There was a difficulty main effect [$F_{(2,106)} = 5.84$, $p = .004$, ${}_{p}\eta^2 = .099$]. Follow-up *t*-tests (Bonferroni corrected) revealed that the garden-path sentences (ambiguous S-coordination) led to fewer high closure ratings ($M = .65$, $SD = .18$) than non-garden-path sentences (NP-coordination) ($M = .72$, $SD = .15$) [$t_{(53)} = 3.46$, $p = .003$, ${}_{p}\eta^2 = .184$]. The unambiguous S-coordination sentences were intermediate in this respect ($M = .69$, $SD = .16$), i.e. not significantly higher than the ambiguous S-coordinations [$t_{(53)} = 1.69$, $p = .294$, ${}_{p}\eta^2 = .051$] and not significantly lower than the NP-coordinations [$t_{(53)} = 1.72$, $p = .276$, ${}_{p}\eta^2 = .053$]. Otherwise, first key endings were less likely ($M = .60$, $SD = .19$) to receive a high closure rating than second-key ratings ($M = .77$, $SD = .13$) [Key, $F_{(1,53)} = 75.64$, $p < .001$, ${}_{p}\eta^2 = .588$]. These two factors did not interact [Key × Difficulty, $F_{(2,106)} < 1$].

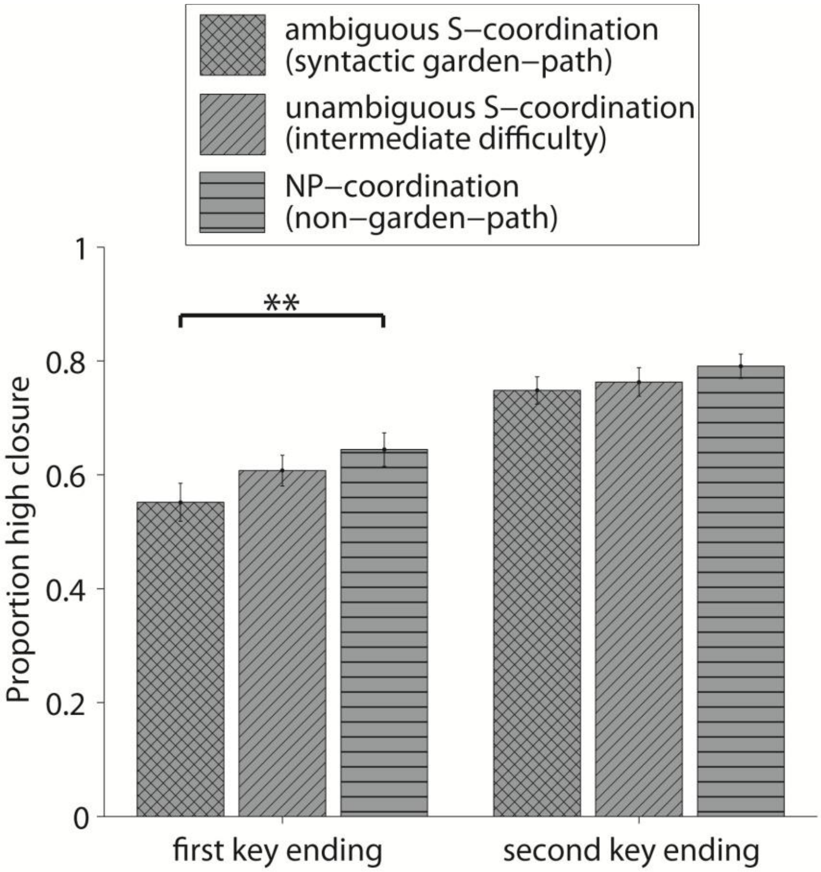


Figure 2|S1. Experiment 1: Closure ratings of critical language trials including intermediate difficulty condition. Error bars = SEM. ** p < .01

2|13.2.3 Analysis of language effect with z-scored digit span as co-variate

The critical closure data of the musico-linguistic part of the experiment were analyzed in 2 (Difficulty: ambiguous S-coordination, NP-coordination) × 2 (Key: first key ending, second key ending) ANCOVAs with *z*-scored digit span as the co-variate. Before standardization, the spread in the digit span data was deemed sufficient to warrant this analysis (overall digit span: *M* = 16.78, *SD* = 3.39; forward digit span: *M* = 8.56, *SD* = 1.88; backward digit span: *M* = 8.20, *SD* = 2.29). Using overall digit span as the co-variate revealed the previously seen effects of Difficulty [$F_{(1,52)} = 11.76$, $p = .001$, ${}_p\eta^2 = .184$] and Key [$F_{(1,52)} = 50.77$, $p < .001$, ${}_p\eta^2 = .494$]. These two factors did not interact [Key × Difficulty, $F_{(1,52)} = 1.82$, $p = .184$, ${}_p\eta^2 = .034$]. None of these effects interacted with overall digit span [Difficulty × Digit Span, $F_{(1,52)} < 1$] [Key × Digit Span, $F_{(1,52)} < 1$] [Key × Difficulty × Digit Span, $F_{(1,52)} = 1.45$, $p = .234$, $_{p}\eta^{2} = .027$].

Using forward digit span as the co-variate similarly revealed the previously seen effects of Difficulty [$F_{(1,52)} = 12.31$, $p = .001$, $_{p}\eta^{2} = .191$] and Key [$F_{(1,52)} = 50.95$, $p < .001$, $_{p}\eta^{2} = .495$]. These two factors did not interact [Key × Difficulty, $F_{(1,52)} = 1.81$, $p = .184$, $_{p}\eta^{2} = .034$]. None of these effects interacted with forward digit span [Difficulty × Digit Span, $F_{(1,52)} = 2.44$, $p = .124$, $_{p}\eta^{2} = .045$] [Key × Digit Span, $F_{(1,52)} < 1$] [Key × Difficulty × Digit Span, $F_{(1,52)} = 1.32$, $p = .255$, $_{p}\eta^{2} = .025$].

Using backward digit span as the co-variate similarly revealed the previously seen effects of Difficulty [$F_{(1,52)} = 12.07$, $p = .001$, $_{p}\eta^{2} = .188$] and Key [$F_{(1,52)} = 50.91$, $p < .001$, $_{p}\eta^{2} = .495$]. These two factors did not interact [Key × Difficulty, $F_{(1,52)} = 1.79$, $p = .187$, $_{p}\eta^{2} = .033$]. None of these effects interacted with backward digit span [Difficulty × Digit Span, $F_{(1,52)} = 1.34$, $p = .252$, $_{p}\eta^{2} = .025$] [Key × Digit Span, $F_{(1,52)} < 1$] [Key × Difficulty × Digit Span, $F_{(1,52)} < 1$].

2|13.3 Exp. 2

2|13.3.1.3 Pre-test 1: sentence completions show the intended misinterpretation of semantic material

In a first pre-test we tested 54 participants who did not take part in experiment 2 (native Dutch speakers, age: $M = 21.7$, $SD = 3.3$, musical training: $M = 4.2$ years, $SD = 4.1$). Their task was to complete sentence beginnings (words one to eight, e.g., 'De programmeur liet zijn **muis** op de tafel …', *The programmer let his **mouse** …*) based on semantic items or filler items. Sentence beginnings were presented fully while the subject typed in a possible sentence ending. Half the sentence beginnings were based on filler items. Subsequently, sentence completions were rated independently by two blind raters (both Dutch native speakers) as to the interpretation of the manipulated word ('**muis**'/'**veldmuis**', ***mouse***/***field vole***: sentence completion indicative of an animal or not?). Many sentence completions (22%) were not rateable according to the raters (uninformative or absent sentence completions) and the two raters sometimes (18% or trials) disagreed about this. When both agreed that an item was rateable their agreement was 99.4%. Any disagreements were resolved by a third blind rater (also a Dutch native speaker). Garden-path and non-garden-path sentence beginnings did not differ according to the rateability of the completions [$t_{(53)} < 1$]. However, using the rateable trials only, the intended word interpretation was overwhelmingly adopted in the non-garden-path condition (93.6%) while this hardly happened in the semantic garden-path condition (13.3%) [$t_{(53)} = 52.18$, $p < .001$, ${}_{p}\eta^2 = .981$].

2|13.3.1.4 Pre-test 2: the strength of the syntactic and semantic manipulations

The aim of the second pre-test was to establish the strength of the difficulty manipulations of the syntactic and the semantic items. We tested 24 participants who did not take part in the first pre-test or the main experiment (native Dutch speakers, age: $M = 21.8$, $SD = 3.1$, musical training: $M = 4$ years, $SD = 4.2$). Stimuli were presented as shown for the main experiment with five differences: 1) the length of word presentation was controlled by the participant (self-paced reading) giving us an online measure of processing difficulty, 2) each trial was followed by a prompt, 3) there was no musical task and music was not presented, 4) after each sentence participants had to rate the acceptability of the sentence (data not presented here), 5) after each trial participants had to rate the overall trial difficulty on a seven point Likert scale (1 = very easy, 7 = very difficult).

For the analysis of the language syntax part, we first analysed difficulty ratings. As expected, object-relative clause trials were rated as more difficult ($M = 3.24$, $SD = 0.93$) than subject-relative clauses ($M = 2.89$, $SD = 0.80$) [$t_{(23)} = 3.96$, $p = .001$, ${}_{p}\eta^2 = .405$]. In order to analyse the reading time data we took the natural logarithm of reading times and defined a trial as an outlier if it fulfilled any of the following three criteria: (1) shorter than 50ms, (2) longer than 2500ms, (3) longer than 2.5 *SD* above the sentence-position-specific mean (only defined as an outlier if this criterion is reached both in analysis by items and by subjects). Outlier values (3.13% of all reading times, 5.98% of critical word reading times) were replaced by the cut-off value. The eighth word was the pre-critical word and as expected it did not show a significant difference between syntax conditions (OR: $M = 509$ ms; SR: $M = 505$ ms)[2] [$t_{(23)} < 1$]. However, both the critical ninth word [$t_{(23)} = 2.83$, $p = .009$, ${}_{p}\eta^2 = .259$] and the post-critical tenth word [$t_{(23)} = 2.91$, $p = .008$, ${}_{p}\eta^2 = .269$] were read longer in the object-relative clause condition (critical: M =

---

[2] For ease of interpretation we report the mean logged values as $e^{\ln(\text{value})}$, i.e. de-logged into standard milliseconds.

723 ms; post-critical: *M* = 665 ms) than in the subject-relative clause condition (critical: *M* = 659 ms; post-critical: *M* = 607 ms).

Carrying out the same analyses for the language semantics part, we also found a significant effect of semantic manipulation on trial difficulty ratings [$t_{(23)}$ = 3.33, *p* = .003, $_{p}\eta^{2}$ = .325], indicating that semantic garden-path sentences were perceived as more difficult (*M* = 2.58, *SD* = 0.90) than the non-garden-path sentences (*M* = 2.26, *SD* = 0.74). The reading time analysis did not reveal a significant difference between conditions on the pre-critical word (GP: *M* = 413 ms; non-GP: *M* = 410 ms) [$t_{(23)}$ < 1], while for both the critical word [$t_{(23)}$ = 3.12, *p* = .005, $_{p}\eta^{2}$ = .297] as well as the post-critical word [$t_{(23)}$ = 4.64, *p* < .001, $_{p}\eta^{2}$ = .484] the same words were read longer in the semantic garden-path condition (critical: *M* = 526 ms; post-critical: *M* = 533 ms) than in the non-garden-path condition (critical: *M* = 482 ms; post-critical: *M* = 448 ms).

In order to compare the effects of the syntax and the semantics manipulations we conducted ANOVAs with two factors: Manipulation (syntax, semantics) and Difficulty (object-relative clause/semantic garden-path, subject-relative clause/non-garden-path). The trial difficulty ratings showed that syntax trials were generally perceived as harder (*M* = 3.06, *SD* = 0.84) than semantics trials (*M* = 2.42; *SD* = 0.79) [Manipulation, $F_{(1,23)}$ = 20.39, *p* < .001, $_{p}\eta^{2}$ = .470] and syntactically or semantically more challenging trials were rated as more difficult (*M* = 2.91, *SD* = 0.81) than less challenging trials (*M* = 2.57, *SD* = 0.69) [Difficulty, $F_{(1,23)}$ = 23.73, *p* < .001, $_{p}\eta^{2}$ = .508]. However, crucially, these two factors did not interact [$F_{(1,23)}$ < 1].

The pattern in the reading time data is similar. For the pre-critical word reading time analysis, words in the syntax part were read longer (*M* = 507 ms) than in the semantics part (*M* = 411 ms) [Manipulation, $F_{(1,23)}$ = 25.75, *p* < .001, $_{p}\eta^{2}$ = .528]. As expected, the difficulty effect as well as the interaction were non-significant [Difficulty, $F_{(1,23)}$ < 1] [Manipulation × Difficulty, $F_{(1,23)}$ < 1]. For the critical word reading time analysis, words in the syntax part were still read longer (syntax: *M* = 690 ms; semantics: *M* = 504 ms) [Manipulation, $F_{(1,23)}$ = 18.85, *p* < .001, $_{p}\eta^{2}$ = .450]. Also, words in more challenging sentences were read longer (OR/GP: *M* = 616 ms; SR/non-GP: *M* = 564 ms) [Difficulty, $F_{(1,23)}$ = 15.77, *p* = .001, $_{p}\eta^{2}$ = .407]. However, these two factors did not interact [Manipulation × Difficulty, $F_{(1,23)}$ < 1]. For the post-critical word, words in the syntactic part were still read longer (syntax: *M* = 636 ms; semantics: *M* = 489 ms) [Manipulation, $F_{(1,23)}$ = 27.27, *p* < .001, $_{p}\eta^{2}$ = .542]. Furthermore, there was an effect of difficulty (OR/GP: *M* = 596 ms; SR/non-GP: *M* = 521 ms) [$F_{(1,23)}$ = 27.39, *p* < .001, $_{p}\eta^{2}$ = .544]. These two factors marginally interacted [Manipulation × Difficulty, $F_{(1,23)} = 3.05$, $p = .094$, ${}_{p}\eta^2 = .117$] indicating that the semantic garden-path effect was slightly greater than the syntax effect in the reading times of the post-critical word.

2|13.4.1.1 Combined analysis of syntax effect in experiments 1 and 2 with musical training as co-variate

A 2 (Experiment: one, two) × 2 (Difficulty: ambiguous S-coordination/object-relative clause, NP-coordination/subject-relative clause) × 2 (Key: first key ending, second key ending) mixed between- and within-subjects ANCOVA with *z*-scored musical training as a co-variate exhibited the same significant effects seen in the equivalent ANOVA without a co-variate [Difficulty, $F_{(1,113)} = 17.47$, $p < .001$, ${}_{p}\eta^2 = .134$] [Key, $F_{(1,113)} = 89.36$, $p < .001$, ${}_{p}\eta^2 = .442$] [Difficulty × Key, $F_{(1,113)} = 4.19$, $p = .043$, ${}_{p}\eta^2 = .036$]. The factor experiment was still without effect [Experiment, $F_{(1,113)} < 1$] [Difficulty × Experiment, $F_{(1,113)} = 1.38$, $p = .243$, ${}_{p}\eta^2 = .012$] [Key × Experiment, $F_{(1,113)} < 1$] [Difficulty × Key × Experiment, $F_{(1,113)} < 1$]. Musical training did not modulate any of these effects [Musical Training, $F_{(1,113)} = 1.19$, $p = .278$, ${}_{p}\eta^2 = .010$] [Difficulty × Musical Training, $F_{(1,113)} < 1$] [Key × Musical Training, $F_{(1,113)} = 1.05$, $p = .307$, ${}_{p}\eta^2 = .009$] [Difficulty × Key × Musical Training, $F_{(1,113)} = 1.61$, $p = .207$, ${}_{p}\eta^2 = .014$].

2|13.4.1.2 Results by items

In order to check whether the syntax effect does not only generalise across subjects but also across items, we further analysed the data with linguistic items (*F*2-analysis) and musical items (*F*3-analysis) as random factors. That is, we averaged music ratings not by participants, but instead across participants by items.

Experiment 1. Two 2 (Task: language, arithmetic) × 2 (Difficulty: ambiguous S-coordination/hard, NP-coordination/easy) × 2 (Key: first key ending, second key ending) ANOVAs were performed. Task is a between-items factor in the analysis by linguistic items (*F*2-analysis) but a within-items factor in the analysis by musical items (*F*3-analysis). All other factors were within-items in both ANOVAs. The analyses exhibited a significantly greater difficulty effect in the language task than the arithmetic task by linguistic items only [Task × Difficulty, $F2_{(1,98)} = 6.25$, $p = .014$, ${}_{p}\eta^{2} = .060$; $F3_{(1,9)} = 3.58$, $p = .091$, ${}_{p}\eta^{2} = .284$]. Otherwise, the Task factor showed a main effect by musical items only [Task, $F2_{(1,98)} = 3.24$, $p = .075$, ${}_{p}\eta^{2} = .032$; $F3_{(1,9)} = 15.25$, $p = .004$, ${}_{p}\eta^{2} = .629$], while the other two main effects were significant by linguistic and musical items [Difficulty, $F2_{(1,98)} = 10.32$, $p = .002$, ${}_{p}\eta^{2} = .095$; $F3_{(1,9)} = 5.38$, $p = .046$, ${}_{p}\eta^{2} = .374$] [Key, $F2_{(1,98)} = 40.50$, $p < .001$, ${}_{p}\eta^{2} = .292$; $F3_{(1,9)} = 19.41$, $p = .002$, ${}_{p}\eta^{2} = .683$]. The Key and Task factors interacted by musical items only [$F2_{(1,98)} = 2.57$, $p = .112$, ${}_{p}\eta^{2} = .026$; $F3_{(1,9)} = 9.14$, $p = .014$, ${}_{p}\eta^{2} = .504$] while Key and Difficulty did not [$F2_{(1,98)} < 1$; $F3_{(1,9)} < 1$]. The three-way interaction was not significant [Task × Difficulty × Key, $F2_{(1,98)} < 1$; $F3_{(1,9)} = 1.33$, $p = .278$, ${}_{p}\eta^{2} = .129$].

Experiment 2. Two 2 (Manipulation: syntax, semantics) × 2 (Difficulty: object-relative clause/semantic garden-path, subject-relative clause/non-garden-path) × 2 (Key: first key ending, second key ending) ANOVAs were performed. As in experiment 1, Task is a between-items factor in the analysis by linguistic items (*F*2-analysis) but a within-items factor in the analysis by musical items (*F*3-analysis). All other factors were within-items in both ANOVAs. The analyses exhibited a significantly greater difficulty effect in the syntax part than in the semantics part by musical items only [Manipulation × Difficulty, $F2_{(1,78)} = 2.71$, $p = .103$, ${}_{p}\eta^{2} = .034$; $F3_{(1,9)} = 15.07$, $p = .004$, ${}_{p}\eta^{2} = .626$]. Otherwise, two of the three factors showed a main effect [Manipulation, $F2_{(1,78)} = 6.57$, $p = .012$, ${}_{p}\eta^{2} = .078$; $F3_{(1,9)} = 5.71$, $p = .041$, ${}_{p}\eta^{2} = .388$] [Difficulty, $F2_{(1,78)} < 1$; $F3_{(1,9)} < 1$] [Key, $F2_{(1,78)} = 132.78$, $p < .001$, ${}_{p}\eta^{2} = .630$; $F3_{(1,9)} = 10.70$, $p = .010$, ${}_{p}\eta^{2} = .543$]. The Key and Manipulation factors did not interact [$F2_{(1,78)} < 1$; $F3_{(1,9)} < 1$], neither did the Key and Difficulty factors [$F2_{(1,78)} = 1.35$, $p = .250$, ${}_{p}\eta^{2} = .017$; $F3_{(1,9)} < 1$]. The three-way interaction was not significant either [Manipulation × Difficulty × Key, $F2_{(1,78)} = 1.01$, $p = .317$, ${}_{p}\eta^{2} = .013$; $F3_{(1,9)} = 1.89$, $p = .201$, ${}_{p}\eta^{2} = .174$].

Experiments 1 and 2 combined. Two 2 (Experiment: one, two) × 2 (Difficulty: ambiguous S-coordination/object-relative clause, NP-coordination/subject-relative clause) × 2 (Key: first key ending, second key ending) ANOVAs were performed. Experiment is a between-items factor in the analysis by linguistic items (*F*2-analysis) but a within-items factor in the analysis by musical items (*F*3-analysis). All other factors were within-items in both ANOVAs. The analyses exhibited the main effects of Difficulty and Key which we observed before [Difficulty, $F2_{(1,98)} = 10.81$, $p = .001$, ${}_{p}\eta^{2} = .099$; $F3_{(1,9)} = 10.63$, $p = .010$, ${}_{p}\eta^{2} = .541$] [Key, $F2_{(1,98)} = 82.09$, $p < .001$, ${}_{p}\eta^{2} = .456$; $F3_{(1,9)} = 19.63$, $p = .002$, ${}_{p}\eta^{2} = .686$]. The factor Experiment was without effect [Experiment, $F2_{(1,98)}$

= 2.25, p = .137, $_p\eta^2$ = .022; $F3_{(1,9)}$ = 1.71, p = .223, $_p\eta^2$ = .160] [Difficulty × Experiment, $F2_{(1,98)}$ = 1.05, p = .308, $_p\eta^2$ = .011; $F3_{(1,9)}$ < 1] [Key × Experiment, $F2_{(1,98)}$ < 1; $F3_{(1,9)}$ = 1.12, p = .318, $_p\eta^2$ = .110] [Difficulty × Key × Experiment, $F2_{(1,98)}$ < 1; $F3_{(1,9)}$ < 1]. The Difficulty × Key interaction was significant by musical items only [$F2_{(1,98)}$ = 1.87, p = .175, $_p\eta^2$ = .019; $F3_{(1,9)}$ = 5.17, p = .049, $_p\eta^2$ = .365].

2|13.6 Post-test: Closure ratings without concurrent task

In order to check whether the double-task paradigm we chose led to unusual music ratings, we ran a post-test on the music material only. Sixty new participants (age: M = 23.1, SD = 4.4; musical training: M = 5.1 years, SD = 5.3) rated the closure of 80 music sequences, as done in each block of experiment 1 or experiment 2. However, there was no concurrent task in the visual modality. The procedure and stimuli were otherwise the same as in the experiments.

As can be seen in Figure 2|S2, the general pattern in this post-test mirrors what is seen in the experiments. All participants gave more high closure ratings to authentic cadence endings than no-cadence endings ($M_{\text{authentic cadence}}$ = .72 > $M_{\text{no cadence}}$ = .09, difference SD = .63). There were also more high closure ratings for second-key endings (M = .60, SD = .15) than first-key endings (M = .56, SD = .15). However, as opposed to the results of experiments 1 and 2, this difference was not significant [$t_{(59)}$ = 1.21, p = .232, $_p\eta^2$ = .024]. As shown in Figure 2|S2, this likely reflects the unexpectedly low number of high closure ratings of second key endings, compared to experiments 1 and 2. The implications for the role of split attention on the ability to integrate chords into a new harmonic key are arguably beyond the scope of this paper. Here, we just want to point out that our claims on shared syntactic processing in music and language are based on first-key endings which appear unaffected by the difference between split-attention (experiments 1 and 2) and full attention (post-test).

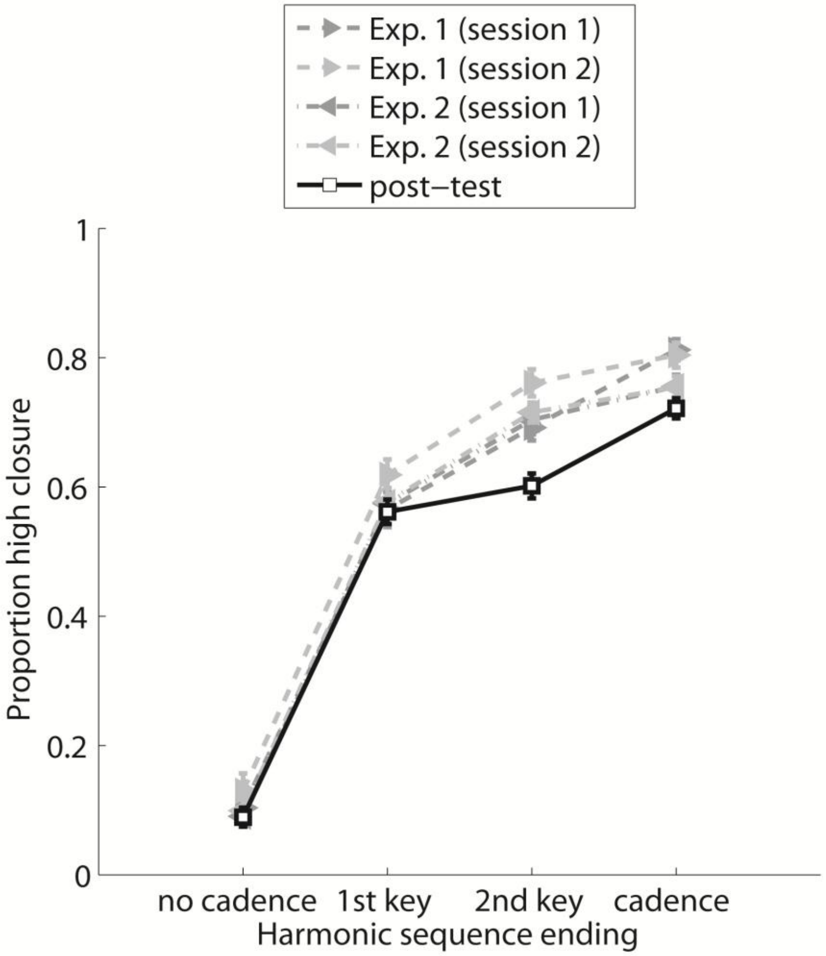


Figure 2|S2. Post-test: Closure ratings of music material without second task. Error bars = SEM.

2|13.7 Additional analyses of experiments 1 & 2: does task order influence the results?

An anonymous reviewer asked whether the order of language syntax and control trials influenced the findings we present. It should be noted that task order was counter-balanced. Therefore, a systematic influence is very unlikely. Still, for the sake of completeness, we present the analysis of music closure ratings with the additional within-subjects factor Order.

The ANOVA of the language block of experiment 1 (see *section 2|2.2.1*) with the additional factor Order reveals the same significant effects as before (Difficulty, Key; *p*s < .01). The factor Order was not significant, neither as a main effect [$F_{(1,52)} = 3.41$, $p = .071$, ${}_p\eta^2 = .061$], nor as an interaction with any other effect (*p*s > .1). The ANOVA of the syntax block of experiment 2 (see *section 2|3.2.2.1*) with the additional factor Order reveals the same significant effects as before (Difficulty, Key; *p*s < .05). The factor Order was not significant, neither as a main effect [$F_{(1,60)} < 1$], nor as an interaction with any other effect (*p*s > .2).

We also carried out an ANOVA of the combined syntax data of experiments 1 and 2 with the factors Order (language syntax trials run first or second), Experiment (experiment 1, experiment 2), Difficulty (ambiguous S-coordination/object-relative clause, NP-coordination/subject-relative clause) and Key (first key ending, second key ending). The first two factors were between-subjects, the other two factors were within-subjects. The significant effects reported in the main analysis without the factor Order were still significant (Difficulty, Key, Difficulty × Key; *p*s < .05). The factor Order was marginally significant ($F_{(1,112)} = 3.06$, $p = .083$, ${}_p\eta^2 = .027$), indicating a slightly higher proportion of high closure ratings in trials run after the control task ($M = .695$) compared to before ($M = .650$). No other effect was significant (*p*s > .2). This shows that task order did not interact with any of the effects of interest.

Chapter 3

Music and language syntax interact in Broca's area: an fMRI study

Instrumental music and language are both syntactic systems, employing complex, hierarchically-structured sequences built using implicit structural norms. This organization allows listeners to understand the role of individual words or tones in the context of an unfolding sentence or melody. Previous studies suggest that the brain mechanisms of syntactic processing may be partly shared between music and language. However, functional neuroimaging evidence for anatomical overlap of brain activity involved in linguistic and musical syntactic processing has been lacking. In the present study we used functional magnetic resonance imaging (fMRI) in conjunction with an interference paradigm based on sung sentences. We show that the processing demands of musical syntax (harmony) and language syntax interact in Broca's area in the left inferior frontal gyrus (without leading to music and language main effects). A language main effect in Broca's area only emerged in the complex music harmony condition, suggesting that (with our stimuli and tasks) a language effect only becomes visible under conditions of increased demands on shared neural resources. In contrast to previous studies, our design allows us to rule out that the observed neural interaction is due to: (1) general attention mechanisms, as a psychoacoustic auditory anomaly behaved unlike the harmonic manipulation, (2) error processing, as the language and the music stimuli contained no structural errors. The current results thus suggest that two different cognitive domains - music and language - might draw on the same high level syntactic integration resources in Broca's area.

Kunert, R., Willems, R.M., Casasanto, D., Patel, A.D., Hagoort, P. (2015). Music and language syntax interact in Broca's area: an fMRI study. *PLoS ONE, 10* (11), e0141069. doi:10.1371/journal.pone.014106

3|1 Introduction

Music and language are uniquely human abilities which, despite their obvious differences, appear to share more than just a common population of users. Specifically, it has been proposed that one overlapping aspect is found in syntactic processing (Patel, 2003). Syntactic processing – whether in language or in music – involves the integration of discrete elements (e.g., words, tones/chords) into higher order structures (e.g., sentences in language and harmonic sequences in music) according to a set of combinatorial principles that are implicitly understood by members of a culture (Patel, 2003). Using functional magnetic resonance imaging (fMRI), the present study aimed to find neural evidence for shared syntactic integration resources recruited by both music and language.

In the present study we defined music syntax processing as harmonic structure processing, in line with many previous studies (e.g., Fedorenko et al., 2009; Patel et al., 1998). Harmony in Western tonal music refers to the organization of pitches in terms of scales, chords, and keys. The basic 'pitch material' of Western tonal/harmonic music (henceforth, tonal music) consists of 12 pitches per octave, each representing one of 12 octave-equivalent 'pitch classes' (e.g., all the C-notes on a piano keyboard). When playing in a musical 'key', a subset of 7 out of 12 pitch classes (in-key tones) is emphasized. Therefore, once a listener has derived a sense of key, e.g., C-major, from a musical piece (for a computational model see Tillmann et al., 2000) she or he expects certain tones – for example in-key tones such as C – more strongly than others – out-of-key tones such as C# (Krumhansl, 1979; Krumhansl & Kessler, 1982). Thus, in tonal music, incoming tones are evaluated in terms of a harmonic framework into which they are continuously integrated.

Do musical and linguistic syntactic processing overlap in the brain? On the one hand, it is known that sensitivity to linguistic syntax and to tonal harmony can dissociate after brain damage, suggesting independence of these two domains (e.g., Peretz, 1993). On the other hand, there is evidence that linguistic syntactic processing and tonal harmonic processing involve similar brain responses (Koelsch, Gunter, Friederici, & Schröger, 2000; Maess et al., 2001; Musso et al., 2015; Patel et al., 1998; for a review see Patel, 2013). To resolve this paradox, the 'Shared syntactic integration resource hypothesis' or SSIRH (Patel, 2003) posited a distinction between domain-specific representations in long-term memory (e.g., stored knowledge of words and their syntactic features, and of chords and their harmonic features) and shared neural resources which act upon these representations as part of structural processing. This "dual-system" model considers syntactic processing to involve the interaction (via long-distance neural connections) of "resource networks" (hypothesized in frontal brain regions) and "representation networks" (hypothesized in temporal brain regions). Patel (2003) posited that resource networks are recruited when structural integration of incoming elements in a sequence is costly; that is, when it involves the rapid and selective activation of low-activation items in representation networks. Cognitive theories of syntactic processing in language (dependency locality theory; Gibson, 2000) and of tonal harmonic processing in music (tonal pitch space theory; Lerdahl, 2001) were used to specify the notion of processing cost.

In both models, incoming elements incur large processing (activation) costs when they need to be mentally connected to existing elements from which they are "distant" in a cognitive sense (e.g., in music, distant in tonal pitch space rather than in terms of physical distance in Hz; in language, distant in terms of the number of intervening words between a syntactic head and the to-be-integrated word). According to the SSIRH, in such circumstances, activity in frontal brain regions increases in order to rapidly activate specific low-activation representations in temporal regions via reentrant connections. Put another way, music and language share limited neural resources in frontal brain regions for the activation of stored structural information in temporal brain regions (for a similar model specific to language see Hagoort, 2005, 2013).

The SSIRH predicts that since neural resources for structural integration are limited, simultaneous costly integrations in harmony and language should lead to interference. Testing this prediction requires experiments which present music and language simultaneously, and which align points of difficult structural integration in the two domains. This prediction has been supported in several studies which presented chord sequences and sentences (two using ERPs: Koelsch, Gunter, et al., 2005; Steinbeis & Koelsch, 2008; and two using behavioral methods: Hoch et al., 2011; Slevc et al., 2009) or melodies and sentences (one using ERPs: Carrus et al., 2013; and one using behavioral methods: Fedorenko et al., 2009), see Kunert & Slevc (2015) for an overview. For example, the behavioral study of Fedorenko et al. (2009) (which informed the design of the current neural study) manipulated linguistic syntactic integration difficulty via the distance between dependent words. These researchers manipulated the structure of embedded relative clauses as shown below (italicized):

a) The boy *that helped the girl* got an "A" on the test.

b) The boy *that the girl helped* got an "A" on the test.

The sentences were sung to melodies (one note per word) which did or did not contain an out-of-key note on the last word of the relative clause: 'girl' in (a), 'helped' in (b). According to dependency locality theory (Gibson, 2000), this word is associated with a distant structural integration in (b) (between 'helped' and 'that') but not in (a). A control condition was included for an attention-getting but non-harmonically deviant musical event: a 10 dB increase in volume on the last word of the relative clause. After each sentence, participants were asked a comprehension question, and accuracy was assumed to reflect processing difficulty. The results revealed an interaction between musical and linguistic processing: comprehension accuracy was lower for sentences with distant versus local syntactic integrations (as expected), but crucially, this difference was larger when melodies contained an out-of-key note. The control condition (loud note) did not produce this effect: the difference between the two sentence types was of the same size as that in the conditions which did not contain an out-of-key note.

However, the brain areas underlying such interaction effects are unclear. Overall, a great number of brain lesion, electrophysiological and hemodynamic brain imaging studies converge in highlighting one key region for syntax processing in either music or language when studied separately: Broca's area (Bookheimer, 2002; Drai & Grodzinsky, 2006; Friederici, Wang, Herrmann, Maess, & Oertel, 2000; Koelsch, Fritz, Schulze, Alsop, & Schlaug, 2005; Maess et al., 2001; Sammler et al., 2011). Thus this region may be the locus of the interaction effect, either in the left hemisphere and/or in the right hemisphere homologue of this area (Embick, Marantz, Miyashita, O'Neil, & Sakai, 2000; Friederici et al., 2000; Koelsch, Fritz, et al., 2005; Maess et al., 2001; Tillmann, Koelsch, et al., 2006).

In searching for interactions between language and music in Broca's area, the current study was mindful of a confound identified by Rogalsky et al. (2011). Many previous experiments using brain measures have operationalized syntactically challenging processing in language as syntactic violation processing (Carrus et al., 2013; Koelsch, Gunter, et al., 2005;

Maidhof & Koelsch, 2011; Steinbeis & Koelsch, 2008). Therefore, general error processing may be shared between music and language, rather than syntactic processing. We used a language manipulation and a music manipulation which did not involve syntactic violations.

Motivated by the hypothesis that Broca's area was a neural site of interaction between linguistic and musical syntactic processing, the present study specifically focused on the activation pattern of Broca's area and its right hemisphere homologue in response to structural manipulations of music and language. Participants heard songs containing either a syntactically easy construction containing only a local dependency (SR: subject-extracted relative clause) or a difficult construction containing a non-local dependency (OR: object-extracted relative clause; see Gibson, 1998). Sentences were sung *a cappella* and the critical word which disambiguated between these two linguistic options was either sung on a regular tone (in-key tone which is easy to integrate in the prevailing harmonic context) or on an irregular tone (out-of-key tone which is not easy to integrate harmonically). Thus, the time point of integration difficulty in music was aligned with the one in language.

Note that neither integration difficulty involved errors. Both types of sentences used in the current study were fully grammatical, and differed in syntactic complexity. Similarly, the use of an out-of-key tone in some of the musical melodies increased their complexity in terms of tonal-harmonic structure (Eerola, Himberg, Toiviainen, & Louhivuori, 2006), but such tones would not be considered 'errors' because they are common stylistic elements in tonal melodies. For example, the melodies of Schubert's lieder often contain out-of-key notes, which are considered to play an important role in the pattern of tension and resolution within the melodies (Lerdahl, 2013).

As noted above, a previous behavioral study in English using a similar design showed an interaction between linguistic and musical conditions in terms of sentence comprehension (Fedorenko et al., 2009). As in that study, we included a control condition involving a non-syntactic auditory anomaly – presenting the critical tone in-key but 10dB SPL louder – in order to rule out the possibility that any acoustic irregularity would elicit the predicted interaction. (This loudness increment was identical to that used in Fedorenko et al. (2009).)

It was hypothesized that Broca's area would be sensitive to the increased processing difficulty of a concurrent syntactic integration challenge in both music and language.

Furthermore, this brain area is not predicted to be sensitive to the interaction between language syntax and a perceptually salient loudness increase at the critical sentence position, as the latter is not syntactic in nature but instead merely acoustic.

3|2 Materials & Methods

3|2.1 Ethics Statement

Written informed consent was obtained from all participants prior to measurement and the study received ethical approval from the local reviewing committee ‘‘CMO Arnhem Nijmegen’’ (CMO no 2001/095 and amendment ‘‘Imaging Human Cognition’’ 2006, 2008), in accordance with the Research involving human subjects Act, following the principles of the Declaration of Helsinki.

3|2.2 Participants

19 healthy participants were included in the final analysis (mean age = 22 years, range 18 – 27). No subject had a known history of neurological, language related or hearing problems and all had normal or corrected-to-normal vision. Five additional participants were excluded due to technical difficulties or excessive movement. The remaining 7 men and 12 women were all right handed, native speakers of Dutch with little formal musical training (mean training = 1.9 years, *SD* = 2.3). All were naïve as to the purpose of the study and were paid for their participation.

3|2.3 Stimuli

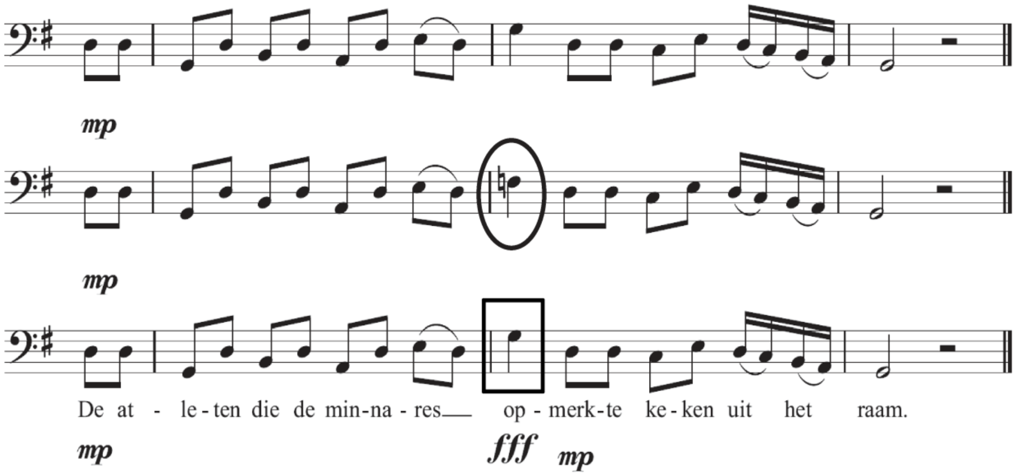


Figure 3|1. Example stimuli. The top melody shows the in-key condition in which no note is out-of-key (all notes are in G-major). The middle melody shows the out-of-key condition in which only the tone coinciding with the stressed syllable of the relative clause verb (circled) is out-of-key. The bottom melody shows the auditory anomaly condition in which all notes are in G-major but the critical tone is 10dB louder (boxed). The lowest pitch used across all melodies was F#2 (92.5 Hz) and the highest was E4 (329.6 Hz). The Dutch sentence in the figure means: *The athletes that the mistress noticed looked out of the window.*

The stimuli were constructed in a fully factorial design. The language dimension had two levels: either a stimulus sentence included a subject-extracted relative clause (SR) or an object-extracted relative clause (OR), as shown in (3|1). The music dimension had three levels: a melody included either only in-key tones (in-key), or only in-key tones except for one tone which was out-of-key (out-of-key), or only in-key tones with one tone being sung unusually loudly (auditory anomaly). This resulted in 120 stimulus sextuplets: 120 sentences in two linguistic versions and three musical versions, totaling 720 stimuli (120 × 2 × 3). Example stimuli can be accessed online: https://sites.google.com/site/rikunert/CV/example_stimuli_kunert_willems_casasanto_patel_hagoort .

(3|1)

a) Subject-extracted (SR)

De atleet die de minnaressen opmerkte keek uit het raam.

Literal: The athlete$_{singular}$ *that the mistresses*$_{plural}$ *noticed*$_{singular}$ *looked*$_{singular}$ *out of the window.*

English translation: The athlete that noticed the mistresses looked out of the window.

b) Object-extracted (OR)

De atleten die de minnares opmerkte keken uit het raam.

Literal: The athletes$_{plural}$ *that the mistress*$_{singular}$ *noticed*$_{singular}$ *looked*$_{plural}$ *out of the window.*

English translation: The athletes that the mistress noticed looked out of the window.

The language materials consisted of 120 Dutch sentences each in two versions, as can be seen in (3|1): the critical relative clause verb ('opmerkte') agreed in number either with the matrix clause noun phrase ('De atleet') in the subject-extracted version or with the relative clause noun phrase ('de minnares') in the object-extracted version. By ensuring that these two noun phrases differed in grammatical number we forced the listener to disambiguate the sentence and interpret it as one of the two syntactic versions. Disambiguation was only possible at the moment of listening to the relative clause verb.

Sentences were on average 10 (standard deviation = 1.3) words long with the disambiguating relative clause verb always being the sixth word. The final syllable of the relative clause, which distinguishes between the SR and OR versions, was sung on any beat within a 4/4 bar (11.6% on the first beat, 31.7% on the second, 24.2% on the third, 6.7% on the final beat, the remainder on off-beat notes). The matrix subject was plural in half of the SR sentences, i.e. the grammatical number of the first noun phrase was not indicative of the linguistic condition.

In order to ensure that participants would process the full sentences, a linguistic task checked language comprehension by use of prompts relating to some part of the stimulus sentence (e.g., ‘Iemand merkte de atleet op.’ *Somebody noticed the athlete*.). Prompts required a true/false response. Half the comprehension prompts checked for matrix clause understanding. The other half focused on the relative clause (as in the aforementioned example). In order to avoid task-specific strategies we also created (1) more challenging passive voice prompts and (2) prompts with ‘someone’ (‘iemand’) as a singular subject possibly representing either a plural or a singular noun phrase in the song (see example prompt). Within each comprehension prompt version half the prompts matched the content of the songs.

Each of the two sentence stimulus versions was combined with three versions of a melody (in-key, out-of-key, auditory anomaly). All melodies were composed specifically for this study by a professional composer (Jason Rosenberg, www.jasonrosenberg.org). The three music versions of each of the 120 melodies differed only in terms of the tone sung on the stressed syllable of the disambiguating relative clause verb in terms of pitch (in-key versus out-of-key) or loudness (in-key normal volume versus in-key auditory anomaly [loud volume]), see Figure 3|1. The in-key and auditory anomaly conditions did not differ in pitch. Melodies were rhythmically diverse and on average 10.2 seconds long (standard deviation = 1.3) at a tempo of 70 beats per minute, i.e. a quarter note corresponded to a nominal duration of 857 ms. The beginning of each melody established a strong sense of key. The three music conditions were in the same key and differed only by one note. This critical tone coincided with the stressed syllable of the relative clause verb, and was either part of the established key (in-key normal volume, auditory anomaly [in-key but loud volume]) or not (out-of-key normal volume). The melodies were composed in such a way that the location of the out-of-key note was musically plausible from the standpoint of harmonic tension-resolution patterns (Lerdahl, 2013). Rhythmically, the critical note was always a quarter note in length, and occurred on various beats (44.2% on the first beat, 36.7% on the second, 6.7% on the third, 12.5% on the fourth beat). Each of the twelve major keys was used 10 times (10 × 12 = 120 sets). Melodies were in the baritone range.

After stimulus design, the 120 (sets) × 2 (SR and OR versions) × 2 (in-key and out-of-key versions) sung sentences were recorded in a soundproof room at the Max Planck Institute for

Psycholinguistics in Nijmegen by a 34 year old male Dutch baritone. The singer was an amateur (Jan-Mathijs Schoffelen) who had been trained for 16 years in total (piano and voice). First, each of the 480 songs (four per set) was recorded separately in each of the linguistic and harmonic conditions. Afterwards, all recordings were normalized for loudness level. Next, steps were taken to control for acoustic cues prior to the critical verb. Specifically, we cut out the verb recording of one harmonic version and pasted it into the audio stream of the other. This created two harmonic versions of each sentence with identical recordings except for the critical verb. After this splicing step the new song signal was adjusted in order to avoid the audibility of the verb recording exchange. To exclude any possible systematic influence of this processing step it was ensured that an equal number (exactly half) of in-key and out-of-key recordings were left unchanged. Next, the auditory anomaly condition of each sentence was created. Of the resulting four files the in-key versions were chosen and the critical tone's loudness was increased by 10 dB following Fedorenko et al. (2009). All audio manipulations were done with the program Audacity version 1.3 (audacity.sourceforge.net).

3|2.4 Procedure

Of each of the 120 stimulus sextuplets, each participant heard both linguistic versions, i.e. a total of 240 trials (120 × 2). However, each linguistic version of a stimulus sextuplet was only presented in one music condition. Still, overall, each participant heard an equal number of trials in each music condition. Following an event-related design, the stimuli were ordered pseudo-randomly with the following constraints: (1) no more than three consecutive trials with the same prompt condition (the prompt matches the sentence or not), (2) no more than three consecutive trials of the same music-language condition, and (3) at least ten trials between any stimulus set's SR and OR versions. For every three participants a new pseudo-randomized stimulus order was used. Within each such participant-triplet, for each trial the musical condition was counterbalanced. The stimuli were presented to the participants using MR-compatible non-magnetic earphones (Sensimetrics, model S14) which also dampened scanner noise. Volume was set at a subject-specific, comfortable level before the start of the experiment.

Participants were asked to concentrate on the linguistic dimension of the sung sentences. As in most previous studies examining interactions between linguistic and musical syntactic processing (e.g., Koelsch, Gunter, et al., 2005; Slevc et al., 2009), there was no musical task. That is, we relied on the musical structure being processed implicitly. The experiment was organized as follows. Four example trials preceded the experimental session. Experimental trials were divided into eight blocks of 30 sung sentences. After four blocks participants could rest for approximately ten minutes while an anatomical MRI scan was acquired.

Each trial was organized as follows. After a stimulus was played a comprehension prompt was displayed visually through a projector from outside the scanner room. Subjects saw it through a nonmagnetic mirror attached to the head-coil. Within 10 seconds they had to press a button to indicate whether the prompt was true according to the preceding sung sentence or not. Except for the example trials, no feedback was given. Stimulus onset was jittered with respect to volume acquisition by randomly varying the intertrial interval (time between response to the previous trial's prompt and the song-onset of the next trial) between 3.5 and 6 seconds. During the intertrial interval as well as during the song presentation a fixation cross was displayed centrally. An experimental session lasted approximately 100 minutes.

3|2.5 fMRI Data Acquisition

The experiment was carried out in a 1.5 Tesla MRI scanner (Siemens Avanto, Siemens Medical Systems, Erlangen, Germany). Thirty-three axial slices were acquired (3.5 mm × 3.5 mm in-plane resolution, 3 mm slice thickness, 0.51 mm slice spacing, field of view [FOV] = 224 mm) covering the whole brain. We used a single-shot echo-planar imaging (EPI) sequence (repetition time [TR] = 2140 ms, echo time [TE] = 40 ms, 90° flip-angle [FA]). In the middle of the scanning session a 3-D T1 scan was acquired (176 slices, voxel size = 1 mm × 1 mm × 1 mm, TR = 2250 ms, TE = 2.95 ms, FA = 15°, sagittal orientation).

3|2.6 fMRI Data Analysis

Analysis was carried out using SPM8 (www.fil.ion.ucl.ac.uk/spm). The first five volumes of each functional run were discarded. In order to compensate for small head movements, images were realigned to the first image by means of rigid body registration. Slice timing correction was applied by means of linear interpolation to the onset of the first slice. All functional datasets were individually co-registered using the participants' individual high-resolution anatomical images. Afterwards, this co-registered EPI dataset was normalized to Montreal Neurological Institute (MNI) space. The time series were high pass filtered with a cut-off frequency of 128 seconds and images were spatially smoothed using an 8mm FWHM Gaussian kernel.

The statistical evaluation was performed using the general linear model. The model was generated with a synthetic hemodynamic response function modeled on the manipulated song region, i.e. the start of the critical verb until the end of the song. We separately modeled the six conditions of interest and included two nuisance regressors (dummy variables for run1 and run2) to capture the effect of functional scanning run as well as 18 nuisance regressors derived from the motion correction algorithm. These modeled variability in all three rotations and all three translations due to linear motion, quadratic motion and the first derivative of linear motion (6 motion types × 3 quantifications = 18 regressors; see Lund, Nørgaard, Rostrup, Rowe, & Paulson, 2005). Statistical analysis was performed by computing contrast maps for each condition for each participant separately including all of his or her trials (independent of behavioral performance), and the subsequent group analysis involved calculating interaction and main effects in a full factorial ANOVA with factors language (SR, OR) and music (in-key, out-of-key, auditory anomaly). In this way participant was treated as a random factor ('random effect analysis'). The multiple comparisons problem ensuing from this massive univariate approach was dealt with by applying a topological feature based false discovery rate correction at the .05 level (peak-based FDR; Chumbley & Friston, 2009; Chumbley, Worsley, Flandin, & Friston, 2010).

The region definitions used in the structural region of interest (ROI) analysis we derived from the Automated Anatomical Labeling library (Tzourio-Mazoyer et al., 2002). The chosen ROIs were those where overlapping activation sites between music harmony and language syntax had been reported (see Introduction): bilateral inferior frontal gyrus (IFG) pars opercularis and pars triangularis, i.e. Broca's area and its right hemisphere homologue. The Marsbar ROI toolbox version 0.42 (Brett, Anton, Valabregue, & Poline, 2002) was used to derive average contrast values across the 3567 and 3550 voxels of size 2 × 2 × 2 mm^3, in the left and right structural ROIs respectively, based on data generated during the first level analysis with SPM8. Please note that we are aware of the literature describing structural and functional differences between different parts of Broca's area (e.g., Bookheimer, 2002). Nonetheless, we only defined a single Broca's area ROI for three reasons: 1) Patel's SSIRH does not specify which part of Broca's area should show the predicted interaction between music and language, 2) previous studies which investigated music and language separately found syntax-processing related activations in both pars opercularis (music: Koelsch, Fritz, et al., 2005; language: Snijders et al., 2009) and pars triangularis (music: Koelsch, Fritz, et al., 2005; Maess et al., 2001; language: Segaert, Menenti, Weber, Petersson, & Hagoort, 2012), 3) we aimed to reduce the number of ROIs in order to have sufficient statistical power after controlling for the number of comparisons (Bonferroni method), i.e. the number of structural ROIs.

The ROI data were not normally distributed. Using the SPSS implementation of the Kolmogorov-Smirnov test to check the distributions of OR-SR difference scores within each music condition and ROI revealed that two distributions were significantly different from normal [left and right hemisphere, out-of-key: $D\text{s}_{(19)} > .19$, $p\text{s} < .05$]. In order to maximize power, these *p*-values of the normality test are not corrected for multiple comparisons. In order to account for the non-normal data distribution, inferential analyses of the ROI data were carried out using random permutation based tests which require no parametric assumptions. In terms of the dependent *t*-tests this amounts to creating a null hypothesis *t*-distribution by randomly applying condition labels to data points within each participant 50,000 times and testing the effect of interest on the randomized data each time. The proportion of randomly obtained *t*-values equal or greater than the true *t*-value represents the likelihood of obtaining the *t*-statistic under the null hypothesis, i.e. the *p*-value. Similarly, the random permutation based ANOVA randomized labels within each participant but otherwise in an unrestricted way across experimental factors (Manly, 2006). ANOVA *p*-values were Bonferroni corrected for two ROIs and within each ROI *t*-test *p*-values were corrected for three comparisons. Only the corrected *p*-values are reported.

It has recently been argued that doing region of interest analysis with the same ROI across the whole group of participants is a statistically insensitive procedure (Nieto-Castañón & Fedorenko, 2012). Therefore we complemented our previous ROI analysis with a functional ROI (fROI) analysis using the spm_ss toolbox (Nieto-Castañón & Fedorenko, 2012). For each subject separately, we extracted the top 10% of voxels (357 voxels) in the left IFG (pars opercularis and pars triangularis, taken from the AAL template) which exhibited the highest *t*-values in the OR > SR contrast (averaged across music conditions). Strictly speaking the voxels did not need to be adjacent, but in practice they mostly are. In order to ensure the independence of data for fROI identification and activity estimation, we used either the first or the second scanning run for fROI building and the left-out run for estimation of activation during the conditions of interest (see Kriegeskorte, Simmons, Bellgowan, & Baker, 2009). Responses were averaged across the two partitions (' 2-fold cross-validation procedure', see Fedorenko, McDermott, Norman-Haignere, & Kanwisher (2012) for a similar approach). Thus, each subject had a different fROI in Broca's area. Data from the fROI were used to derive average contrast values across voxels. Inferential analyses were again carried out using random permutation based tests and *t*-test *p*-values were corrected for three comparisons (Bonferroni method). Only the corrected *p*-values are reported.

3|3 Results

3|3.1 Behavioral Results

Participants answered one comprehension prompt after each trial. The accuracy rates revealed that no participants scored at or below chance level, i.e. not with an accuracy below 56% (binomial distribution, $p < .05$). Scores ranged between 66% and 90% ($M = 78\%$). A 2 (prompt type: matrix or relative clause) × 2 (language: SR or OR) × 3 (music: in-key, out-of-key, or auditory anomaly) dependent ANOVA revealed three effects. First, there was a main effect of prompt type [$F_{(1,18)} = 143.56$, $p < .001$, ${}_{p}\eta^2 = .889$, ${}_{p}\omega^2 = .882$], such that prompts targeting main clause understanding were easier to answer (88%) than prompts targeting relative clause understanding (68%). Furthermore, a main effect of linguistic condition was found [$F_{(1,18)} = 43.90$, $p < .001$, ${}_{p}\eta^2 = .709$, ${}_{p}\omega^2 = .693$] indicating that prompts after SR sentences were answered more accurately (86%) than those after OR sentences (70%). Furthermore, these two main effects interacted [$F_{(1,18)} = 51.18$, $p < .001$, ${}_{p}\eta^2 = .740$, ${}_{p}\omega^2 = .725$]. Follow-up *t*-tests revealed that the difference between SR and OR sentences is significant for both kinds of prompts albeit larger for those targeting relative clause comprehension [$t_{(18)} = 7.05$, $p < .001$] than those targeting main clause comprehension [$t_{(18)} = 3.85$, $p < .01$]. This supports the idea that OR sentences were indeed more challenging than SR sentences. However, this difficulty did not interact with the music factor [$p > .3$]. The three-way interaction was not significant. It should be borne in mind that the behavioral measure was designed to ensure adequate neural processing instead of showing the previously reported behavioral interaction effect (Fedorenko et al., 2009).

Therefore, the crucial test of our hypothesis lies in the neural data analysis. We will return to this point in the discussion section.

3|3.2 fMRI Results

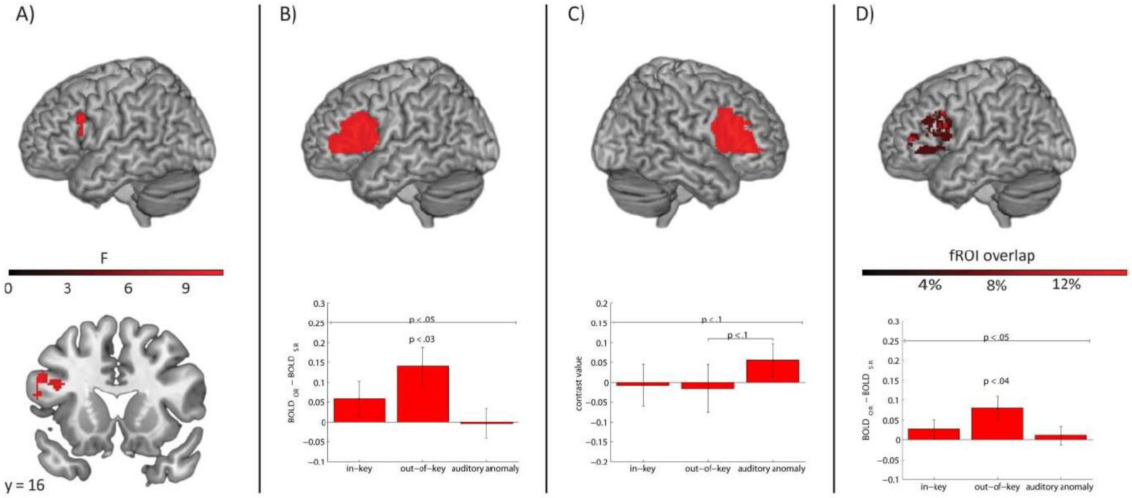


Figure 3|2. fMRI results. A) The language main effect (OR > SR) found in the whole-brain analysis (p < .005 uncorrected, cluster size = 87 voxels). **B)** Left hemisphere structural ROI. The BOLD effect of the linguistic manipulation is shown (OR - SR) with the associated *p*-value of a paired *t*-test above the bar. The significance level of the interaction effect is denoted above the line. Bars represent the activity difference (OR-SR) to sequences in which the stressed syllable of the critical word was sung in-key, out-of-key or unusually loudly (auditory anomaly). **C)** Right hemisphere structural ROI. The BOLD effect (compared to implicit baseline) is shown for each music condition. The *p*-value of a dependent *t*-test comparing two music conditions can be seen above the respective bars. The significance level of the music main effect is denoted above the line. **D)** Left hemisphere functional ROI. fROIs were individually defined in the left structural ROI. The inter-subject overlap in fROI locations is shown in the top panel. See methods for details. The BOLD effect is shown for the three different music conditions separately. Error = SEM. All *p*-values in structural ROI analyses are Bonferroni adjusted.

3|3.2.1 Whole-brain Analysis

For the whole brain analysis, no cluster emerged for any of the main effects or their interaction with a probability of $p < .05$ (FDR corrected). In order to see whether our data set replicates previous findings of language syntax-related effects in left prefrontal areas, we lowered the statistical threshold ($p < .005$ uncorrected) and identified the biggest cluster (87 voxels) with a peak at [-54; 18; 28], see Figure 3|2A. The cluster represents increased activity to OR sentences compared to SR sentences and it covers parts of the IFG pars opercularis and pars triangularis, showing that our data set can replicate previous findings albeit only at a reduced statistical threshold.

3|3.2.2 Structural ROI analysis

The predicted interaction between the language and music factors was found in left IFG [$F = 4.14$, $p < .05$] but not right IFG [$F = 1.68$, $p > .4$]; see Figure 3|2B. Follow-up *t*-tests showed that the significant interaction in left Broca's area emerged because the OR > SR contrast was only significant in the out-of-key condition [$t = 2.93$, $p < .03$] but not in the in-key condition [$t = 1.30$, $p > .5$] or the auditory anomaly condition [$t < 1$]. Similar analyses in the right ROI revealed no significant OR > SR effect in any of the music conditions [all p values > .2]. The language main effect was not significant in either region of interest [left: $F = 4.00$, $p > .1$; right: $F = 1.76$, $p > .4$]. However, the music main effect was marginally significant in the right hemisphere region of interest [$F = 3.26$, $p < .1$] but not in the left one [$F = 2.63$, $p > .1$]; see Figure 3|2C. The former was due to a marginally greater activation in the auditory anomaly condition compared to the out-of-key condition [$t = 2.28$; $p < .1$]. The contrast with the in-key condition did not approach significance [$t = 2.00$; $p > .1$], nor did the in-key vs. out-of-key contrast [$t < 1$].

3|3.2.3 functional ROI analysis

The 10% of voxels in the left IFG which exhibited the strongest language effect were used to construct an fROI for each subject separately. Figure 3|2D (top panel) shows that the overlap of included voxels was small across participants, reflecting known individual differences in the location of language-related activity peaks in Broca's area (Xiong et al., 2000). Similar to the structural ROI, this language syntax-related fROI exhibited the predicted interaction between the language and the music factors [$F = 3.27$, $p < .05$], see Figure 3|2D (bottom panel). Follow-up t-tests again showed that the significant interaction in the fROI emerged because the OR > SR contrast was only significant in the out-of-key condition [$t = 2.78$, $p < .04$] but not in the in-key condition [$t = 1.09$, $p > .5$] or the auditory anomaly condition [$t < 1$]. The language main effect - indicative of an OR > SR pattern [$F = 3.89$, $p < .1$] - as well as the music main effect [$F = 2.65$, $p < .1$] were only marginally significant. The latter was reflecting a pattern previously seen in the right hemispheric structural ROI: greater activity to the auditory anomaly condition compared to the other two music conditions.

3|4 Discussion

The present study aimed to provide brain-imaging support for the proposal that syntax processing in music and language interact in the human brain. To this end we adopted an interference paradigm. We found a statistical interaction between music and language processing in Broca's area, corresponding to BA44 and BA45 in the left inferior frontal gyrus. This music-language interaction even emerged when restricting the analysis to voxels within Broca's area which are involved in language syntax processing. This suggests that at least some of the neural resources in Broca's area that process syntactic relations between words in language are also sensitive to syntactic relations between tones in music, and that syntactic integration in language is not wholly independent of syntactic integration in music. Note that this non-independence is not due to shared general attention resources as an auditory anomaly led to a different activation pattern.

Specifically, the interaction between music and language emerged when participants heard a stimulus containing a syntactically challenging sentence (object-extracted relative clause instead of subject-extracted relative clause) sung on a melody containing a syntactically challenging tone (out-of-key instead of in-key), with the tone located at the precise point in the melody where the linguistic syntactic integration difficulty occurred. In this case an interaction pattern emerged (see Figures 3|2B and 3|2D). This is indicative of an even greater integration difficulty in this condition compared to what would be expected from integrating challenging words and tones entirely independently of each other. In order to check whether non-syntactic auditory anomalies would also show such a pattern we included a control condition in which the critical tone was sung in-key but unusually loudly. In Broca's area this control condition did not lead to activity patterns similar to the out-of-key condition, even in numerical terms. Instead of affecting the left hemisphere Broca's area, the control condition seemed to activate the right hemisphere homologue of Broca's area. The implications of these findings for our understanding of music and language are discussed below.

3|4.1 A Common Role for Broca's Area in the Music and Language Networks

The current study has found some support for a common syntactic processing role of Broca's area in music and language. This fits with results showing that musical training is associated with structural changes in this area (Bermudez, Lerch, Evans, & Zatorre, 2009; Gaser & Schlaug, 2003; James et al., 2014; Sluming et al., 2002) and altered language syntax processing (Fitzroy & Sanders, 2013; Jentschke & Koelsch, 2009). Damage to this brain area is also known to lead to processing deficits in both language and music in non-musicians (Sammler et al., 2011).

However, the results of four recent fMRI studies might appear to contradict a common role for Broca's area in the music and language networks. Two found common brain areas but differing music and language activation patterns in them using multi-voxel pattern analysis (MVPA) (Abrams et al., 2011; Rogalsky et al., 2011). In two other studies Fedorenko et al. (Fedorenko, Behr, & Kanwisher, 2011; Fedorenko et al., 2012) found different activated brain regions when comparing a music-localizer based on a scrambling manipulation to a language-localizer based on the reading of sentences versus lists of non-words. However, none of these studies specifically manipulated syntactic structure in language and tonal/harmonic structure in music, while leaving other aspects of sequence structure intact (cf. Peretz, Vuvan, Lagrois, & Armony, 2015b; Slevc & Okada, 2014). It is also worth keeping in mind that the SSIRH actually predicts the overlap between music and language to be partial, not complete. The question which we attempted to answer in this study was whether music and language share *any* circuitry at the level of syntactic processing, as suggested by the music-language interaction in Broca's area.

3|4.2 Non-syntactic Overlap between Music and Language

Despite the evidence for a shared syntactic processing mechanism in the left inferior frontal gyrus, alternative explanations for our results could be proposed. First of all, any auditory anomaly might draw attention away from language and thus interact with linguistic processing. We reject this explanation because a control condition consisting of a sudden 10dB loudness increase did not lead to a similar pattern of results compared to the harmonic violation. This is striking since, as opposed to the subtle harmonic violation which interacted with language processing, the loudness increase evoked a marginally significant brain correlate in the right hemisphere's inferior frontal gyrus. This more salient non-syntactic manipulation, however, did not interact with language processing. This supports a shared syntactic neural architecture between music and language. Furthermore, the finding is in line with previous behavioral and ERP studies which found that neither a loudness anomaly nor a timbral anomaly leads to the same music-language interactions as seen with harmonic manipulations (Fedorenko et al., 2009; Koelsch, Gunter, et al., 2005; Slevc et al., 2009).

Another recent alternative explanation has been Rogalsky et al.'s (2011) proposal that music and language processing exhibit a link only in tasks which involve the processing of violations. However, the current study elicited a music-language interaction by using relatively easier or more difficult linguistic constructions which were without any errors, as well as a musically plausible in-key/out-of-key tone manipulation (Lerdahl, 2013). Moreover, the brain activation response we find is not indicative of linguistic error processing which is associated with relatively more right-lateralized prefrontal activation sites (Indefrey, Hagoort, Herzog, Seitz, & Brown, 2001), as opposed to the relatively left-lateralized effect here. Thus, the overlap we found does not appear to be elicited only under the exceptional circumstances of processing violations (see also behavioural studies without error manipulations: Fedorenko et al., 2009; Slevc et al., 2009).

Still, some studies have reported interactions between music and semantic language manipulations (Perruchet & Poulin-Charronnat, 2013; Steinbeis & Koelsch, 2008). The present study does not directly address semantic language processing, but its design could be extended to investigate the neural differences between semantic-harmonic and syntactic-harmonic interactions. Thus, more research is required in order to address the syntax-specificity of the interaction we found in the present study.

Besides attention and violation processing, it could also be suggested that the observed activation differences reflect decision-making related processes. Binder, Liebenthal, Possing, Medler, & Ward (2004) have shown that a cluster in the lateral part of the left IFG is associated with decision making performance in a syllable differentiation task. However, such an explanation is unlikely to reflect the pattern seen here because a decision was only required *after* a song was heard, upon seeing a comprehension prompt. Furthermore, the kind of prompt was variable and unpredictable. For example, half the comprehension prompts did not focus on the relative clause manipulation at all. Thus, activation differences due to a decision process are unlikely as decision making started after the song, i.e. after the time interval which the current fMRI analysis investigated.

3|4.3 The Role of the Right Inferior Frontal Gyrus

The right hemisphere homologue of Broca's area did not show activity related to linguistic or harmonic processing, which was partly surprising given that previous brain imaging studies reported an involvement in both cases (e.g., Embick et al., 2000; Tillmann, Koelsch, et al., 2006). In contrast to these studies, we employed a task-unrelated, subtle harmonic manipulation which was based on a single tone (the smallest possible alteration of melody). This manipulation might not have been strong enough to reliably activate right hemisphere areas involved in musical harmonic processing (e.g., Tillmann, Koelsch, et al., 2006). In future work, one could increase the salience of the tonal/harmonic manipulation, e.g., by using a melody sung over instrumental musical chords, or over an instrumental melody with several notes per sung word, so that the critical word was accompanied by several out-of-key notes. The subtle effect we find could be taken to suggest that the tone manipulation is only able to modulate linguistic processing already triggered by the language task. Instead of syntactic or harmonic processing we found a marginal attention-related effect in the right inferior frontal gyrus. Our control condition, a salient loudness increase, seemed to activate this region, likely due to its involvement in the bottom-up attention network (Corbetta & Shulman, 2002; Fox, Corbetta, Snyder, Vincent, & Raichle, 2006).

3|4.4 Limitations

The absent behavioural effect of music on language might appear surprising. However, it does not necessarily contradict the proposal for shared syntactic processing resources in the left inferior frontal gyrus. In comparison to a previous behavioral study (Fedorenko et al., 2009) which did find a behavioral effect with a similar paradigm, several aspects of our study might have lowered the sensitivity of our behavioral measure. First, we used a diverse set of comprehension prompts which were partly challenging in themselves (e.g., passive voice prompts) and which often did not focus on the relative clause manipulation. This was necessary in order to discourage unnatural task strategies, a problem Fedorenko et al. (2009) were not faced with due to linguistic differences such as a word order manipulation in English vs. a number agreement manipulation in Dutch. Second, in order to reduce the duration of scanning sessions, our stimulus list did not include fillers. Third, our stimuli were rhythmically and linguistically more diverse, possibly increasing ecological validity at the expense of reducing the effect size. In sum, by focusing the present study on exploring a neuronal effect we did not optimize the design for finding a behavioral effect.

The neural effects we find could appear weak. Concerning the language main effect, we only find a marginally significant language syntax effect in Broca's area, and that result only emerges in the functional region of interest analysis. This weak effect might be a consequence of the syntactic manipulation we used. Dutch participants could have misheard the number of the relative clause verb in the more difficult object-extracted relative clauses, and therefore 'default' to the more common subject-extracted relative clause version. Such a process is considerably less likely in an object- versus subject-relative clause manipulation based on word order, such as used in (Fedorenko et al., 2009). Thus, future work might employ a word-order based syntactic contrast. Similarly, the influence of music on the language effect was relatively weak, see Figure 2|2B and Figure 2|2D. This might simply mirror the rather subtle music manipulation in combination with our particular choice of syntactic constructions. Moreover, pilot work reported by Fedorenko et al. (2009) suggests that music-language interaction effects might be enhanced by an increased rate of presentation (the average rate in (Fedorenko et al., 2009) was 1.78 words/sec, versus 0.98 in the current study). Future work is needed to test whether the effects found here generalize to other music and language manipulations.

3|5 Conclusion

The present study aimed to test the hypothesis that music and language share neural resources for syntactic processing in Broca's area. The predicted interactive pattern between music and language demands was indeed found in this part of the brain. This is the first direct evidence which suggests that music and language syntactic processing interact in Broca's area.

3|6 Acknowledgements

We would like to thank the following people for help in stimulus creation: Jason Rosenberg, Dan Acheson, and Jan-Mathijs Schoffelen. The sentence material was largely based on a stimulus set developed by Dan Acheson. We are grateful for Paul Gaalman's help during data acquisition, to Sean J. Fallon for discussing earlier versions of this draft, and to Ev Fedorenko for suggesting fROI analyses.

3|7 Author Contributions

Conceived and designed the experiment: RK RMW ADP DC PH. Performed the experiment: RK. Analyzed the data: RK RMW. Wrote the paper: RK RMW ADP.

Chapter 4

When do music harmony and language syntax processing interact? An MEG study

Music and language are two communication systems relying on the structured organization of elements such as words or tones into higher order structures like sentences or melodies. Previous studies have suggested that comprehending the structured organization of music in terms of its harmonic properties partly relies on neural resources which are also involved in the processing of the syntactic structure of sentences. In the present magneto-encephalography (MEG) study we employed an interference design based on sung sentences in order to elucidate the temporal dynamics of shared music-language resources involved in structural processing. Unexpectedly, we found no evidence for shared music-language syntactic processing resources, neither in our behavioural nor in our neural data. A marginally significant posterior event-related field effect was confounded by attentional capture processes, suggesting that previous studies which failed to control for attentional capture might similarly be confounded. Overall, the present results do not support the notion of shared syntactic processing between music and language. We include a list of suggestions for why we, in contrast to many previous studies, failed to reveal shared music-language resources. The discussion of our design might help in optimizing future investigations into the temporal dynamics of music and language processing.

Kunert, R., Jan-Mathijs Schoffelen, Roel M. Willems, Jason C. Rosenberg, Aniruddh D. Patel, Peter Hagoort (unpublished). When do Music Harmony and Language Syntax Processing Interact? An MEG Study.

4|1 Introduction

Music and language are two communication systems which depend on the precise timing of sounds. Through implicitly learned rules (Tillmann et al., 2000), listeners understand that an underlying structure connects the sounds in meaningful ways. As a result, basic elements like words in language and tones/chords in music can be integrated into overarching structures like sentences in language and harmonic sequences/melodies in music (Patel, 2008). These parallels between music and language motivated the proposal for shared music-language structural integration resources in the brain (Patel, 2003; Patel et al., 1998). The present paper investigates the temporal dynamics of these shared resources using magneto-encephalography (MEG), following an approach previously used in a functional magnetic resonance imaging (fMRI) study (Kunert et al., 2015).

In the present paper musical structure refers to the implied harmony of tonal melodies. In the Western tradition, pitches are variously grouped. First of all, all pitches are manifestations of one of only 12 octave-equivalent pitch classes (e.g., all C tones on a piano). Moreover, a subset of 7 pitch classes forms a scale (e.g., C-major scale). When playing in a given harmonic key (e.g., C-major) these seven pitch classes (in-key tones, e.g., C, D, E, F, G, A, B) are emphasized. Once a listener has derived a sense of key, out-of-key tones (e.g., C#) are less expected than in-key tones (Krumhansl, 1979; Krumhansl & Kessler, 1982). As a result, out-of key tones are thought to be more challenging to harmonically integrate than in-key tones (Lerdahl, 2001).

Does the brain use the same neural resources for structural processing in music harmony and in language syntax? Brain damaged patients can experience language problems without music problems and vice versa (Peretz & Coltheart, 2003; but see Patel et al., 2008), suggesting that music and language engage brain circuitry which is sufficiently distinct to be selectively disrupted. However, brain responses to structurally challenging stimuli are very similar between music and language (Patel et al., 1998), suggesting that there are commonalities in processing across cognitive domains.

In order to resolve this seemingly paradoxical set of findings for both independent processing of music and language (lesion studies) and shared processing resources, Patel (2003,

2008) proposed the shared syntactic integration resource hypothesis (SSIRH). The SSIRH hypothesizes that there are two distinct components to structural processing in music and language. On the one hand, representations which music and language do not share need to be stored in long term memory (hypothesized in the temporal lobe). On the other hand, these domain-specific representations need to be integrated in sentences/melodies using shared music-language structural integration resources (hypothesized in prefrontal brain areas). When dependencies are costly, e.g. because a tone in the melody is out-of-key (Lerdahl, 2001) or because the word requires the re-activation of a previous word held in working memory in order to provide a successful parse (Gibson, 1998), resource networks are recruited.

One prediction derived from this account is that the limited capacity of resource networks should lead to suboptimal processing when harmonically costly tones coincide with syntactically costly words. Such interference effects have indeed been shown in many behavioural (Fedorenko et al., 2009; Hoch et al., 2011; Kunert & Slevc, 2015; Kunert, Willems, & Hagoort, 2016; Slevc et al., 2009), electro-physiological (Carrus et al., 2011, 2013; Koelsch, Gunter, et al., 2005; Steinbeis & Koelsch, 2008), and hemodynamic brain activity (Kunert et al., 2015) studies. The present study adopts a similar interference design, one previously used in a behavioural and an fMRI study (Fedorenko et al., 2009; Kunert et al., 2015). In this paradigm sentences include an embedded relative clause (in italics) which is either challenging (**a**) or not (**b**):

a) The boy *that the girl helped* got an "A" on the test.

b) The boy *that helped the girl* got an "A" on the test.

The sentences are sung to melodies containing an out-of-key tone on the last word of the relative clause which requires a challenging long-distance integration in **a)** but not in **b)**, according to dependency locality theory (Gibson, 1998, 2000). A control condition includes a non-harmonically deviant tone which is thought to grab the listener's attention. The answers to comprehension prompts which followed each sentence revealed impaired comprehension when difficult sentences were sung with an out-of-key note compared to only in-key notes (Fedorenko et al., 2009; but see Kunert et al., 2015). Moreover, a pre-frontal brain area (Broca's area) thought to be involved in structural integration in language and music (Hagoort, 2005, 2013; Patel, 2008) evidenced a similar interaction.

However, the temporal dynamics of the shared music-language resources are very much unclear, with effects related to shared music-language structural integration resources claimed in nearly all time windows commonly associated with cognitive processing (roughly 150 - 800 ms post stimulus onset). Patel et al. (1998), using an event-related potential (ERP) design with either music or language stimuli, identified a positive component from 450 to 750 ms (P600) which was indistinguishable between a language syntax contrast and a music harmony contrast. Other researchers found an earlier language syntax evoked ERP component, the left anterior negativity (LAN; 300 - 400 or 450 ms), which evidences a lower amplitude when a syntactic error coincides with a harmonically challenging chord or unexpected tone (Carrus et al., 2013; Koelsch, Gunter, et al., 2005; Steinbeis & Koelsch, 2008). Both the P600 and the LAN time windows have also been associated with structural music-language interactions in terms of the oscillatory dynamics of electrical brain responses (Carrus et al., 2011). Finally, an even earlier ERP component associated with harmonic processing, the early right anterior negativity (ERAN; 160 - 220 or 260 ms) has also been shown to be modulated by language syntax errors (Maidhof & Koelsch, 2011; Steinbeis & Koelsch, 2008). In summary, there is next to no time-window after 160 ms in which structural music-language interactions have *not* been found.

A possible interpretation for this lack of temporal precision of music-language interactions might lie in the fact that some of the aforementioned EEG results are confounded by attentional and/or error related processing. In terms of the error processing confound, language syntax processing is operationalised as (morpho-)syntactic error processing in all of the aforementioned electrophysiological investigations (Rogalsky et al., 2011). In many studies, there is no non-syntactic error contrast to control for a possible error processing confound (Koelsch, Gunter, et al., 2005; Maidhof & Koelsch, 2011; Steinbeis & Koelsch, 2008). Moreover, none of the aforementioned EEG studies control for the effect of a harmonically deviant tone/chord on attention. Presumably, such unexpected tones/chords grab attention in similar ways as acoustic stimuli without structural properties.

Therefore, the current study is based on a paradigm which uses a syntactic contrast which does not involve syntactic errors. Moreover, we included an acoustic deviant control condition (10dB loudness increase) in order to control for the effect of a harmonically unexpected tone on attention. Sung sentences were heard during the acquisition of the magnetic-encephalogram (MEG) providing excellent temporal and very good spatial resolution of brain activity. In an event-related field (ERF) analysis, we expected the behavioural and fMRI interaction between music and language (Fedorenko et al., 2009; Kunert et al., 2015) to translate into an interference effect in one of the aforementioned ERP time-windows. In a subsequent time-frequency analysis we focused on changes in ongoing oscillatory brain activity in response to the stimuli. While previous music-language investigations identified the delta and theta bands as frequency bands where structural processing of music and language interact (Carrus et al., 2011), language-specific studies have proposed the beta band (Bastiaansen & Hagoort, 2006; Bastiaansen, van Berkum, & Hagoort, 2002) to reflect syntactic integration processes which the SSIRH hypothesizes to be shared with music harmony processing (Patel, 2003, 2008).

4|2 Methods

4|2.1 Participants

A total of 48 participants finished the experiment of whom 18 were rejected because of below chance task performance on critical trials (< 56% correct, binomial distribution; *N* = 16) or due to low data quality (*N* = 2). The final sample (15 male, 15 female) consisted of 30 Dutch native speakers with little musical training (*M* = 2.4 years, *SD* = 2.8) who all reported being right handed. No subject had a known history of neurological, language related or hearing problems and all had normal or corrected-to-normal vision. Written, informed consent was provided before the experiment began.

4|2.2 Stimuli

The 120 critical items have been used before and are described in detail elsewhere (Kunert et al., 2015). Briefly, critical stimulus construction followed a fully factorial design crossing the factors Language (subject-extracted relative clause, SR; object-extracted relative clause, OR; see example (4|1)) and Music (melody with only *in-key* tones, melody with only in-key tones except for critical tone being *out-of-key*; melody with only in-key tones with critical tone being 10dB louder forming an *auditory anomaly*; see Figure 4|1). Melodies were composed by author Jason C. Rosenberg (a professional composer), sung (combining sentences and melodies) *a cappella* by an amateur baritone singer (16 years of music experience, author Jan-Mathijs Schoffelen) at 70 bpm, and edited by author Richard Kunert. Comprehension prompts targeted the SR versus OR ambiguity in simple active voice sentences.

Ninety-six filler items were constructed in two versions differing only by one word. Their melodies were different to critical melodies but equally included three music versions for each sentence (in-key, out-of-key, auditory anomaly). Filler item syntax was based on various constructions including passive voice and simple complement phrases. Their length was 8 to 18 words (*M* = 9.6). In terms of singer, composer, pitch range and speed, the filler stimuli were very similar to critical stimuli.

(4|1)

a) Subject-extracted (SR)

De atleet die de minnaressen opmerkte keek uit het raam.

Literal: *The athlete*$_{singular}$ *that the mistresses*$_{plural}$ *noticed*$_{singular}$ *looked*$_{singular}$ *out of the window*.

English translation: *The athlete that noticed the mistresses looked out of the window.*

b) Object-extracted (OR)

De atleten die de minnares opmerkte keken uit het raam.

Literal: *The athletes*$_{plural}$ *that the mistress*$_{singular}$ *noticed*$_{singular}$ *looked*$_{plural}$ *out of the window.*

English translation: *The athletes that the mistress noticed looked out of the window.*

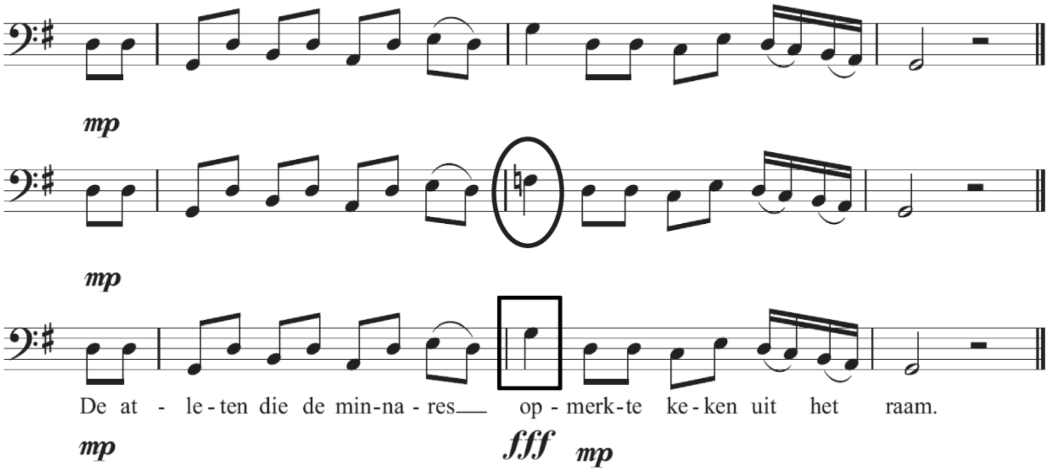


Figure 4|1. Example stimuli in the in-key condition (top), out-of-key condition (middle) and auditory anomaly condition (bottom). The critical tone (highlighted if deviating in implied harmony (circle) or volume (square)) coincides with the stressed syllable of the relative clause verb distinguishing the language conditions. The Dutch sentence in the figure means: *The athletes that the mistress noticed looked out of the window*. Figure taken with permission from Kunert et al. (2015).

4|2.3 MEG data acquisition

We used a 275 axial gradiometer system (CTF) situated at the Donders Centre for Cognitive Neuroimaging in Nijmegen, The Netherlands. The MEG signal was digitized at a sampling frequency of 1200 Hz. In order to continuously monitor the participant's head position relative to MEG sensors, three coils were attached to the participant's head (nasion, left and right ear canals; see Stolk, Todorovic, Schoffelen, & Oostenveld, 2013). Between blocks, participants were asked to reposition their head if necessary. Generally, deviations greater than 5 mm from the original position were rare. Three bipolar Ag/AgCl electrode pairs were used in order to measure the horizontal and vertical electro-oculogram, and the electro-cardiogram.

4|2.4 Procedure

Participants listened to a total of 432 trials divided over two sessions, each session consisting of six blocks. 240 trials were critical, consisting of each item (*N* = 120) in its SR and its OR versions. 192 trials were fillers, consisting of each item (*N* = 96) in two minimally different versions. Each linguistic version of a critical or filler item was presented in just one music condition. The stimuli were ordered in a pseudorandom way with the following constraints: (1) each block started with two filler trials, (2) at least 10 trials between the different language versions of an item, (3) no more than 5 consecutive same answer trials, (4) no more than 3 consecutive same language conditions, (5) no more than 3 consecutive same music conditions, and (6) no more than 4 consecutive fillers or critical trials. Trial sequences were organized in triplets. In each triplet the music conditions were counter-balanced. Every participant received a different trial sequence (10 randomization triplets in total).

Participants listened via custom-made non-metallic ear-phones. Volume was set at a subject-specific, comfortable level before the start of the experiment. Participants only had a linguistic task checked via comprehension prompts after each trial. There was no musical task (see Kunert et al., 2015). That is, we relied on the musical structure being processed implicitly.

Each trial started with a visual blinking signal presented for 1500 - 2000 ms, encouraging the participants to blink if necessary, followed by a fixation cross for 200 ms, followed by the auditory presentation of the stimulus during the continued visual display of the fixation cross. Blinking was discouraged at this point. Afterwards, a comprehension prompt was shown. Subjects had to respond within 10 s via a button press. At the end of each block, participants received task feedback. An experimental session lasted approximately 2.5 hours including set-up.

4|2.5 MEG analysis

4|2.5.1 Preprocessing

The MEG analyses employed FieldTrip, an open source toolbox programmed in MATLAB for analyzing M/EEG data (Oostenveld, Fries, Maris, & Schoffelen, 2011). We used a semiautomatic artifact identification procedure (www.fieldtriptoolbox.org/tutorial/automatic_artifact_rejection) in order to identify eye-movement and muscle contraction artifacts as well as jump artifacts of the SQUIDs (superconducting quantum interference devices). After additional visual inspection, all data segments containing artifacts were excluded from further analysis.

4|2.5.2 Event-related field analysis

The ERF analysis proceeded as follows. First, each sensor's average pre-trial baseline (200 ms) was computed for each session and participant separately. Second, the MEG signal surrounding the linguistic disambiguation point (t_0 = the earliest point in time at which the listener could distinguish the SR and OR conditions) was extracted (t_0 - 200 ms until t_0 + 800 ms), baseline-corrected, detrended, low-pass filtered at 40Hz, averaged according to conditions, and transformed from an axial to a planar gradiometer representation. Third, the signal was *z*-scored for each participant and channel separately. Finally, the contrast waves OR minus SR in each music condition (in-key, out-of-key, auditory anomaly) were computed based on the resulting *z*-scored MEG signal.

Statistical inference was done for two time-windows previously implicated in syntactic-harmonic music-language interactions: the early (right) anterior negativity at 160 - 260 ms and the left anterior negativity at 300 to 400 ms after stimulus onset (Steinbeis & Koelsch, 2008). The music-evoked early anterior negativity has been source-localized in the left and right inferior frontal gyrus (Garza Villarreal, Brattico, Leino, Østergaard, & Vuust, 2011; Maess et al., 2001). A language-evoked left anterior negativity has been found in the temporal lobe (Service, Helenius, Maury, & Salmelin, 2007). Given the plurality of potential sources (and resulting sensor-level signals), we decided to apply a non-parametric data-driven approach identifying spatial (sensor) clusters which pass a significance criterion ($p < .05$) for the average MEG signal in each time window (10,000 random permutations in order to derive a sampling distribution, Maris & Oostenveld, 2007).

4|2.5.3 Time-frequency analysis

The time-frequency analysis started by computing each participant's and sensor's average pre-trial baseline (200 ms long). Second, the signal was time-locked to the linguistic disambiguation point as described above for the ERF analysis (-1 s $< t_0 < 1$ s), baseline corrected for each channel and frequency separately, demeaned, and transformed from an axial gradiometer representation into a planar gradiometer representation. The power for every frequency between 2 Hz and 30 Hz (in steps of 2 Hz) was extracted through a sliding-window based multitaper time-frequency transformation based on the Fast Fourier Transform using Slepian sequences as tapers. We smoothed the frequency space by including the surrounding 4 frequency bins for each frequency's power estimate. The time-window for time-frequency transformations was 500 ms long (a moving window between t_0 - 1 s and t_0 + 1 s in steps of 50 ms). After baseline-line correction, the signal was averaged for each condition separately and difference scores for the language contrast were computed. The inferential statistical evaluation followed the approach taken in the ERF analysis by identifying spatial clusters passing a significance criterion.

4|3 Results

4|3.1 Behaviour

Behavioural results are based on participants' critical trial comprehension prompt answers. Accuracy scores ranged between 57% and 93% (*M* = 70%). A 2 (Language: SR or OR) × 3 (Music: in-key, out-of-key, or auditory anomaly) dependent ANOVA revealed only an effect of Language [$F_{(1,29)} = 48.28$, $p < .001$, ${}_{p}\eta^2 = .625$, $\omega^2 = .604$], indicating that prompts after SR sentences were answered more accurately (85%) than those after OR sentences (65%). The Music main effect was not significant [$F_{(2,58)} < 1$], nor was the predicted Language × Music interaction [$F_{(2,58)} < 1$].

4|3.2 Event-related fields

In a first step, the language main effect (OR > SR) was investigated but no spatial cluster passed the significance criterion in the ERAN time window ($p > .2$), the LAN time window ($p > .7$), or indeed the P600 time window between 450 and 700 ms after the linguistic disambiguation point ($p > .4$). We also looked for a spatio-temporal cluster anywhere between 160 and 700 ms without restriction. None emerged ($p > .06$).

In order to investigate the music × language interaction effect, we assessed spatial clusters in time windows of interest in which the language contrast (OR > SR) was different in the in-key and the out-of-key music conditions. No spatial cluster passed the significance criterion for either the ERAN time window ($p > .2$) or the LAN time window ($p > .06$). The response in one cluster passed a more liberal significance criterion ($p = .066$), see Figure 4|2. It is obvious that the neural response difference of the OR > SR contrast [$F_{(2,58)} = 7.40$, $p = .001$, ${}_{p}\eta^2 = .203$, $\omega^2 = .173$] is driven by activation differences between the in-key and the out-of-key conditions [$t_{(29)} = 4.23$, $p_{\text{corrected}} < .001$, $d = .775$, Bonferroni correction of *p*-value for two planned comparisons]. However, the difference between the auditory anomaly (attentional control) condition and the out-of-key condition is not significant [$t_{(29)} = 1.28$, $p_{\text{corrected}} = .422$, $d = .233$], suggesting that an attentional effect cannot be excluded for this cluster’s response.

In order to test the involvement of hypothesized shared music-language structural processing in prefrontal areas (Patel, 2003, 2008) we also defined a left and a right frontal region of interest consisting of the 33 sensors (all those named frontal according the standard channel naming convention of the CTF 275 system). Neither the left region of interest [$F_{(2,58)} < 1$] nor the right region of interest [$F_{(2,58)} < 1$] showed a main effect across the OR > SR contrast across the three levels of music (in-key, out-of-key, auditory anomaly). Finally, in order to fully describe the data, we also looked for a spatio-temporal cluster anywhere between 160 and 700 ms without restriction. None emerged ($p > .42$).

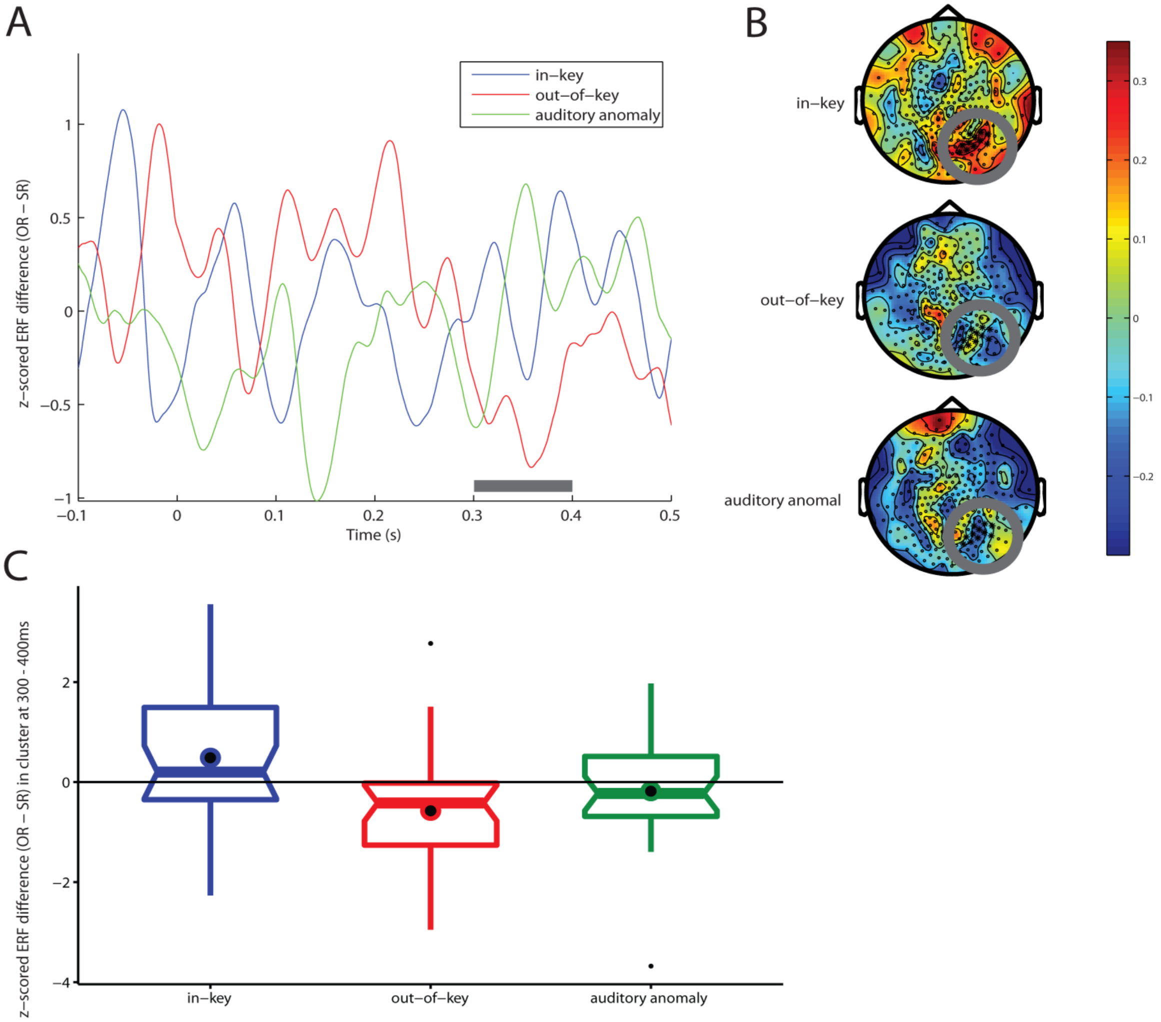


Figure 4|2. Neural response during the LAN time-window in the right posterior cluster. A) The low-pass filtered (10 Hz) response of the cluster highlighted in B (see grey circle). **B)** The topoplots of the LAN time-window highlighted in A (see grey square on x-axis). **C)** Neural response (OR > SR) in three different music conditions during LAN time-window in posterior cluster. Note that the cluster only passed a very liberal significance criterion (*p* = .066).

4|3.3 Time-frequency analysis results

In a first step, the language main effect (OR > SR) was investigated by assessing frequency-spatio-temporal clusters across frequencies (2 - 30 Hz), sensors, and time (100 - 800 ms). One positive cluster emerged (*p* = .005) covering all frequency bands from 4 to 30 Hz and all time-points from 100 to 700 ms. As shown in Figure 4|3, it had a changing spatial distribution depending on time and frequency driven mostly by frontal and parietal channels.

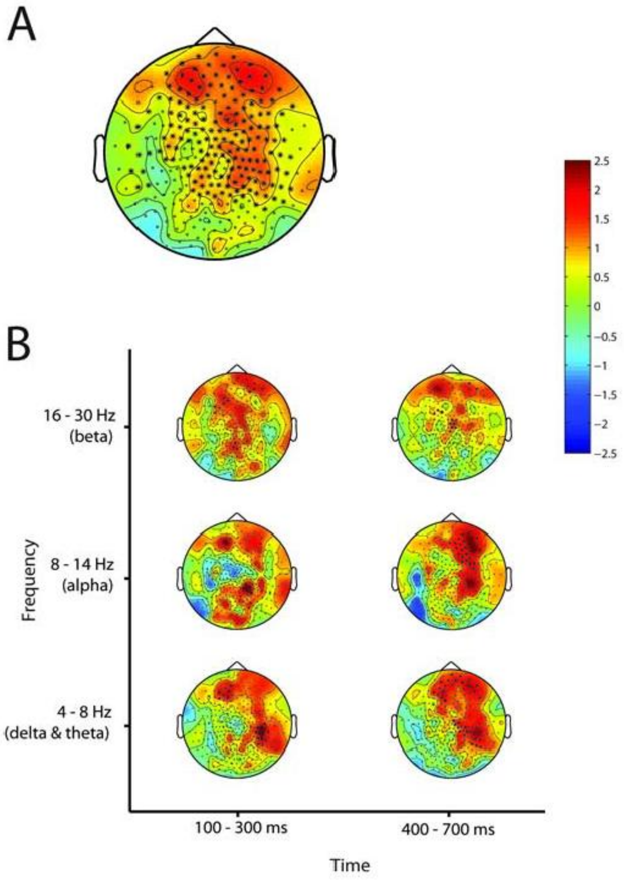


Figure 4|3. A frequency-spatio-temporal cluster with greater amplitude to OR than to SR sentences. A) Topoplot with all channels belonging to the cluster highlighted with asterisks showing the average signal over time (100 - 700 ms) and frequency (4 - 30 Hz). **B)** Development of the cluster over time and frequency. Note that the cluster passed the significance criterion (*p* = .005).

Regarding the predicted music × language interaction, there is one report of similar interactions in oscillatory brain responses by Carrus et al. (2011) in the delta-theta band (2 - 7Hz) between 350 - 700 ms. No similar effect emerged for the difference between the OR > SR contrasts in the in-key and the out-of-key conditions here. Even when looking at the delta (2 - 5 Hz) and theta bands (5 - 7 Hz) separately, no spatial cluster passed the significance criterion. Finally, in an effort to fully describe the data, we also decided to assess spatio-temporal clusters of the same difference between 100 and 800 ms after the disambiguation point for different frequency bands. No effects were observed in the delta ($p = .293$), theta ($p = .291$), alpha (7 - 13 Hz; $p = .288$), and beta bands (13 - 30 Hz; $p = .290$). Likewise, there were no effects when investigating frequency-spatio-temporal clusters ($p = .477$) assessed across frequencies (2 - 30Hz), sensors and time (100 - 800 ms). Similarly, the same analysis restricted to those time points, frequencies and sensors which were involved in the aforementioned OR > SR language main effect revealed no frequency-spatio-temporal clusters which passed the significance threshold ($p = .400$).

4|4 Discussion

4|4.1 Summary

The present study aimed to elucidate the temporal dynamics of shared music-language resources involved in structural integration. As opposed to similar, previous investigations of this question which focused on electro-magnetic brain signals, we avoided error and attention related confounds. Unexpectedly, we found no evidence for any shared music-language resources, whether in terms of a previously reported interaction between implied music harmony and language syntax in behaviour (Fedorenko et al., 2009), or in terms of ERF or time-frequency analyses of the MEG signal. While these findings could be taken to suggest a lack of shared music-language resources, we are hesitant to draw strong theoretical conclusions for the reasons discussed below.

4|4.2 Behavioural results

As opposed to a previous behavioural study of structural music-language interactions, we failed to find lower object-relative clause understanding when the stressed syllable of the relative-clause verb was sung on an out-of-key note, rather than an in-key note or an unusually loud note (Fedorenko et al., 2009). This lack of an effect is actually in line with the behavioural results of a previous fMRI study using the same critical stimulus set (Kunert et al., 2015). The authors of the fMRI study speculated that differences in experimental design might be responsible for the contrasting behavioural findings between Fedorenko et al. (2009) and them. However, in the present study these differences were reduced. In line with Fedorenko et al. (2009) but not Kunert et al. (2015), we used only a single kind of simple comprehension prompts and many filler trials. Moreover, a recent Dutch study has shown that a very similar linguistic stimulus set can lead to interactions with music when words are presented visually (Kunert et al., 2016). This renders the lack of a behavioural effect in the current study rather unexpected.

However, our stimuli were still more rhythmically diverse than in many previous studies (Fedorenko et al., 2009; Kunert et al., 2016). Moreover, the critical syntactic difference was based on the number agreement of the relative clause verb with a preceding noun phrase. Acoustically, such a difference is easily missed. Perhaps many participants defaulted, if unsure, to the simpler subject-relative clause interpretation. This might explain why 16 participants had to be excluded for poor behavioural performance and why the remaining participants showed relatively poor performance on object-relative clauses (M = 65% correct here compared to 81% in Fedorenko et al., 2009). Such a very low level of OR comprehension might have been difficult to further reduce through a music manipulation. Given that Fedorenko et al. (2009) used a word order manipulation and Kunert et al. (2016) used visually presented language stimuli, it was probably less easy to miss the critical language manipulation in these cases.

4|4.2 Neural results

We found no evidence for music (operationalised as the in-key versus out-of-key contrast) affecting language (operationalised as the OR versus SR contrast). However, when lowering the statistical threshold to also include marginally significant effects, a right posterior cluster emerged. Its response pattern powerfully shows why the lack of a control condition accounting for attentional capture effects of harmonically deviant tone/chords in previous studies (e.g., Steinbeis & Koelsch, 2008) is problematic for the interpretation of these studies' results. This cluster could easily have been misinterpreted as providing weak support for shared music-language resources in a right posterior brain area as a result of the modulation of the OR-SR difference by the music factor. However, the OR-SR difference was similarly influenced by the auditory anomaly condition, suggesting that the effect we measured represents a change in linguistic processing as a result of a previous attentional capture by the music domain. Whether harmonic or acoustic (unusual volume) properties led to the attentional capture appears not to be of great importance to this modulation of linguistic processing.

It is possible that we failed to observe other, previously reported, neural music-language interference effects because as opposed to previous studies, our music manipulation did not exactly coincide with the linguistic manipulation. While both targeted the relative clause verb, the music manipulation targeted its stressed syllable instead of the usually later, final morpheme targeted by the language manipulation. In 90% of the stimuli this led to a different syllable being targeted by the music and the language manipulations (mean syllable onset asynchrony of music- and language manipulation syllables = 887 ms). Previous investigations used only monosyllabic words (Fedorenko et al., 2009) or visual language presentation (e.g., Steinbeis & Koelsch, 2008) in order to exactly align music and language manipulations and, thus, maximize the potential to observe music-language interactions. A possible follow-up study could, therefore, improve on our design by singing the entire relative-clause verb on out-of-key notes.

4|4.3 Limitations

Next to the aforementioned limitations of our study related to the acoustically very subtle language manipulation and the slight asynchrony of our music and language manipulations, there are more general limitations to our design regarding the signal to noise ratio. First, there were only 40 trials per language × music condition, comparable to the 39 trials per condition in (Koelsch et al., 2005; Steinbeis & Koelsch, 2008) and 60 trials per condition in (Carrus et al., 2013). An average critical stimulus duration of 10.2 seconds did not allow for a higher trial count if one is limited to sessions of no more than 2.5 hours each. Longer sessions would likely be very taxing for the sustained attention abilities of participants, reducing task performance. Instead, we suggest that future investigations should pilot the stimulus tempo, increasing it if possible (as in Fedorenko et al., 2009).

Second, the behavioural data suggest that participants were often not even aware of the exact syntactic structure they had just heard. This suggests that the neural OR versus SR contrast includes many allegedly challenging OR trials which were really perceived as easier SR trials. Future studies might want to focus on the correctly answered trials only, which was not possible here due to the low trial count. Whether avoiding the limitations of our design truly leads to the discovery of the time-course of music-language interaction effects remains a question for future research.

4|5 Conclusions

The present study aimed to elucidate the temporal dynamics of shared music-language structural processing resources. Unexpectedly, the influence of such resources was not visible in our behavioural or MEG data. We hesitate to interpret this as evidence against the existence of shared music-language structural processing resources. Instead, we propose that our design was sub-optimal for finding answers to our research question. Future studies should take our limitations into account in order to optimize their design. In summary, future studies should consider 1) reducing the rhythmic diversity of stimuli, 2) use an acoustically salient linguistic manipulation which is not easily misheard, 3) temporally exactly align the point of structural integration difficulty in the language and music dimensions of the stimuli, 4) increase the presentation tempo, and 5) analyze the subset of trials with correct behavioural responses.

4|6 Acknowledgements

We would like to thank Dan Acheson for providing most of the stimulus material. This work was supported by an NWO Spinoza Prize awarded to PH and a PhD grant from the Max Planck Society to RK.

Chapter 5

Structural processing of music and language: imaging domain-specific neural resources

Music and language are uniquely human communication systems. How does the brain react when being exposed to them? In this manuscript we evaluate a theoretical model claiming that the music and language networks are largely distinct except for specific shared processing resources. Shared processing in the left inferior frontal gyrus is thought to reflect the integration of elements (words/chords) into higher order structures (sentences/harmony). Domain-specific circuitry in temporal areas is thought to store the necessary elements in long term memory. We adopt a neural priming (repetition suppression) paradigm previously used for language syntax investigations and validate it in a behavioural experiment. Participants listen to sentences sharing a syntactic structure with a previous sentence (primed) or not (unprimed), or they listen to piano chord sequences sharing a harmonic structure with a previous piano chord sequence (primed) or not (unprimed). As predicted, the left middle temporal gyrus was sensitive to syntactic priming in language but not to harmonic priming in music, suggesting that this brain region indeed includes neural representations specific to language. The inferior frontal gyrus reacted to neither syntactic priming in language nor to harmonic priming in music. Additional analyses reveal that with the current stimuli and tasks this brain area is involved in language and music processing in general. Overall, the current results support the evaluated model: in middle temporal brain areas domain-specific circuitry prevails. We find no effect for shared syntactic resources for language and music in this study, and we discuss reasons for why this may be the case.

Kunert, R., Roel M. Willems, Jason C. Rosenberg, Aniruddh D. Patel, Peter Hagoort (submitted). Structural processing of music and language: imaging domain-specific neural resources.

5|1 Introduction

This article focuses on two specific faculties, music and language, and asks to what extent these two cognitive domains rely on functional computations which are domain-specific (Fedorenko & Varley, 2016; Patel, 2008, 2013; Peretz & Coltheart, 2003). Using behavioural and functional magnetic resonance imaging (fMRI) experiments, we investigate whether there are domain-specific brain regions in temporal cortex responsive to either instrumental music's harmonic structure[3] or a linguistic stimulus's syntactic structure, as predicted by Patel's shared syntactic integration resource hypothesis (SSIRH; Patel, 2003, 2008).

Patel's SSIRH distinguishes between domain-general and domain-specific processing resources (2003, 2008). The proposal for domain-general processing resources is based on music and language both being communication conventions relying on richly structured auditory sequences. Their discrete elements (e.g., words in language, tones/chords in music) are combined according to implicitly learned rules in order to form higher order structures (e.g., sentences in language, harmonic sequences in music). A common brain area displaying domain-general processing of both language syntax and music harmony should result in language responses being influenced by music harmony manipulations, as seen in behaviour (Fedorenko et al., 2009; Fiveash & Pammer, 2014; Hoch et al., 2011; Slevc et al., 2009) and electro-encephalography (EEG; Carrus et al., 2011, 2013; Koelsch, Gunter, et al., 2005; Steinbeis &

---

[3] In the Western tonal tradition, harmony is represented by pitch organization in terms of scales and chords (simultaneous soundings of multiple pitches). Adjacent pitches (e.g., C and C#) are separated by a semitone and pitches of the same pitch class (e.g., all C notes on a piano) are separated by an octave. Typical scales include seven pitch classes (e.g., C, D, E, F, G, A, B in C major). When playing in a harmonic key (e.g., the key of C major), these pitch classes and their associated chords are emphasized. Different pitch classes/chords fulfill different functions. For example, the first scale degree (the tonic, e.g., C in C major) is the harmonic centre and most stable tone/chord. When preceded by the second most stable chord, the fifth scale degree (the dominant, e.g., G in C major), one speaks of an authentic cadence, which usually signals a moment of closure in tonal music. On the other hand, if the dominant is followed by the sixth scale degree (the submediant, e.g., A in C major), one speaks of a deceptive cadence, which can signal a musical phrase extension.

Koelsch, 2008). Furthermore, vice versa, responses to music harmony should be influenced by language syntax manipulations, as seen in behaviour (Kunert et al., 2016) and EEG (Steinbeis & Koelsch, 2008). These findings constitute compelling evidence against the notion of purely modular processing of music and language.

However, even though the SSIRH is mostly associated with shared processing resources common to music and language, it also predicts domain-specific resources based, in part, on obvious formal differences between both domains. For example, one of the basic ecological functions of language is to convey propositional thought. Despite some proposals for semantic processing in music (Koelsch, 2011), it remains difficult to imagine even basic semantic operations in music, e.g., compositionality (Slevc & Patel, 2011), negation (Jackendoff, 2009), and translation (Could a piece in the classical Western tradition like Dvořák's *9th* Symphony be translated into the classical Indian music system without loss of meaning?) Such formal differences between music and language might underlie neuropsychological dissociations between language problems and music problems after brain damage (Peretz & Coltheart, 2003), and an absence of interactions between music harmony and language *semantics* in behaviour (Kunert et al., 2016; Slevc et al., 2009; but see Perruchet & Poulin-Charronnat, 2013).

Based on evidence for shared as well as non-shared processing in music and language, Patel (2003, 2008) proposed the shared syntactic integration resource hypothesis (SSIRH). This dual-system model invokes shared resource circuitry in frontal brain regions involved in structural processing and domain-specific representation circuitry in temporal brain regions. Domain-specific representations are thought to constitute the stored knowledge of words and their syntactic features or of tone/chords and their harmonic features. Via long-distance neural connections, these domain-specific representations are recruited by the more frontally located shared resources during structural processing. This model can explain why 1) language and music problems can at times be dissociated after brain damage (impairment of temporal domain-specific representation circuitry), 2) there are neural and behavioural interactions between music harmony and language syntax (involvement of shared frontal structural processing circuitry), and 3) there are no consistent neural and behavioural interactions between music harmony and language semantics (functional specificity of shared circuitry for *structural* processing).

However, despite the good account the SSIRH provides for the experimental evidence so far, one of its central predictions (non-shared temporal music-language circuitry) has so far not been investigated. In order to investigate this prediction we borrow a paradigm used to investigate a dual-system model designed to explain the neural basis of language comprehension, the Memory-Unification-Control (MUC) model by Hagoort (2005, 2013). Specifically, a number of language studies have found evidence for both pre-frontal syntactic integration resources and temporal memory-related regions (Menenti, Gierhan, Segaert, & Hagoort, 2011; Segaert et al., 2012; Snijders et al., 2009). We adopt a syntactic priming paradigm used to find both kinds of processing resources in the brain and develop a music version of it (experiment 1). Then, we turn to fMRI (experiment 2) to investigate whether syntactic priming of language leads to repetition suppression effects in the left inferior frontal gyrus (integration resources) and the posterior middle temporal gyrus (language-specific memory representations) as seen before and predicted by the Memory-Unification-Control model. Moreover, harmonic structure priming of music is not predicted to lead to a repetition suppression effect in the language-specific posterior middle temporal gyrus (music-specific memory representations are located elsewhere).

5|2 Methods common to experiments 1 & 2

5|2.1 Analysis

We have deposited all music materials, raw data, and code for re-creating the analyses and figures on the OSF website at https://osf.io/59drt/?view_only=fff7bc27c4454b489454c6c2455e9c1c. We report results from standard null hypothesis significance testing throughout. Moreover, we include two different Bayesian analyses. First, in order to quantify relative model support for the null hypothesis (of no difference between conditions) versus model support for the alternative hypothesis (of a convincing difference between conditions; standard Cauchy prior with $r = \sqrt{2}/2$) we report Bayes factors using the BayesFactor package in R (Morey, Rouder, & Jamil, 2015; Rouder, Speckman, Sun, Morey, & Iverson, 2009). We follow standard convention for interpreting the model support indicated by Bayes factors (Jeffreys, 1961): $1 < BF < 3$ represents support that is not worth more than a bare mention; $3 < BF < 10$ represents substantial support, $10 < BF < 30$ strong support, and 30 < BF < 100 very strong relative model support. $BF_{01} > 1$ supports the null hypothesis, while $BF_{10} > 1$ supports the alternative hypothesis. Second, we also summarize the Bayesian posterior distribution which formally represents the belief in condition differences. The 95% Credible Interval based on 100,000 samples is a measure of uncertainty about this belief. For this analysis we use Krushke's BEST package in R (Kruschke, 2013; Meredith & Kruschke, 2015).

We adjust the analysis in case of problems with the normality assumption as indicated by visual inspection of distributions and the Shapiro-Wilk's test ($p < .05$). In terms of classical frequentist statistics, a non-parametric test from the family of random permutation based tests is used. These *t*-tests are based on a null hypothesis of *t*-values derived from testing within-subjects randomized data 50,000 times. The proportion of randomly obtained test statistics equal or more extreme than the true test statistic represents the *p*-value (p_{perm}). In terms of Bayesian analyses, there is to our knowledge no non-parametric Bayes factor analysis and, thus, none is reported. The analysis of the posterior distribution using the BEST package does not assume normality.

5|2.2 Stimuli

5|2.2.1 Music

Sixteen auditory music stimuli were created in major keys by Jason Carl Rosenberg, a professional composer (www.jasonrosenberg.org). All music stimuli are available at the aforementioned data repository. These piano chord sequences were all monorhythmic, uniform in amplitude, and 7750 ms long at 130 beats per minute (stimulus onset asynchrony (SOA) for chords = 462 ms) and observed Bach-style voice-leading principles in their top and bottom voices (four voices in total).

In addition to eight filler stimuli without a harmonic modulation, we designed eight critical stimuli including a harmonic modulation, whereby a sequence starts in one harmonic key lasting five beats, then two beats of modulatory pivot chords (harmonies shared by both the first and second harmonic keys), and then 5 beats of chords within a second key, see Figure 5|1. The stimulus ends in an authentic cadence in either the first key (creating an ABA harmonic structure) or the second harmonic key. Critical stimuli were organized in priming pairs sharing three harmonic organization principles: harmonic keys (e.g., D-major for first key, G-major for second key), chord function sequence (i.e. the sequence of tonics (I), dominants (V), subdominants (IV), etc.), and global structure (whether the final authentic cadence indicates a return to the first harmonic key or an end in the second harmonic key). Despite the shared harmonic organization, the stimuli of a priming pair sound very different due to strong stylistic differences.

The critical stimuli were organized in such a way as to ensure that an *un*primed stimulus was preceded by a critical harmonic sequence differing on all three harmonic organization principles. In addition to the eight critical stimuli, an equal amount of filler stimuli were composed. These were very similar to critical sequences except that they did not include any harmonic key changes and they ended either in an authentic cadence or a deceptive cadence. The visual stimuli for the audio-visual matching task were piano roll representations of the auditory stimuli.

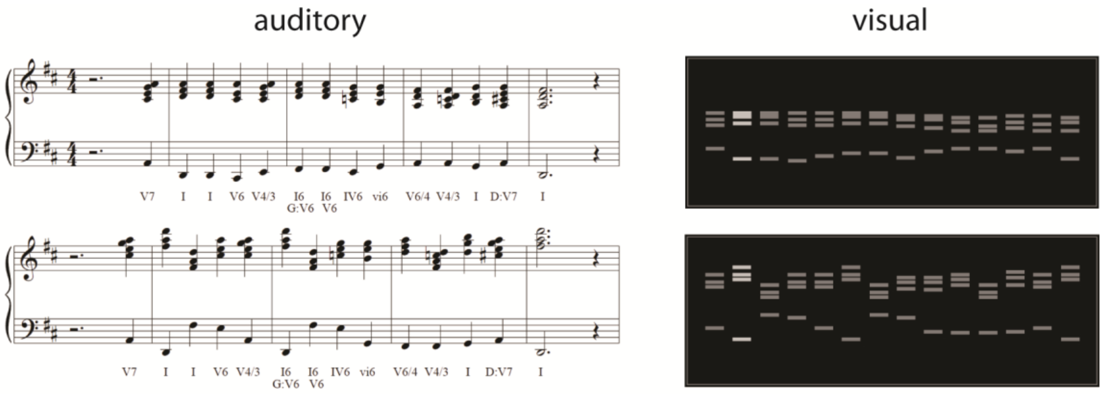


Figure 5|1. Example of a harmonic structure priming pair. Musical stimuli are played with a piano timbre at a constant rate (130 beats per minute). The stimuli in the top and bottom staves are separate stimuli in the experiment. They share three levels of harmonic structure: harmonic keys (D-major transitioning to G-major), chord function sequence (notice the indication of chord functions below the bass line), and global structure (modulation after 5 chords to a second key, modulation back to first key after another 7 chords). Note that despite the shared harmonic structures, the two stimuli vary stylistically and sound very different. On the right, the corresponding correct piano roll notations of the stimuli are shown with the second chord being highlighted. Participants were asked to judge whether the visual stimulus corresponded to the auditory stimulus. You can view and listen to all music stimuli online: https://osf.io/59drt/?view_only=fff7bc27c4454b489454c6c2455e9c1c.

5|2.2.2 Language

The language stimuli have been used before in fMRI repetition-suppression studies (Menenti et al., 2011; Schoot, Menenti, Hagoort, & Segaert, 2014; Segaert, Kempen, Petersson, & Hagoort, 2013; Segaert et al., 2012). All stimuli are in Dutch. Critical stimuli are based on 36 different transitive verbs (e.g., feeding, serving, etc.) performed by four different opposite sex couples (2 couples × man-woman, 2 × boy-girl). The resulting scene is described in active-voice (e.g., *The man feeds the woman.*) or passive-voice sentences (e.g., *The woman is fed by the man.*) presented visually (experiment 1) or auditorily (experiment 2). A priming pair shares both verb and syntactic structure but not actors, see Figure 5|2. If preceded by a non-priming stimulus, only the verb is repeated while the syntactic structure is changed.

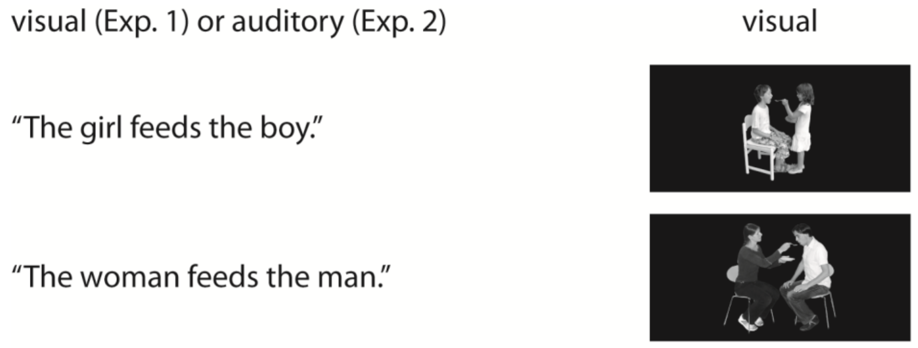


Figure 5|2. Example of a linguistic structure priming pair. On the left, two language stimuli (read silently in experiment 1, listened to in experiment 2) repeating the syntactic structure (active voice) and verb but not the actors. On the right, the corresponding correct photograph. Participants were asked to judge whether the visual stimulus corresponded to the auditory stimulus.

Moreover, filler items are included in order to vary the syntactic structures and lexical items in an experimental session. In the case of fillers, the depicted verbs are intransitive (e.g., singing, running) and only one actor and/or inanimate object is included in the sentence (e.g., *The baby cries.*). The visual stimuli for the matching task are gray-scale photographs depicting the sentence. For each couple, four photographs are taken crossing agent-patient role (male agent and female patient, male patient and female agent) and left-right arrangement (agent right and patient left, agent left and patient right).

5|2.3 Task

Music and language stimuli were presented in different blocks with a similar matching task. Participants were asked to judge whether a visual stimulus represents a sentence/musical piece (90% of trials) or not (10% of trials, only filler trials). Only a mismatch trial requires a button press, ensuring that attention is paid to the stimuli without eliciting motor activity during critical trials. Thus, the task is unrelated to the syntactic/harmonic contrasts of interest.

5|2.4 List composition

Stimuli were arranged in a running priming paradigm such that each target item also serves as the prime of the next target item (Schoot et al., 2014; Segaert et al., 2012). In each music or language block, participants were exposed to 250 trials divided into two parts separated by a brief pause. Music blocks reused stimuli for a maximum of 20 trials per participant following the suggestion that a single exposure might not be enough for a full analysis of the harmonic structure (Koelsch et al., 2013). Language blocks did not repeat trials for each participant.

The 250 trials per block were divided into 60 mini-blocks, half of which contained fillers and half of which contained critical trials. The 30 critical mini-blocks (size = 3 to 7 trials) included 30 trials in each of the four syntax/harmony (language: active/passive, music: first-key ending/second-key ending) × repetition (primed/unprimed) conditions as well as 30 prime only trials at the start of a mini-block. All critical trials were match trials (no response required). The 30 filler mini-blocks (size = 3 or 4 trials) included 74 picture-sentence or picture-music match trials (no response required) and 26 catch trials with a mismatch (button press required). Critical mini-blocks were alternated with filler mini-blocks. Each participant saw a different list of trials. The order of blocks (music first or language first) was counter-balanced. Performance feedback was provided 10 times per block (only within a filler mini-block so as to keep priming conditions intact).

5|3 Experiment 1 (behavioural)

5|3.1 Introduction

We first seek to behaviourally validate the priming task. In terms of music, harmonic priming has previously only been reported between chords (e.g., Tekman & Bharucha, 1998) or within musical phrases (e.g., Tillmann, Bigand, Escoffier, & Lalitte, 2006). We investigate whether it is also present between trials containing different musical sequences which share three levels of harmonic structure. In order to operationalise harmonic priming we make use of closure ratings which indicate how 'well finished' or 'complete' a musical piece sounds. Such judgments have previously been shown to indicate how well a chord ending harmonically fits with the preceding harmonic context (Bigand & Pineau, 1997; Kunert et al., 2016; Tillmann & Lebrun-Guillaud, 2005). It is predicted that a primed chord sequence leads to higher closure ratings compared to an unprimed one because priming facilitates harmonic processing leading to better harmonic fits of chord sequence endings. Only critical sequences ending on an authentic cadence (expected to result in generally high ratings) are tested.

Regarding language, Schoot et al. (2014) have shown that the present stimulus material and priming paradigm leads to syntactic priming effects in speech production. However, syntactic priming effects in *comprehension* have not previously been reported with this paradigm and they are generally less well established (e.g., Weber & Indefrey, 2009). In order to operationalise linguistic syntactic priming we presented sentences visually in a self-paced reading paradigm. We expect a primed sentence to be read faster than an unprimed one because priming facilitates syntactic processing.

5|3.2 Methods

5|3.2.1 Participants

We sampled 68 amateur musicians with at least seven years of formal musical training (*M* = 11 years, *SD* = 3.0 years, 26% male, 91% right handed) and Dutch as their native language. They were aged 24 years on average (*SD* = 7), reported no language impairments and had normal or corrected to normal vision. Written, informed consent was obtained at the start of a testing session. Compensation was provided financially or through course credits.

5|3.2.2 Procedure

A language block included two tasks: a self-paced reading task and a picture matching task. Each trial started with the reading task. After a variable inter-trial interval (ITI) between 0 and 2000 ms, a photograph appeared, followed by a readiness symbol and the central word-by-word presentation of a sentence (white on black background) whose timing was controlled by the subject through button presses. After the last word, the participant was given 1000 ms to press a response button to indicate that the photograph and the sentence did not match. In case of a match, no button press was required.

A music block also included two tasks: a closure rating task and a visual matching task. Each trial started with the visual matching task. After a variable ITI between 0 and 2000 ms, a visual piano roll notation (dark gray bars on black background) was shown without music for a variable SOA of 500 to 1500 ms, followed by the presentation of an auditory stimulus while the relevant chord position was highlighted in the visual stimulus (light gray bars). During auditory stimulation and up to 1000 ms afterwards (while the piano roll notation remained on the screen) participants could press a response button to indicate that picture and music mismatched (no button press required for match trials). Pilot experiments suggested that showing the picture only at the end of the trial, as in the language trials, rendered the music picture-matching task too difficult. Afterwards, participants were asked to rate their feeling of completeness of the auditory chord sequence on a scale from 1 (no completion) to 7 (perfect completion) within 3000 ms.

5|3.3 Results

5|3.3.1 Music

Regarding the closure rating task, we check participants' task adherence by analyzing the filler trials ending either in an authentic cadence (expected high closure rating) or an unusual deceptive cadence ending (expected lower ratings). Indeed, all participants gave higher ratings for authentic cadence endings (M = 5.81, SD = 0.80) than deceptive cadence endings (M = 2.99, SD = 0.85) ($t_{(67)} = 22.09$, $p < .001$, $\mathrm{BF}_{10} = 1.5 \times 10^{29}$, posterior $M_{\mathrm{diff}} = 2.83$, 95% Credible Interval = [2.57; 3.08]). Thus, participants appear to have provided meaningful closure ratings, allowing us to use closure ratings in order to investigate whether primed critical stimuli are more easily harmonically integrated (higher closure ratings) than unprimed sequences. This is indeed the case ($t_{(67)}$= 2.79, p_{perm} = .006, posterior M_{diff} = 0.07, 95% Credible Interval = [0.02; 0.12]), see Figure 5|3A. Finally, regarding the visual matching task, participants performed very well at 92% correct (*SD* = 6.53, *d* prime: *M* = 2.98, *SD* = 1.21), suggesting that the task is appropriate for use in the MRI scanner.

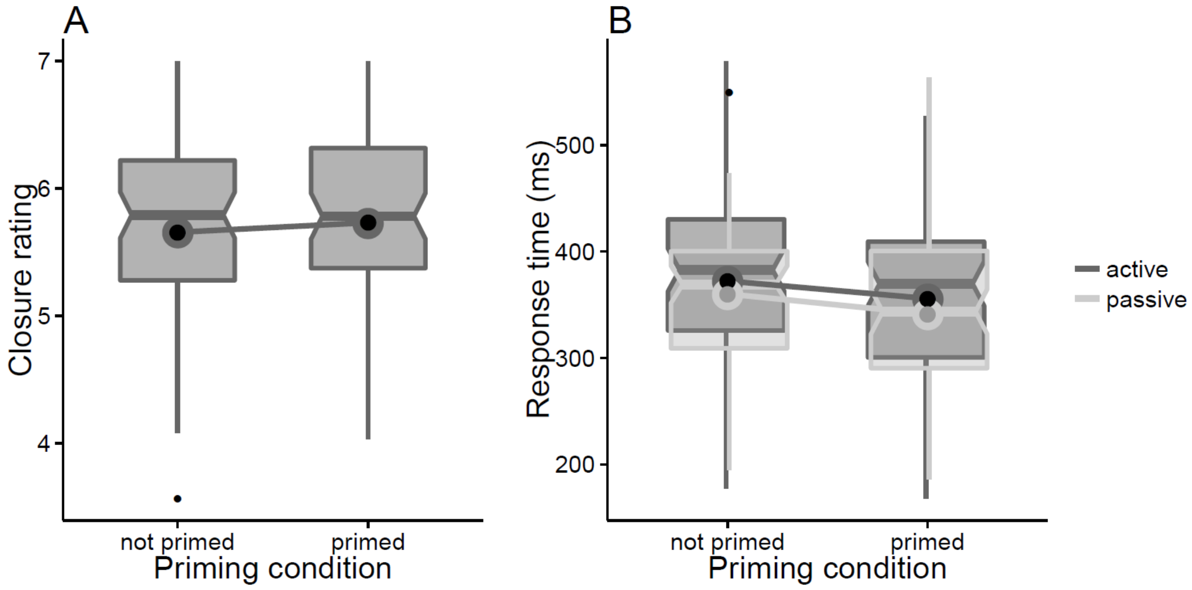


Figure 5|3. Behavioural priming effect in music (A) and language (B). A) Closure ratings of chord sequences preceded by a music stimulus with the same harmonic structure ('primed') or a different structure ('not primed'). Note that the difference between ratings of primed and unprimed sequences is small (M_{diff} = 0.07) but statistically significant (*p* = .006) as well as convincing in a Bayesian analysis (95% Credible Interval = [0.02; 0.12]). **B)** Reading times of active (dark) and passive (light) sentence verbs following a sentence with the same syntactic structure ('primed', e.g., active voice sentence preceded by active voice sentence) or the opposite syntactic structure ('not primed', e.g., active voice sentence preceded by passive voice sentence). Box plots display the interquartile range (IQR; top of box: upper quartile, middle line: median, bottom of box: lower quartile) together with whiskers representing the data points within a range corresponding to 1.5 times the IQR. Notches represent 95% Confidence Intervals of the medians. Big dots show mean values.

5|3.3.2 Language

Regarding the self-paced reading task, participants read the primed sentences faster than the unprimed sentences. To illustrate this pattern we focus on the verb reading time of all critical trials. In terms of outlier rejection, individual reading times are replaced by the cut-off value if they are 1) smaller than 50 ms, 2) greater than 2500 ms, or 3) more than 2.5 *SD* greater than the subject-specific mean of this segment (M = 4.1 % outliers, SD = 1.0 % across participants). Priming effects are seen both for active sentences ($t_{(67)} = 4.52$, $p_{\text{perm}} < .001$, posterior M_{diff} = 14.80 ms, 95% Credible Interval = [8.30; 21.65]) and passive sentences ($t_{(67)} = 5.62$, $p < .001$, BF_{10} = 36,300, posterior M_{diff} = 18.30 ms, 95% Credible Interval = [11.68; 24.99]), see Figure 5|3B. Regarding the picture matching task, all participants (except for one who never pressed the button) performed well at 96% (*SD* = 2.92, *d* prime: *M* = 3.44, *SD* = 1.16).

5|3.4 Discussion

Experiment 1 shows that the current paradigm results in musical and linguistic priming effects in terms of behaviour, suggesting that the stimulus material is well suited to investigate the neural basis of harmonic and syntactic priming in experiment 2.

5|4 Experiment 2 (fMRI)

5|4.1 Introduction

The aim of the neural investigation in experiment 2 is to reveal non-shared components of the music and language networks. Patel’s SSIRH (2003, 2008) predicts temporal regions to include domain specific brain circuitry related to musical representations or linguistic representations. While the concrete location of such a music-specific processing area is unknown (though see Norman-Haignere, Kanwisher, & McDermott, 2015), a host of language experiments suggests that the posterior middle temporal gyrus provides syntactic representations which the SSIRH hypothesises to be specific to language (Menenti et al., 2011; Segaert et al., 2012; Snijders et al., 2009).

The prediction regarding shared music-language processing depends on how the role of shared processing circuitry is defined. According to the SSIRH (Patel, 2003, 2008), shared music-language resources rapidly and selectively increase the activation level of an unexpected element's (word's/chord's) representation in posterior regions up to threshold after which structural integration can take place. According to this view, only the linguistic contrast should reveal shared processing as it includes relatively more expected (active voice) and less expected (passive voice) sentences. However, the musical material does not include a contrast between unexpected and expected chords as used in previous studies (Koelsch, Gunter, et al., 2005; Patel et al., 1998). Thus, the critical music stimuli likely do not strongly draw on shared music-language resources involved in structural integration.

However, recent behavioural evidence suggests that this characterization of the role of shared music-language resources might be too narrow. Specifically, Kunert et al. (2016) found that linguistic syntactic processing can interfere with holding a harmonic key online as an integration site for incoming chords. This was interpreted as evidence for shared music-language resources acting as a syntactic working memory system (Fiveash & Pammer, 2014) or as a unification workspace (Hagoort, 2005). Moreover, Van de Cavey & Hartsuiker (2016) found that global hierarchical structures can be primed from music to language comprehension, suggesting that the syntactic working component or unification workspace component partly processes the same syntactic representations in music and language.

Thus, in this view, one would expect the priming of musical harmonic structures to involve shared pre-frontal resources in the left inferior frontal gyrus, a location previously linked to shared music-language resources in an fMRI (Kunert et al., 2015) and a brain lesion study (Sammler et al., 2011). Moreover, this region is involved in syntactic priming in language (Menenti et al., 2011; Schoot et al., 2014; Segaert et al., 2012). Thus, this fMRI experiment will only reveal shared music-language resources involved in structural processing if in addition to increasing the activation level of unexpected words/chords these resources also act as a syntactic working memory or unification workspace.

The precise activation profile we predict is a repetition suppression effect, i.e. the reduction of a brain region's activity as a result of a repeated stimulus feature. Thus, we predict that unprimed harmonic sequences and sentences result in a greater BOLD response than primed sequences/sentences. The opposite pattern (repetition enhancement) is not predicted, and generally difficult to interpret (see Segaert, Weber, de Lange, Petersson, & Hagoort, 2013).

5|4.2 Methods

5|4.2.1 Participants

A total of 32 participants were sampled. None had participated in experiment 1. Five of them were rejected due to sound equipment failure (*N* = 1), excessive motion (*N* = 1), or MRI scanner problems (*N* = 3). The final sample (*N* = 27) were all amateur musicians with at least seven years of formal musical training (*M* = 11 years, *SD* = 3.6 years, 35% male, 100% right-handed) and Dutch as their native language. They were aged 23 years on average (*SD* = 4.9), reported no neurological or language impairments and had normal or corrected to normal vision. Written, informed consent was obtained at the start of a testing session. Compensation was provided financially or through course credits.

5|4.2.2 Procedure

A language block included only the aforementioned picture matching task. Each trial started with a variable ITI between 0 and 2000 ms, followed by a photograph shown 500 to 1500 ms before sentence onset, followed by the audio presentation of a sentence while the photo remained on the screen. Participants could press a button during and up to 1500 ms after the sentence to indicate that the sentence and the picture mismatched. A music block was the same as in experiment 1 except that participants did not perform the closure rating task, instead focusing solely on the visual matching task with played chords being highlighted in the piano roll notation, see Figure 5|1. Button presses were recorded from audio onset until 1000 ms after audio offset.

5|4.2.3 fMRI Data acquisition

Participants were scanned with a Siemens 3-T Skyra MRI scanner (Siemens Medical system, Erlangen, Germany), using a multiecho echo-planar imaging (EPI) sequence, in which images are acquired at multiple time echoes (TEs) following a single excitation [time repetition (TR) = 2.070 s; each volume consists of 34 slices of 3 mm thickness with slice gap of 17%; isotropic voxel size = 3.5 × 3.5 × 3.5 mm; field of view (FOV) = 224 mm, 90° flip-angle (FA)]. The functional images were acquired at the following TEs: TE1 at 9.0 ms, TE2 at 19.3 ms, TE3 at 30 ms, and TE4 at 40 ms. In the middle of the scanning session a 3-D T1 scan was acquired (T1-weighted MPRAGE, 192 slices, voxel size = 1 mm × 1 mm × 1 mm, TR = 2300 ms, TE = 3.03 ms, FA = 8°, FOV = 256 mm, sagittal orientation).

5|4.2.4 fMRI Data analysis

The analysis is carried out using SPM8 (www.fil.ion.ucl.ac.uk/spm). The four echoes are realigned to correct for motion artifacts (estimation of the realignment parameters is done for the first echo and then copied to the other echoes). The four echo images are combined into a single MR volume based on 36 volumes acquired before the actual experiment started using an optimised echo weighting method (Poser, Versluis, Hoogduin, & Norris, 2006). Slice timing correction is applied by means of linear interpolation to the onset of the first slice. Structural and functional data are then co-registered, spatially normalised to a standardized stereotactic space (Montreal Neurological Institute (MNI) template), and spatially smoothed using an 8mm FWHM Gaussian kernel.

First- and second-level statistics are performed using the general linear model framework of SPM8. A synthetic hemodynamic response function modeled the BOLD response from the beginning to the end of the auditory stimulus. There are four regressors modeling critical music trial responses (primed and first key ending, primed and second key ending, not primed and first key ending, not primed and second key ending) and four modeling critical language trial responses (active primed, passive primed, active not primed, passive not primed). Moreover, there are three regressors for music and language, each modeling prime-only trials, and picture-audio mismatch catch trials (filler trials) and picture-audio match trials (filler trials). Two additional regressors account for button click responses and white-noise only null trials.

Furthermore, 18 motion parameters account for linear motion, its quadratic effect and its first derivative (Lund et al., 2005). Finally, two compartment signal parameters (white matter and cerebral spinal fluid) correct for global intensity fluctuations (see Schoot et al., 2014 and Segaert et al., 2012 for a similar approach).

The second-level analysis is performed based on first-level contrast maps of each participant which include all of his/her trials (not taking behavioural performance into account). Participant is treated as a random factor ('random effect analysis'). Cluster size is chosen as the test statistic as standardly implemented in SPM8 (based on an uncorrected voxel-wise threshold of $p < .001$) and only clusters significant at $p < .05$ (corrected for multiple comparisons) are reported together with their peak voxels, see Table 5|1 and Table 5|2. For the region of interest analysis we use Marsbar (Brett et al., 2002) and the same syntactic priming related regions as Menenti et al. (2011) and Schoot et al. (2014): the left inferior frontal gyrus and the left posterior middle temporal gyrus.

5|4.3 Results

5|4.3.1 Behaviour

The visual matching task was performed well during music blocks (*M* = 93% correct, *SD* = 2.9, *d* prime: *M* = 3.43, *SD* = 1.0) and language blocks (*M* = 98% correct, *SD* = 1.0, *d* prime: *M* = 3.99, *SD* = 0.5). This suggests that participants paid attention to the stimuli.

5|4.3.2 Whole brain

5|4.3.2.1 Language

As shown in Figure 5|4 and Table 5|1, there are two regions showing the expected repetition suppression response pattern as a result of repeated language syntax (not primed > primed): left middle temporal gyrus extending into the superior temporal sulcus (BA 21 and BA 22), and the left inferior parietal cortex extending into the intra-parietal sulcus (BA 40). Two regions display unpredicted repetition enhancement effects (primed > unprimed): the right inferior temporal gyrus (BA21 and BA 37) and the left cerebellum (lobules V and VI). The inferior frontal gyrus is unaffected by the syntax repetition manipulation. In order to check whether with this paradigm it reacts to language syntax at all, we contrast reactions to passive and active sentences. This contrast does show reactivity of the left inferior frontal gyrus (BA 46 and BA 44) to syntax, see Table 5|1 for details.

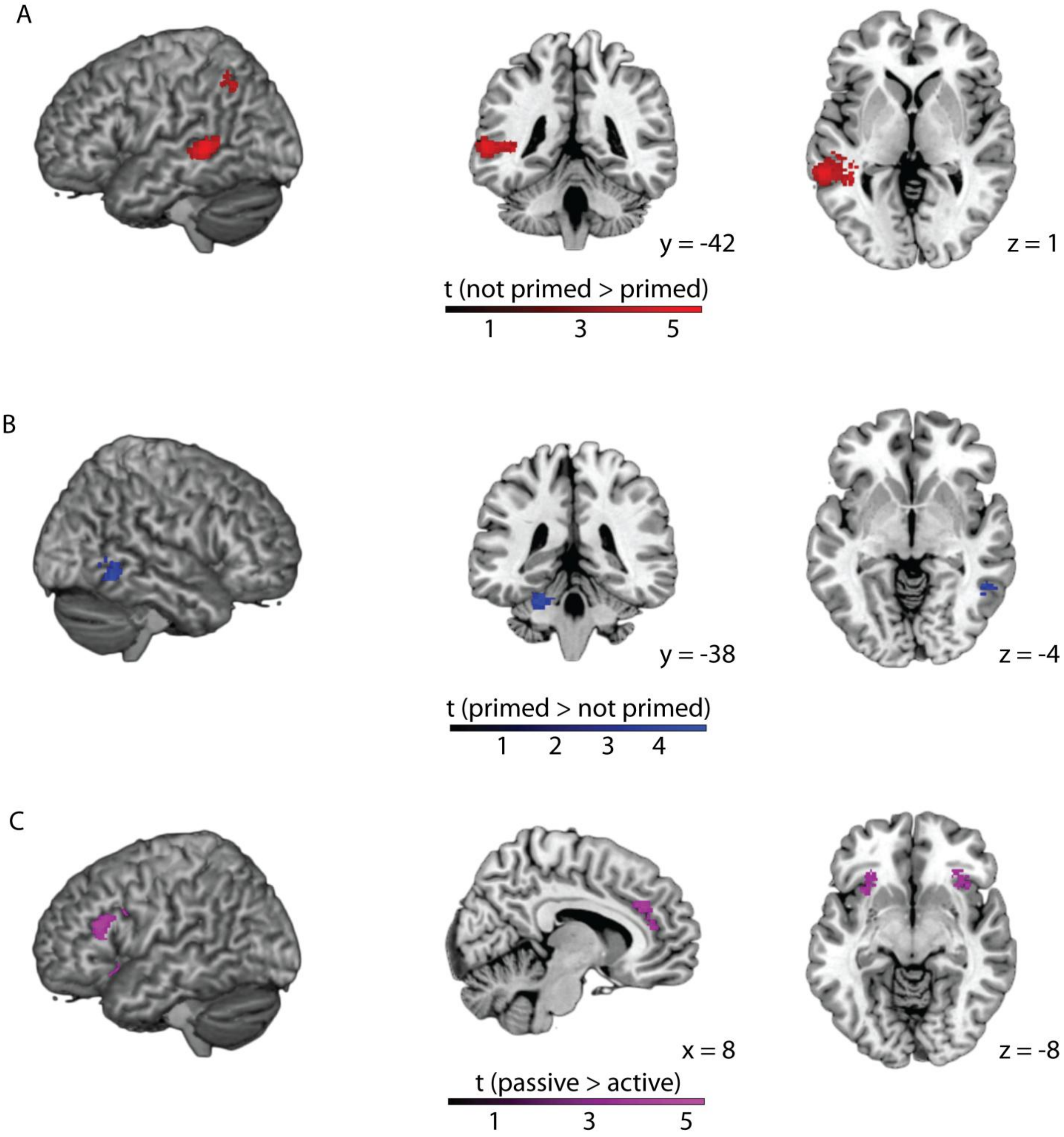


Figure 5|4. BOLD response in language blocks. A) Positive BOLD response difference (repetition suppression) to syntactically primed versus unprimed sentences. **B)** Negative BOLD response difference (repetition enhancement) to syntactically primed versus unprimed sentences. **C)** BOLD response to passive voice > active voice sentence contrast (independent of priming status).

Table 5|1. Whole brain analysis of language effects

Anatomical label	**BA**	**Global and local maxima (MNI)**			**Cluster-level**		**Voxel-level**
		X	***Y***	***Z***	***K***	***P*****(corrected)**	***Z***
not primed > primed: positive syntactic repetition effect (repetition suppression)							
Left middle temporal gyrus	21	-60	-38	0	526	<.001	4.51
Left middle temporal gyrus	22	-50	-24	0			4.36
Left middle temporal gyrus	22	-54	-42	4			4.19
Left inferior parietal lobule	40	-34	-56	40	133	.002	4.15
Left inferior parietal lobule	40	-42	-56	50			3.34
primed > not primed: negative syntactic repetition effect (repetition enhancement)							
Right inferior temporal gyrus	37	52	-54	-12	110	.008	4.05
Right inferior temporal gyrus	37/22	50	-50	-4			3.75
Right middle temporal gyrus	37	48	-58	0			3.44
Left Cerebellum (V)	-	-22	-38	-28	107	.009	4.00
Left Cerebellum (VI)	-	-26	-44	-24			3.97
Left Cerebellum (V - VI)	-	-30	-36	-32			3.22
passive > active (independent of priming status)							
Left hippocampus	-	-30	-16	-10	115	.007	4.38
Close to left olfactory cortex	-	-22	10	-12			3.64
Left inferior frontal gyrus (pars orbitalis)	47	-30	28	-10			3.62

Right anterior cingulate cortex	32	14	38	14	133	.003	4.13
Right anterior cingulated cortex	32	8	34	28			3.85
Right anterior cingulated cortex	32	5	36	16			3.45
Left inferior frontal gyrus (pars triangularis)	45/46	-46	26	20	172	<.001	4.09
Left inferior frontal gyrus (pars opercularis)	9	-36	12	30			3.65
Right inferior frontal gyrus (pars orbitalis)	47	32	32	-4	82	.036	3.99
Right insula	-	38	20	-8			3.64
Right insula	47	30	20	-12			3.52

5|4.3.2.2 Music

No whole-brain significant cluster related to harmonic priming emerges with a voxel threshold of *p* = .001 (uncorrected). In order to fully describe the data, we lower the voxel threshold to *p* = .005 (uncorrected) and report whole-brain significant clusters (*p* < .05 corrected for multiple comparisons) at this threshold. As shown in Figure 5|5A and Table 5|2, there are two clusters showing a repetition suppression effect to repeated music harmony (not primed > primed): an occipital cluster (BA 18) extending into the cerebellum and a subcortical cluster in the right putamen. The inferior frontal gyrus is not responsive to the harmonic structural repetition manipulation. No brain area displays repetition enhancement effects.

Finally, we attempt to investigate harmonic processing difficulty akin to syntactic processing difficulty which was investigated by means of the passive > active contrast. However, there were no critical trials with difficult to integrate chords as used in previous studies (Hoch et al., 2011; Koelsch et al., 2005; Slevc et al., 2009; Steinbeis & Koelsch, 2008). Therefore, we assume that a music stimulus' average closure rating in experiment 1 (independent of priming status) reflects harmonic processing difficulty (see Kunert et al., 2016) and construct a regressor capturing the parametric effect of a sequence's average closure rating. This regressor captures the difference between authentic cadence endings (all critical trials, half of filler trials) and harmonically less expected deceptive cadence endings (half of filler trials) as well as variability within each cadence type ending. However, this additional regressor cannot be included in the general GLM reported so far because its inclusion renders other regressors unestimable. Therefore, a new GLM modeling only music block data with the regressors 'correct filler trial', 'catch filler trial', 'parametric modulation to harmonic processing difficulty of all music stimuli', as well as the aforementioned nuisance regressors is run.

Bilaterial perisylvian regions, parietal, prefrontal, temporal, and occipital/cerebellar brain areas react to music harmony integration difficulty, see Figure 5|5B and Table 5|2. There are two things to note. First, the left middle temporal activation to music harmony processing difficulty observed with this analysis lies posterior and inferior to the middle temporal gyrus area revealed by the linguistic repetition suppression effect, compare Figures 5|4A and 5|5B. Second, the music-harmony related cluster encompassing mostly the left superior temporal gyrus extends into the left inferior frontal gyrus (BA 45). It does not, however, overlap with the inferior frontal gyrus activation related to the language syntax contrast which is more superior, see Figure 5|4C. In general, this additional analysis shows that with our stimuli the brain reacts quite differently to music harmony (Figure 5|5B) than to language syntax (Figure 5|4C).

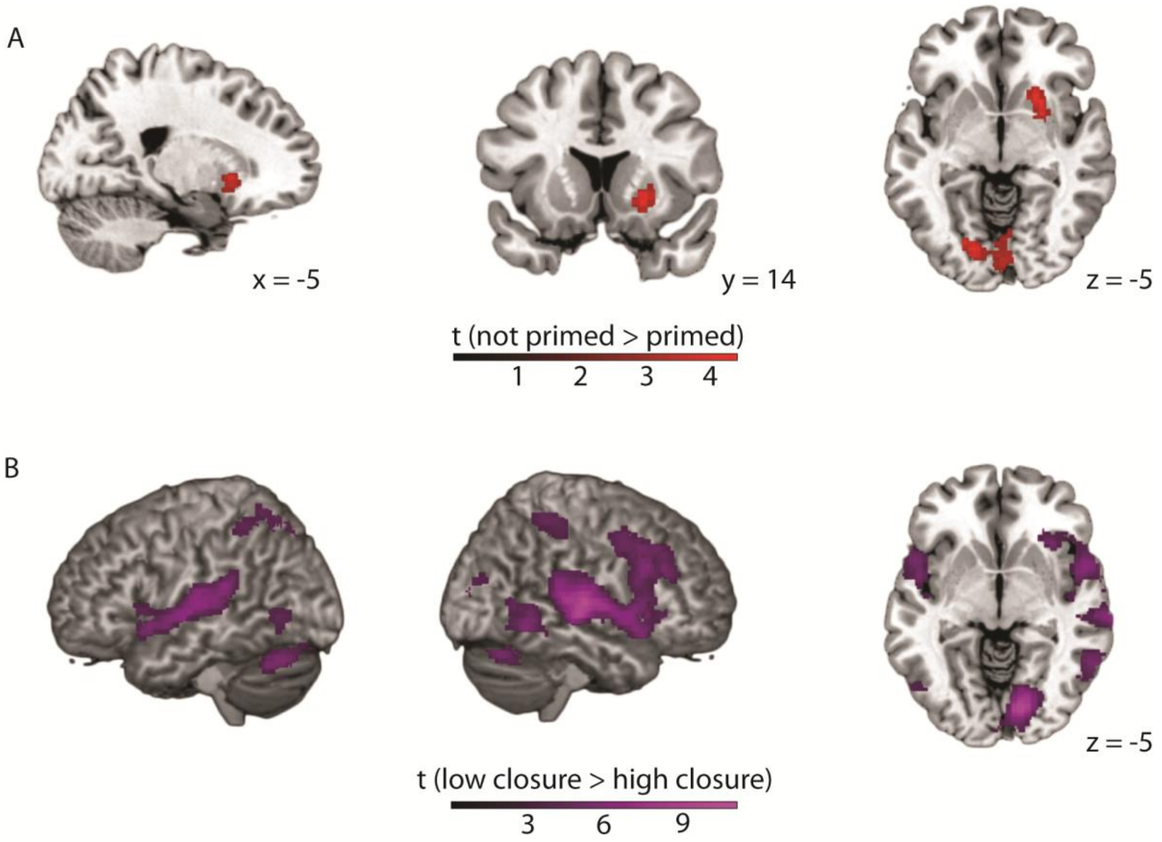


Figure 5|4. BOLD response in music blocks. A) Positive BOLD response difference (repetition suppression) to harmonically, structurally primed versus unprimed piano chord sequences. **B)** Parametric modulation of BOLD signal with harmonic difficulty (as defined by harmonic closure ratings from experiment 1) independent of priming status.

Table 5|2. Whole brain analysis of music effects

Anatomical label	**BA**	**Global and local maxima (MNI)**			**Cluster-level**		**Voxel-level**
		X	***Y***	***Z***	***K***	***P*(corrected)**	***Z***
not primed > primed: positive harmonic repetition effect (repetition suppression)							
Left lingual gyrus	18	-16	-82	-4	397	.001	3.77

Left cerebellum (VI)	-	-6	-72	-12			3.51
Left fusiform gyrus	18	-22	-76	-6			3.41
Right putamen	-	24	14	-4	218	.026	3.67
Right putamen	-	26	4	-6			3.54
primed > not primed: negative harmonic repetition effect (repetition enhancement)							
No significant clusters							
parametric modulation by harmonic closure (independent of priming status)							
Left superior temporal gyrus	22	-48	-12	0	2191	< .001	5.24
Left superior temporal gyrus	22	-52	-20	6			5.18
Left superior temporal gyrus	40	-66	-20	12			5.10
Right lingual gyrus	18	14	-82	-4	1231	< .001	6.67
Left calcarine gyrus	17	-4	-94	-10			3.69
Right middle occipital gyrus	19	30	-86	18			3.52
Right superior temporal gyrus	22	68	-18	4	5481	< .001	6.52
Right superior temporal gyrus	42	62	-20	10			5.89
Right superior temporal gyrus	42	64	-30	8			5.85
Left cerebellum (VIII)	-	-20	-68	-42	2194	< .001	5.53
Left cerebellum (VII)	-	-12	-74	-42			5.20
Left cerebellum (VII)	-	-34	-68	-28			4.83
Right inferior temporal	21	62	-54	-8	421	< .001	4.59

gyrus							
Right middle temporal gyrus	21	46	-60	2			4.07
Right middle temporal gyrus	21	56	-50	0			3.96
Left superior parietal lobule	7	-20	-66	54	515	< .001	4.07
Left superior parietal lobule	7	-22	-58	60			3.76
Left inferior parietal lobule	7	-44	-44	52			3.71
Left middle temporal gyrus	37	-50	-68	0	193	.020	3.97
Left middle occipital gyrus	37	-38	-64	0			3.51
Left middle occipital gyrus	37	-42	-72	-2			3.50
Right inferior parietal lobule	40	44	-42	54	237	.008	3.97
Right inferior parietal lobule	40	50	-36	54			3.83
Right inferior parietal lobule	40	38	-40	44			3.21

5|4.3.3 Regions of interest

As predicted, the anatomical region of interest in the left middle temporal gyrus does not display a repetition suppression effect to the repetition of harmonic structure ($t_{(26)} = 1.31$, $p = 0.202$, $BF_{01} = 2.28$, posterior $M_{diff} = 0.05$, 95 % Credible Interval = [-0.03; 0.12]), see Figure 5|6A. In contrast to this, and in line with our predictions, the repetition of linguistic syntactic structures results in repetition suppression ($t_{(26)} = 4.55$, $p < 0.001$, $BF_{10} = 239.97$, posterior $M_{diff} = 0.76$, 95% Credible Interval = [0.41; 1.12]), see Figure 5|6B. The neural priming effect is greater in language than in music ($t_{(26)} = 4.38$, $p < .001$, $BF_{10} = 161.19$, posterior $M_{diff} = 0.69$, 95% Credible Interval = [0.36; 1.04]).

The left inferior frontal gyrus region of interest does not display repetition suppression effects in either cognitive domain. Neither music ($t_{(26)} = 1.20$, $p = 0.239$, $BF_{01} = 2.56$, posterior $M_{diff} = 0.05$, 95% Credible Interval = [-0.04; 0.15]) nor language ($t_{(26)} = 0.14$, $p = 0.891$, $BF_{01} = 4.87$, posterior $M_{diff} = 0.03$, 95% Credible Interval = [-0.37; 0.42]) display a repetition suppression effect related to harmonic/syntactic structure, see Figure 5|6C and Figure 5|6D. In order to check whether this brain region is responsive to language syntax and music harmony at all, we run the passive > active sentence contrast (independent of priming status) and the parametric modulation to harmonic processing difficulty contrast (independent of priming status) in this ROI. Neither evidences a strong response even though the language contrast is statistically significant ($t_{(26)}$ = 2.45, p = 0.021, BF_{10} = 2.49, posterior M_{diff} = 0.40, 95% Credible Interval = [0.05; 0.75]) while the music contrast is not ($t_{(26)}$ = 1.46, p = 0.157, BF_{01} = 1.91, posterior M = 0.03, 95% Credible Interval = [-0.02; 0.09]).

We decided to further investigate the absent repetition suppression effects in the left inferior frontal gyrus ROI. It could be hypothesized that only a subset of voxels in and around this brain region is responsive to syntax. Moreover, the precise voxels probably differ from person to person. Therefore, we use a functional region of interest approach extracting the 10% most responsive voxels of each participant to the passive voice > active voice contrast in Broca's area (pars opercularis and pars triangularis) as defined by the Automatic Anatomical Labeling toolbox (Nieto-Castañón & Fedorenko, 2012; Tzourio-Mazoyer et al., 2002; see Kunert et al., 2015 for a similar approach). The neural priming effect in syntax-responsive voxels is subsequently estimated. Two-fold cross-validation is applied in order to ensure independence between localization of voxels and estimation of the language syntax priming response (Kriegeskorte et al., 2009). Neither music ($t_{(26)}$ = 0.44, p = 0.660, BF_{01} = 4.48, posterior M = 0.01, 95% Credible Interval = [-0.03; 0.05]) nor language ($t_{(26)}$ = 1.18, p = 0.249, BF_{01} = 2.62, posterior M_{diff} = 0.04, 95% Credible Interval = [-0.04; 0.13]) display a repetition suppression effect related to harmonic/syntactic structure in voxels which are both located in Broca's area and responsive to language syntax, see Figure 5|6E and Figure 5|6F.

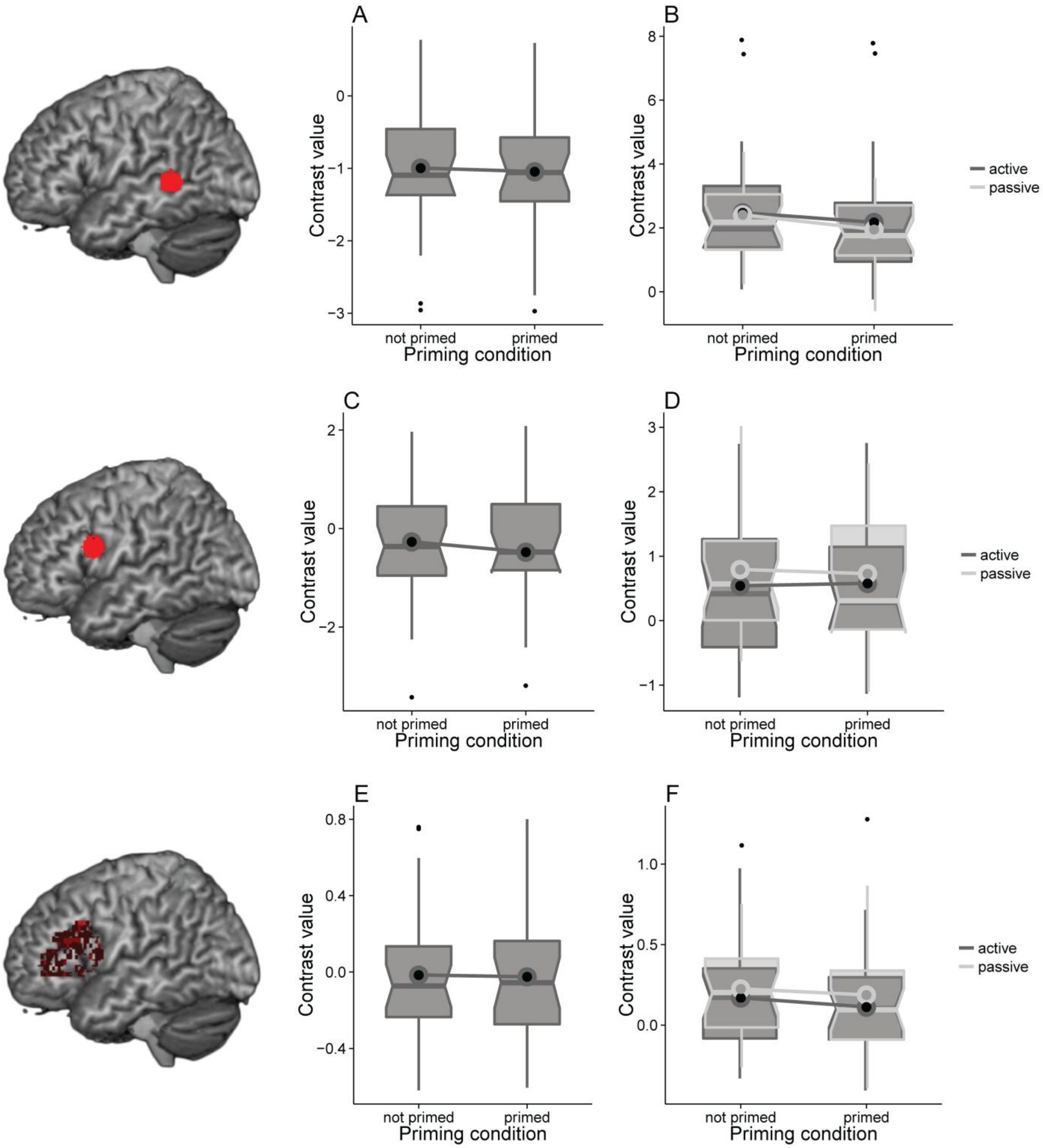


Figure 5|5. Region of interest results. A) BOLD response to harmonic structural priming in the left middle temporal gyrus. This anatomical region of interest is not hypothesized to respond to music structure. Its location is based on similar language studies of syntactic priming. **B)** BOLD response to linguistic, syntactic priming in the same region. This ROI is hypothesized to respond to language syntax. **C)** Music results in the left inferior frontal gyrus. This brain region is not reactive to music harmony. Its location is based on similar language studies of syntactic priming. **D)** Language results in the left inferior frontal gyrus. This ROI is not reactive to syntactic priming (contrary to prediction) but does react to language syntax (passive voice > active voice, in line with predictions). **E)** BOLD response to harmonic structural priming in the functional region of interest (fROI defined as the 10% most active Broca's area voxels to the passive > active contrast for each participant separately). Again, no repetition suppression effect to harmonic structural priming is observed. **F)** Language results in the same functional region of interest. Again, no convincing repetition suppression effect to syntactic priming is observed. Box plots display the interquartile range (IQR; top of box: upper quartile, middle line: median, bottom of box: lower quartile) together with whiskers representing the data points within a range corresponding to 1.5 times the IQR. Notches represent 95% Confidence Intervals of the medians. Big dots show mean values.

5|4.4 Discussion

In experiment 2 we sought to find evidence for domain-specific circuitry in the temporal lobe. This prediction receives support. While language stimuli lead to a syntactic repetition suppression effect in the left middle temporal gyrus, music stimuli do not lead to the same pattern. Instead, harmonic repetition suppression effects are observed in the occipital lobe and the right putamen.

The right putamen has previously been found to be involved in syntactic ambiguity processing (Snijders et al., 2009), in line with lesion evidence for basal ganglia (of which the putamen is a part) involvement in the inhibition of competing alternatives during language processing (Copland, 2006). Perhaps harmonic processing involves the selection of appropriate harmonic interpretations and chord functions via basal ganglia mediated circuitry (see Basal ganglia involvement in musical unexpectancy processing: Koelsch et al., 2002; Seger et al., 2013; Tillmann et al., 2003). Because a primed sequence is preceded by the exact same harmonic keys and chord function sequence, it is seen as less ambiguous, requiring less basal ganglia activity to select the appropriate harmonic interpretation and inhibit competing alternatives. The application of the behavioural priming paradigm presented in experiment 1 to neuropsychological patients with basal ganglia damage could provide a critical test for this suggestion.

We find no evidence for common syntactic and harmonic priming-related BOLD response patterns in the left inferior frontal gyrus. In terms of music, such a pattern was not predicted by the SSIRH itself, but instead by an additional characterization of shared music-language resources as a syntactic working memory or a unification workspace (Fiveash & Pammer, 2014; Kunert et al., 2016). However, both a whole-brain analysis as well as more focused anatomical and functional region of interest analyses fail to find evidence for any harmony-related activity in this brain region. This is not in contradiction to the SSIRH (Patel, 2003, 2008) which only predicts a left inferior frontal involvement in the case of difficult-to-integrate chords which were not used here. An additional analysis of brain activity related to harmonic processing difficulty independent of inter-trial priming might relate to this prediction. This analysis does reveal some limited left inferior prefrontal involvement (in line with the intra-trial chord priming literature, e.g., Tillmann et al., 2003). Thus, while this area appears to be involved in music harmony processing in general, its precise role (e.g., as a syntactic working memory or a unification workspace) could not be elucidated here.

In terms of language, as opposed to previous studies (Menenti et al., 2011; Schoot et al., 2014; Segaert, Kempen, et al., 2013; Segaert et al., 2012), we find no evidence for the left inferior frontal gyrus being sensitive to syntactic priming. However, it displays more activity to the more challenging passive voice sentences than active voice sentences, suggesting that this brain region is involved in language processing generally even though its precise role cannot be elucidated here. It might be that we fail to find pre-frontal priming-related activation patterns because participants only performed a comprehension task which might lead to a task set with shallow/incomplete parsing routines (Ferreira et al., 2002). In support, language studies with fMRI repetition suppression paradigms without a production component do not consistently reveal a pre-frontal involvement (Noppeney & Price, 2004; Sammler et al., 2010; but see Santi & Grodzinsky, 2010; Weber & Indefrey, 2009).

5|5 General discussion

Two experiments are presented with the aim of critically evaluating the proposal for domain-specific brain circuitry related to structural processing of music and language. The current results provide support for the model: the non-shared model component in temporal cortex behaves as expected. In detail, a temporal brain region (middle temporal gyrus) does not respond to music but shows a language syntax priming response. Thus, this region is a good candidate for storing language-specific knowledge about words and their syntactic properties in long term memory, as hypothesized by the SSIRH (Patel, 2003, 2008) and the language-specific MUC model (Hagoort, 2005, 2013).

Two regions (right putamen and left occipital cortex) exhibit a similar neural response profile for music. Given that these music-specific regions lie outside the predicted temporal cortex, we are hesitant to locate the storage of musical elements (tones/chords) and their structural properties there. The exact function of these regions for music harmony processing should be elucidated by future research. Concerning the right putamen, we hypothesise a critical role of this brain structure for harmonic ambiguity resolution.

The inferior frontal gyrus results are partly unexpected. Neither music nor language stimuli lead to strong priming effects in this brain region. However, recent behavioural priming (Van de Cavey & Hartsuiker, 2016) and language-music interference evidence (Kunert et al., 2016) suggest that shared resources might be involved in syntactic working memory or related functions (Fiveash & Pammer, 2014). This would predict a priming effect in the left inferior frontal gyrus which was not found. The absent pre-frontal priming effect could also relate to participants' task strategy. In the current investigation, participants might have simply abandoned the syntactic analysis once difficulties were encountered, rather than involving the inferior frontal gyrus. This might explain why other investigations of repetition suppression to music (melody: Sammler et al., 2010) without an active task also did not reveal a prefrontal involvement. Support for this speculation could come from a repetition of experiment 2 with a music production task which is predicted to show a repetition suppression effect to harmonic structure in the left inferior frontal gyrus if this region fulfills a syntactic working memory function (Fiveash & Pammer, 2014).

The SSIRH (Patel, 2003, 2008) did not predict a repetition suppression effect for music in this brain region given that participants were not exposed to difficult-to-integrate chords in critical music stimuli. However, when investigating the brain activity related to harmonic processing difficulty of all music stimuli, including filler trials with deceptive cadence endings which are harmonically relatively unexpected, some inferior pre-frontal involvement was found. While this activation profile is perhaps not strong enough to provide strong support for the SSIRH's shared music-language resources in pre-frontal areas, it shows that this data set does not contradict the proposal for left pre-frontal involvement in music processing either.

5|6 Conclusion

The current set of experiments provides evidence for distinct neural resources related to the storage of linguistic representations in long term memory which are not shared with music. This suggests that syntactic processing of music and language is not wholly shared (Patel, 2003, 2008; Peretz & Coltheart, 2003). Instead, the music and the language networks share a subset of their extensive circuitry for higher order structural processing. However, given the stimulus material and tasks, we failed to reveal such a shared component which previous behavioural and neural studies suggest exists (Kunert & Slevc, 2015; Kunert et al., 2015). Therefore, this study answers just a part, albeit an important part, of the question of how music and language relate to each other in the brain.

5|7 Acknowledgements

We would like to thank Katrien Segaert and Lotte Schoot for providing the linguistic stimulus material and help in designing the study. This work was supported by an NWO Spinoza Prize awarded to PH and a PhD grant from the Max Planck Society to RK.

Chapter 6

An independent psychometric evaluation of the PROMS measure of music perception skills

The Profile of Music Perception Skills (PROMS) is a recently developed measure of perceptual music skills which has been shown to have promising psychometric properties. In this paper we extend the evaluation of its brief version to three kinds of validity using an individual difference approach. The brief PROMS displays good discriminant validity with working memory, given that it does not correlate with backward digit span (r = .04). Moreover, it shows promising criterion validity (association with musical training (r = .45), musicianship status (r = .48), and self-rated musical talent (r = .51)). Finally, its convergent validity, i.e. relation to an unrelated measure of music perception skills, was assessed by correlating the brief PROMS to harmonic closure judgment accuracy. Two independent samples point to good convergent validity of the brief PROMS (r = .36; r = .40). The same association is still significant in one of the samples when including self-reported music skill in a partial correlation (r_{partial} = .30; r_{partial} = .17). Overall, the results show that the brief version of the PROMS displays a very good pattern of construct validity. Especially its tuning subtest stands out as a valuable part for music skill evaluations in Western samples. We conclude by briefly discussing the choice faced by music cognition researchers between different musical aptitude measures of which the brief PROMS is a well evaluated example.

Kunert, R., Roel M. Willems, Peter Hagoort (2016). An independent psychometric evaluation of the PROMS measure of music perception skills. *PLoS ONE, 11* (7), e0159103. doi:10.1371/journal.pone.0159103.

6|1 Introduction

Perceptual music skills differ widely in the population: from amusic individuals who exhibit impaired music listening skills (Peretz, Champod, & Hyde, 2003) to highly proficient people scoring highly on musical skill measures (Law & Zentner, 2012). There is growing interest in these inter-individual differences, partly because evidence is accumulating that musical and non-musical faculties are related. For example, music skills have been linked to native and non-native language abilities (Anvari, Trainor, Woodside, & Levy, 2002; Slevc & Miyake, 2006).

However, progress in music cognition has been hampered by an absence of modern, objective measurement tools which are both fast as well as easy to administer and psychometrically validated. A variety of novel musical skill measures has been proposed to fill this gap (Gingras, Honing, Peretz, Trainor, & Fisher, 2015; Peretz et al., 2013; Schaal, Bauer, & Müllensiefen, 2014; Ullén, Mosing, Holm, Eriksson, & Madison, 2014; Wallentin, Nielsen, Friis-Olivarius, Vuust, & Vuust, 2010). This publication is concerned with the psychometric evaluation of one measure for perceptual music skills (the profile of music perception skills; PROMS (Law & Zentner, 2012); www.zentnerlab.com/psychological-tests/the-profile-of-music-perception-skills) which has already been adopted by researchers interested in music cognition (Pasinski, Hannon, & Snyder, 2016).

As most other music skill measures, the PROMS requires participants to judge whether a reference and a probe stimulus are the same or not. The comparison can be performed based on different music features such as melody, rhythm, or tuning. While the full PROMS is based on nine subtests assessing a different music feature each, the brief version - which we focus on here - comprises only four (comparison of stimuli based on melody, tuning, tempo, or rhythmic accent) taking about half an hour to administer. The administration time is somewhat higher than that of other novel musical skill measures aimed at adults (11 minutes for SMDT (Ullén et al., 2014); 18 minutes for the MET (Wallentin et al., 2010); 20 - 25 minutes for the Gold-MSI (Müllensiefen, Gingras, Musil, & Stewart, 2014)). However, in return, the brief PROMS provides more subtests (4 for PROMS; 2 for MET; 3 for SMDT; Gold-MSI includes 2 music measures according to (2014) and 4 according to the website http://www.gold.ac.uk/music-mind-brain/gold-msi/) allowing for a more fine-grained assessment of music skills and subskills.

This raises the question of what the PROMS actually purports to assess. We define the measured concept - music perception skills - along the same lines as Law & Zentner (2012). The focus lies on rather elementary aspects of music which can be found across musical systems and traditions, such as the use of discrete pitch, tempo, precise rhythms and melodic lines (Brown & Jordania, 2013). While this does not render the PROMS culture-free, it allows for insights into musicality which are potentially wider ranging than the Western cultural context. We are uncommitted as to the origin of musical skills, whether they are due to deliberate practice (Ericsson, Krampe, & Tesch-Römer, 1993) or talent/giftedness (Macnamara, Hambrick, & Oswald, 2014). The advantage of musical skill measures such as the PROMS is that they can be used to answer questions regarding the origins of musical skills instead of relying on skill operationalisations (e.g., musical instrument proficiency) which require musical expertise. Thus, the PROMS aims to measure basic music abilities in the general population, including musically trained and untrained individuals.

Whether the PROMS achieves its aim of measuring music perception skills has been assessed psychometrically. Law and Zentner (2012) reported high internal consistency as well as good test-retest reliability for the brief PROMS. Furthermore, the full PROMS, with which the brief version correlates at $r = .95$, has been shown to have convergent validity with other measures of music ability, i.e. it appears to measure the same concept as established music ability tests. Also, it exhibits criterion validity with various measures of music achievement, i.e. it is related to ‘real world’ variables of musical skill such as musical training and musicianship status. Furthermore, Law and Zentner (2012) investigated its discriminant validity, i.e. whether the PROMS is *not* associated with a task measuring an unrelated concept. The unrelated task was a gap detection task in which participants had to detect short gaps of silence in white noise. Given that none of the correlations between sub-test scores and gap-detection performance reaches significance, performance on the PROMS cannot be equated with nonmusical auditory discrimination abilities.

While these results impressively demonstrate the good test properties of the PROMS, open questions remain. First of all, Law and Zentner (2012) established discriminant validity solely with a gap-detection task even though it is known that similar music skill measures suffer from a working memory confound. That is to say that holding a reference stimulus in mind in order to compare it to a probe stimulus requires working memory resources whose efficiency and size can consequently influence test scores, as seen for similar music ability measures (e.g., Anvari et al., 2002; Hansen, Wallentin, & Vuust, 2013; Wallentin et al., 2010). Therefore, we evaluate the relation of the brief PROMS to a standard measure of working memory: digit span. Its forward subscore is usually thought to measure short-term memory (store information temporarily) while its backward subscore is related to working memory (holding information online for active processing).

Secondly, we sought to replicate and extend Law and Zentner's (2012) assessment of criterion validity by comparing brief PROMS scores to values measuring musical training (years of training, musicianship status) as before. However, we also include self-reported musical talent in this assessment in order to evaluate whether musical skill as measured by the PROMS is related to musical talent as commonly understood in the general population.

Thirdly, convergent validity of the PROMS has so far only been established through a comparison with other music ability measures which are, crucially, also based on the judged similarity of a reference and a comparison stimulus. Therefore, we investigate the test's convergent validity by comparing it to a different kind of task. This new task is based on a music feature (harmony) which is not directly measured by any of the subtests of the brief PROMS. Furthermore, the task (closure ratings) is different in kind to the usual music ability measures as it does not rely on an auditory discrimination between two stimuli. Therefore, convergent validity aims to see whether the assessment of perceptive music skills in one task (PROMS) is related to the same assessment in a completely different task (harmony judgments).

A music ability measure with good psychometric properties is essential in order to investigate the influence of music skills on cognition as well as the underlying reason for why some people are 'good at music'. Therefore, it is important to rigorously assess newly developed measures of music ability before they are widely adopted. The current investigation does just that for the brief PROMS.

6|2 Methods

6|2.1 Ethics Statement

Written informed consent was obtained from all participants prior to measurement and the study received ethical approval from the local reviewing committee ‘‘CMO Arnhem Nijmegen’’ (CMO no 2001/095 and amendment ‘‘Imaging Human Cognition’’ 2006, 2008), in accordance with the Research involving human subjects Act, following the principles of the Declaration of Helsinki.

6|2.2 Participants

Our data are based on main task, pretest, post-test, and pilot participant data principally acquired for an independent research question (Kunert et al., 2016). The full sample consists of 161 Dutch participants (37 males) aged 18 to 64 (*M* = 22.80; *SD* = 4.75), with little formal musical training (*M* = 4.39 years; min = 0; max = 20; *SD* = 4.46). Most participants (54%) described themselves as non-musicians, 39% as amateur and 7% as semi-professional musicians. 14% reported being left-handed. All were paid for their participation or received undergraduate course credit. Not all participants took part in all aspects of the study. Therefore, the number of people available for specific analyses differs between 53 and 160 (see sample sizes per analysis in brackets below).

6|2.3 Tasks

The brief PROMS was administered in the lab as a web-based test at the end of a testing session. The brief PROMS assesses melody first, followed by tuning, tempo, and rhythmic accent. Participants are asked to judge whether a standard stimulus, which is repeated, is identical to a comparison stimulus. Answers are given on a five-point Likert scale providing a coarse measure of confidence ("definitely same", "probably same", "I don’t know", "probably different", and "definitely different"). Each subtest includes 18 trials.

In the melody subtest participants hear a two-bar monophonic harpsichord melody twice, followed by the probe melody which can differ slightly by one or more tones. The tuning subtest plays a C-chord whose tone E could be mistuned. Participants are asked to judge whether the tuning is the same in the reference and the probe stimulus. The tempo subtest comprises rhythmically and timbrally diverse stimuli which are the same between reference and probe stimulus except, potentially, for their tempo. Finally, the rhythmic accent subtest uses non-melodic rhythmic sequences played with a rim-shot timbre. Participants are asked to detect whether intensity accents are placed on the same notes in the reference stimulus and the probe stimulus.

In order to measure working memory, digit span was acquired in Dutch. The task contains only single digits which are read out by the experimenter (RK) at a rate of approximately one per second. Participants were tested individually. The experimenter asked them to repeat progressively longer sequences of digits in the same order (forward digit span) or in reverse order (backward digit span) until they failed on two trials of the same length (Groth-Marnat, 2001).

The closure rating task, which we use in order to measure musical task accuracy, is based on ten harmonic sequences made up of 14 chords played with a piano timbre at 96 bpm; see Figure 6|1. Participants are asked to judge their feeling of completeness, i.e. to what extent they feel the music stimulus has ended instead of being cut early (seven point Likert scale). Unbeknownst to the participants, sequences end either on an authentic cadence (dominant followed by tonic) or not (dominant followed by supertonic or subdominant). The former usually result in a high closure rating, the latter in a low closure rating. Original sequences are transposed twice resulting in 60 harmonic sequences (10 items × 2 endings × 3 transpositions).

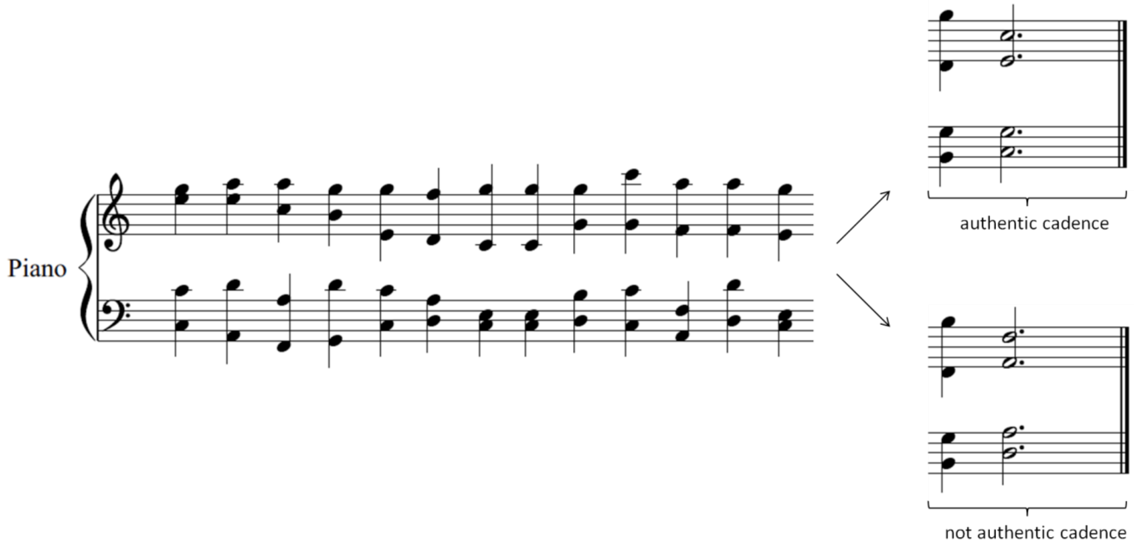



Figure 6|1. Musical task item. Participants are required to rate the closure (feeling of completion) of chord progressions ending either on an authentic cadence (top ending) or not (bottom ending), i.e. a dominant followed by a supertonic (shown here) or followed by a subdominant (not shown). Accuracy refers to the average rating of sequences ending on an authentic cadence minus the average rating of no cadence endings.

60 participants were required to perform the music task in isolation (post-test participants in (Kunert et al., 2016)). Their data are based on all 40 trials (20 ending on authentic cadence, 20 not) which are entered into the analysis under the label ‘full attention’. Given the moderate sample size, we sought to replicate the findings of this sample with a new set of 56 other participants performing the music task while simultaneously solving a reading task (one word per chord presented visually) or an arithmetic task (one number or operator per chord presented visually) (experiment 1 participants in (Kunert et al., 2016)). The trials analyzed here constituted the filler trials in the original study, i.e. their language/arithmetic dimension was variable but relatively easy. All 100 filler trials (50 ending on authentic cadence, 50 not) entered into the analysis labeled as ‘divided attention’.

Our musical task accuracy measure is novel, requiring some basic validation. Regarding its internal consistency, Cronbach’s α is .90 overall (117 participants, 10 items). This suggests good to excellent internal consistency. Regarding its external validity, our musical task accuracy measure does correlate with the number of years of formal musical training ($r_{(117)} = .21$, $p = .02$). This suggests that it is related to at least one real-world measure of musical skill.

6|2.4 Analysis

Raw data and analysis code written in R are available in the supplementary materials. The brief PROMS scores were derived from the web-based feedback screen. Per trial, participants receive 1 point for a correct response chosen with maximum confidence, half a point for a correct response chosen with less confidence, and zero points for an incorrect or ‘I don’t know’ response. The maximum possible score, therefore is 18 points per subtest and 72 points overall. For forward and backward digit span we used the number of correct trials. For the closure rating task we derived a difference score from the ratings (authentic cadence minus no-cadence).

In order to check whether the correlations we report are robust, we compare the reported Pearson correlation values to Spearman rank order correlations and to iterated re-weighted least squares regressions which weigh down data points with large residuals. The results are very similar for all three methods.

We adjust the alpha-level of the correlations’ inferential tests using the Bonferroni correction in order to control for the number of erroneous theoretical inferences. Each kind of validity is taken as an independent theoretical claim following the intuition that the assessed musical skill measure might well be valid on one dimension (e.g., discriminant validity) but not another (e.g., criterion validity). All tables and Figure 6|2 report uncorrected *p*-values while the text reports both uncorrected and corrected *p*-values. The latter correct for three (criterion validity), or two (discriminant and convergent validity) comparisons.

6|3 RESULTS

6|3.1 Test structure

In Table 6|1 we show the pattern of associations of the overall brief PROMS score and its subtests with each other. Each of the subtests correlates very highly with the overall brief PROMS score (all *r*s > .77). Amongst each other, the subtests correlate between *r* = .46 and *r* = .62, i.e. the effect size of the association is mostly large (*r* > .5) according to Cohen's criteria (Cohen, 1992).

Table 6|1. Pearson product moment correlations of the brief PROMS and its subtests with each other. Sample size is given in brackets.

	melody	tuning	tempo	rhythmic accent
brief PROMS total	$r_{(157)}$ = .812***	$r_{(157)}$ = .833***	$r_{(157)}$ = .776***	$r_{(157)}$ = .841***
melody		$r_{(160)}$ = .560***	$r_{(157)}$ = .458***	$r_{(157)}$ = .619***
tuning			$r_{(157)}$ = .559***	$r_{(157)}$ = .580***
tempo				$r_{(157)}$ = .552***

Note. The maximal absolute difference between the Pearson *r* values reported here and their associated Spearman *rho* values is .038 units. † .05 < $p_{\text{uncorrected}}$ < .1 ; *$p_{\text{uncorrected}}$ < .05; **$p_{\text{uncorrected}}$ < .01; ***$p_{\text{uncorrected}}$ < .001

6|3.2 Validity measures

Table 6|2 shows how the different validity measures correlate with each other. This assessment is of course only possible if the same participants provide information for different measures. This is not the case for all combinations of validity measures, see Table 6|2 (*N* = 0).

The two discriminant validity measures (forward and backward digit span) are merely moderately correlated (r = .33), justifying a distinction between short term memory (forward digit span) and working memory (backward digit span). The criterion validity measures (musical training years, musicianship status, self-rated musical talent), on the other hand, are strongly correlated with each other (rs > .64).

Furthermore, it is noteworthy that the correlations of musical task accuracy under *divided* attention with musical skill measures are somewhat greater (rs > .36) than the correlations of musical task accuracy under *full* attention with musical skill measures (rs < .24). This could reflect the different trial counts available for the two different musical skill measures. A measure with more observations (100 trials under divided attention) is probably less noisy and can therefore display a higher correlation with another variable than a measure with less observations (40 trials under full attention). Alternatively, musical training might affect attention abilities which in turn impact harmony perception. According to this speculative account, the influence of musical training on harmony perception is increased when attention is taxed because attention is differently developed for participants with different amounts of musical training.

Table 6|2. Associations among the validity measures. Sample size is given in brackets.

	backward digit span	musical training years	musicianship status	self-rated musical talent	musical task accuracy (full attention)	musical task accuracy (divided attention)
forward digit span	$r_{(101)}$ = .326**	$r_{(101)}$ = .103	$r_{(100)}$ = .078	$r_{(101)}$ = .135	*N* = 0	$r_{(57)}$ = .148
backward digit span		$r_{(101)}$ = .016	$r_{(100)}$ = -.096	$r_{(101)}$ = .003	*N* = 0	$r_{(57)}$ = .252†
musical training years			$r_{(160)}$ = .647***	$r_{(161)}$ = .644***	$r_{(60)}$ = .156	$r_{(57)}$ = .366**
musicianship status				$r_{(160)}$ = .806***	$r_{(60)}$ = .212	$r_{(57)}$ = .419**
self-rated musical talent					$r_{(60)}$ = .230†	$r_{(57)}$ = .420**
musical task accuracy (full attention)						*N* = 0

Note. The maximal absolute difference between the Pearson *r* values reported here and their associated Spearman *rho* values is .082 units. † .05 < $p_{\text{uncorrected}}$ < .1 ; *$p_{\text{uncorrected}}$ < .05; **$p_{\text{uncorrected}}$ < .01; ***$p_{\text{uncorrected}}$ < .001

Table 6|3. Pearson product moment correlations of the brief PROMS and its subtests with measures of validity. Sample size is given in brackets.

	brief PROMS total	melody	tuning	tempo	rhythmic accent
forward digit span	$r_{(97)} = .223$*	$r_{(100)} = .238$*	$r_{(100)} = .053$	$r_{(97)} = .104$	$r_{(97)} = .285$**
backward digit span	$r_{(97)} = .039$	$r_{(100)} = .100$	$r_{(100)} = -\ .082$	$r_{(97)} = .005$	$r_{(97)} = .092$
musical training years	$r_{(157)} = .450$***	$r_{(160)} = .482$***	$r_{(160)} = .416$***	$r_{(157)} = .201$*	$r_{(157)} = .357$***
musicianship status	$r_{(156)} = .475$***	$r_{(159)} = .456$***	$r_{(159)} = .446$***	$r_{(156)} = .289$***	$r_{(156)} = .354$***
self-rated musical talent	$r_{(157)} = .513$***	$r_{(160)} = .527$***	$r_{(160)} = .404$***	$r_{(157)} = .372$***	$r_{(157)} = .366$***
musical task accuracy (full attention)	$r_{(60)} = .364$**	$r_{(60)} = .340$**	$r_{(60)} = .425$**	$r_{(60)} = .094$	$r_{(60)} = .315$*
musical task accuracy (divided attention)	$r_{(53)} = .398$**	$r_{(56)} = .347$**	$r_{(56)} = .421$**	$r_{(53)} = .274$*	$r_{(53)} = .288$*

Note. The maximal absolute difference between the Pearson *r* values reported here and their associated Spearman *rho* values is .065 units. † $.05 < p_{\text{uncorrected}} < .1$; *$p_{\text{uncorrected}} < .05$; **$p_{\text{uncorrected}} < .01$; ***$p_{\text{uncorrected}} < .001$

6|3.3 Discriminant validity

Discriminant validity assesses whether tests measuring unrelated concepts do not correlate with each other. As can be seen in Table 6|3 and Figure 6|2A, the brief PROMS does not correlate with backward digit span (r = .04, $p_{\text{uncorrected}}$ = .71, $p_{\text{corrected}}$ > 1), neither do any PROMS subtests individually (rs ≤ .100). Even though the correlation of the PROMS total score with forward digit span reaches significance (r = .22, $p_{\text{uncorrected}}$ = .03, $p_{\text{corrected}}$ = .06, see Figure 6|2B), the association is weaker than for other music ability tests. These exhibit Pearson correlation coefficient values of $r \geq .4$ (Anvari et al., 2002; Wallentin et al., 2010), i.e. beyond the 95% Confidence Interval of the association we find here (95% CI = .02 - .40). When looking at the four sub-tests separately (see Table 6|3), it becomes clear that the tuning (r = .05) and tempo subtests (r = .10) prevent the overall brief PROMS score from correlating strongly with forward digit span. Overall, the brief PROMS shows surprisingly good discriminant validity with short term memory and working memory.

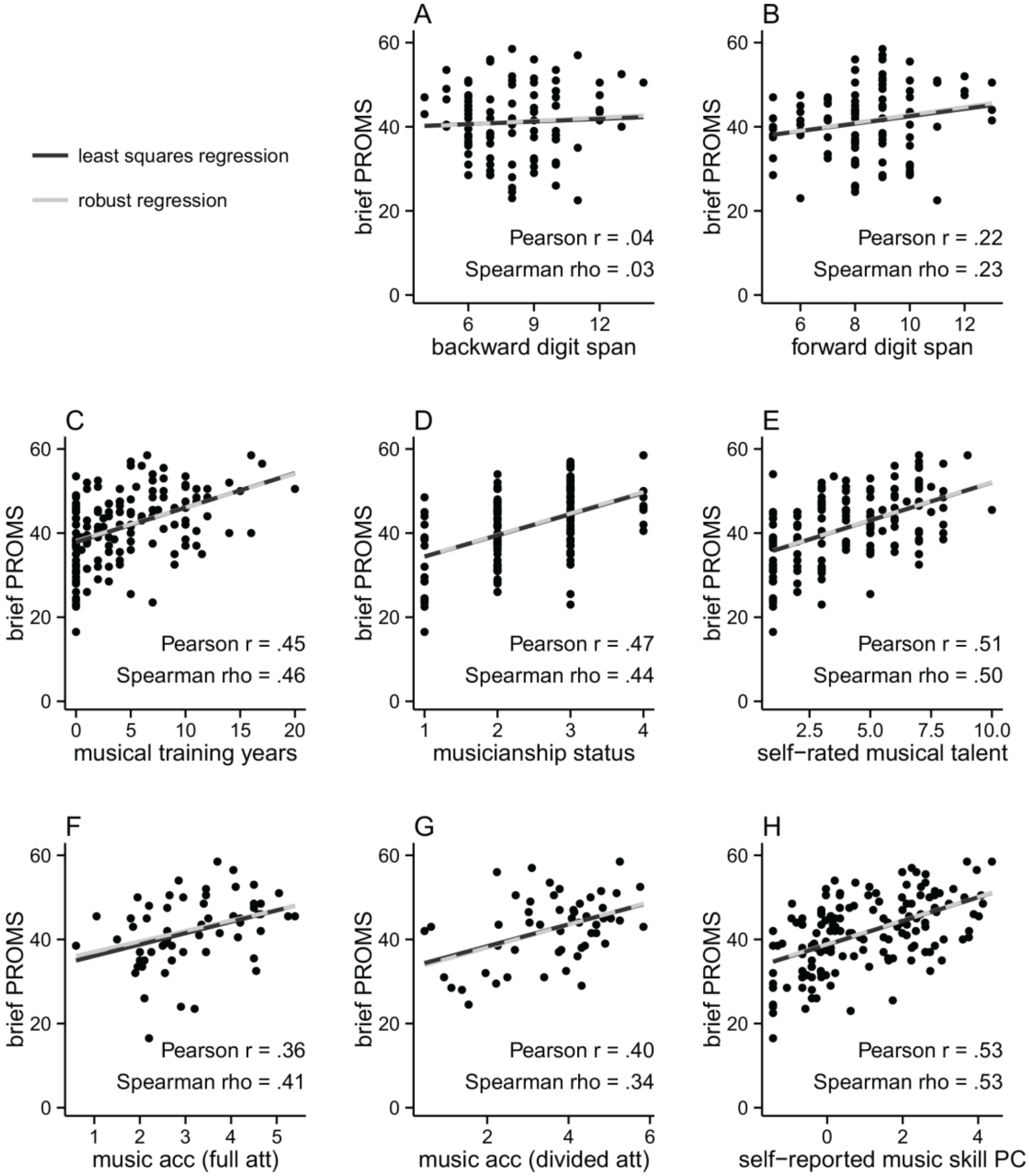


Figure 6|2. Correlation of brief PROMS total scores with validity measures. Two lines are fitted to the data, a linear fit (dark) corresponding to Pearson *r* and a robust fit (light) corresponding to an iterated re-weighted least squares regression. Overlapping lines are plotted as a dashed dark-light line. For inferential tests and PROMS subtest scores, see Table 6|3. music acc (full att) = musical task accuracy (full attention); music acc (divided att) = musical task accuracy (divided attention); self-reported music skill PC = first principal component combining musical training years, musicianship status, and self-rated musical talent.

6|3.4 Criterion validity

Criterion validity assesses a test's association with concrete criteria outside the lab, e.g. teacher assessments or attainments in music exams. We have no access to such data and therefore use three proxies instead. Following Law & Zentner (2012), these are years of musical training and musicianship status (coded: nonmusician [= 1]; music-loving nonmusician [= 2]; amateur musician [= 3]; semiprofessional musician [= 4]; professional musician [= 5]). The third measure is self-rated musical talent (10-point Likert scale; not at all talented [= 0] to extremely talented [= 10]).

Just like Law and Zentner (2012), we find a correlation between the brief PROMS and years of formal musical training ($r = .45$, $p_{\text{uncorrected}} < .001$, $p_{\text{corrected}} < .001$) as well as between the brief PROMS and musicianship status ($r = .47$, $p_{\text{uncorrected}} < .001$, $p_{\text{corrected}} < .001$), see Table 6|3, Figure 6|2C, and Figure 6|2D. The observed association strength is remarkably similar to that reported by Law and Zentner (2012) for the correlation between the brief PROMS and years of formal musical training: $r_{(39)} = .39$. Going beyond associations with measures of musical training, we investigate the relation to self-rated musical talent, see Figure 6|2E. The moderate association strength ($r = .51$, $p_{\text{uncorrected}} < .001$, $p_{\text{corrected}} < .001$) suggests that the brief PROMS partly measures musical talent as understood by relatively untrained participants. All four subtests show association strengths which are not much lower (rs > .36). Overall, this analysis confirms the comparatively good criterion validity of the brief PROMS.

6|3.5 Convergent validity

If two tasks which measure the same construct correlate, they exhibit convergent validity. The brief PROMS correlates well with our measure of musical task accuracy which is based on harmonic closure ratings either acquired in a full attention setting ($r = .36$, $p_{\text{uncorrected}} = .004$, $p_{\text{corrected}} = .009$) or in a divided attention setting ($r = .40$, $p_{\text{uncorrected}} = .003$, $p_{\text{corrected}} = .006$), see Table 6|3, Figure 6|2F, and Figure 6|2G.

In order to check whether the correlation between musical task accuracy and the brief PROMS is due to a third variable influencing both closure ratings and the brief PROMS, we include the self-reported musical skill measures in a partial correlation analysis. Given that musical training years, musicianship status, and self-rated musical talent are highly correlated (see Table 6|2), we summarize them in a single principal component which in this case accounts for 80% of the variance and is correlated with the brief PROMS ($r = .53$, $p < .001$, see Fig 6|2H). In the case of musical task accuracy under full attention, the correlation with the brief PROMS score is still significant when self-reported musical skill measures are held constant through a partial correlation ($r_{\text{partial}} = .30$, $p_{\text{uncorrected}} = .019$, $p_{\text{corrected}} = .038$). However, the correlation between musical task accuracy under *divided* attention and the brief PROMS score is no longer significant after holding self-reported musical skill measures constant ($r_{\text{partial}} = .17$, $p_{\text{uncorrected}} = .232$, $p_{\text{corrected}} = .465$). Overall, these findings suggest that the brief PROMS truly measures musical skills rather than just the ability to compare two auditory stimuli. However, whether this result is a reflection of musical training, talent and musicianship status influencing both the brief PROMS and music harmony perception skills is ambiguous.

6|3.6 Construct validity

At the suggestion of an anonymous reviewer, we also include a validity measure combining all previously reported measures of validity. Westen and Rosenthal (2003) propose $r_{\text{alerting-CV}}$ and $r_{\text{contrast-CV}}$ as measures of general construct validity. Given that not all participants contributed to all validity measures, only $r_{\text{alerting-CV}}$ can be calculated here. It is a measure of fit between the predictions of discriminant, criterion, and convergent validity, and the observed values. It is not associated with a *p*-value.

The ideal pattern of correlations between validity measures and the brief PROMS is shown in Table 6|4's first column. We assume that discriminant validity measures should not be correlated with the brief PROMS ($r = 0$) while criterion and convergent validity measures should correlate as highly as possible (given the test-retest reliability of the brief PROMS at $r = .84$). The resulting $r_{\text{alerting-CV}}$ of .90 (the correlation between the predicted values in column 2 of Table 6|4 and the observed values in column 4) points to an overall very good fit between the ideal pattern of construct validity and the observed pattern. None of the subtests alone displays a poor fit either: melody ($r_{\text{alerting-CV}} = .84$), tuning ($r_{\text{alerting-CV}} = .98$), tempo ($r_{\text{alerting-CV}} = .71$), and rhythmic accent ($r_{\text{alerting-CV}} = .76$).

Table 6|4. Values required for the calculation of the construct validity measure $r_{\text{alerting-CV}}$.

	predicted correlation (r) with brief PROMS	Fischer Z of predicted r (demeaned)	observed r	Fischer Z of observed r
forward digit span	$r = 0$	$Z_{r\_demeaned} = -0.872$	$r = .223$	$Z_r = 0.227$
backward digit span	$r = 0$	$Z_{r\_demeaned} = -0.872$	$r = .039$	$Z_r = 0.039$
musical training years	$r = .84$	$Z_{r\_demeaned} = 0.349$	$r = .450$	$Z_r = 0.485$
musicianship status	$r = .84$	$Z_{r\_demeaned} = 0.349$	$r = .475$	$Z_r = 0.516$
self-rated musical talent	$r = .84$	$Z_{r\_demeaned} = 0.349$	$r = .513$	$Z_r = 0.567$
musical task accuracy (full attention)	$r = .84$	$Z_{r\_demeaned} = 0.349$	$r = .364$	$Z_r = 0.381$
musical task accuracy (divided attention)	$r = .84$	$Z_{r\_demeaned} = 0.349$	$r = .398$	$Z_r = 0.421$

Note. Predicted correlations between validity measures and the brief PROMS are zero for discriminant validity measures. They are the same as the test-retest validity of the brief PROMS for the criterion and convergent validity measures, following the intuition that no measure can better predict brief PROMS scores than the brief PROMS itself

6|4 Discussion

There is no agreement on how to measure musical aptitude even though it provides interesting avenues for research. Newly developed measures of musical skills should be rigorously, psychometrically assessed before being widely applied, as we do here for the brief PROMS (Law & Zentner, 2012). For this evaluation we focus on various measures of validity and show that on all of them the chosen measure of musical perception skills performs well. In terms of discriminant validity, its correlations with short term memory and working memory are low despite the nature of the task which asks participants to hold a stimulus in mind in order to compare it to a second stimulus. Perhaps unsurprisingly, the PROMS subtests with longer stimuli (melody and rhythmic accent) display somewhat stronger (but still weak) correlations with short term memory than the subtests with shorter stimuli (tuning, tempo).

As opposed to the weak correlations with short term memory and working memory, the associations with musical training, musicianship status, and self-rated musical talent are high. This suggests good criterion validity, meaning that the brief PROMS is associated with ‘real world’ measures of the same concept. The observed correlations might not be even higher because of the presence of musical sleepers (musically untrained people with great music perception skills, see Figure 6|2C or Figure 6|2H top left corner) and sleeping musicians (musically trained people with surprisingly poor music perception skills, see Figure 6|2C or Figure 6|2H bottom right corner) (Law & Zentner, 2012).

Furthermore, in two independent samples, brief PROMS scores correlate well with a different kind of music measure based on closure ratings of harmonic sequences. We take this as a sign of good convergent validity - the brief PROMS measures an underlying concept (musical perception skill) also measured by the closure rating task. This result is surprising because no PROMS subtest actually requires any harmonic understanding, suggesting that the brief PROMS captures a general form of musical aptitude which generalizes to unrelated tasks, as hypothesized by Law and Zentner (2012). The pattern of correlations of the subtests suggests that next to general musical aptitude there are also music perception subskills which can be more or less developed in the same person. Specifically, the harmonic judgment accuracy measure correlates well with the only subtest including chords (tuning subtest) but weakly with timing-related subtests (tempo, rhythmic accent). This suggests that the brief PROMS can, at least to some degree, measure a profile of strengths and weaknesses in music perception skills. It remains ambiguous whether this pattern of association between the brief PROMS and the harmony perception task is due to both these tasks being affected by musical training or whether the association between the brief PROMS and harmony perception goes beyond musical training. A partial correlation analysis reveals contradictory findings for two independent samples. Future research might elucidate this point.

Finally, an overall construct validity analysis (Westen & Rosenthal, 2003), which combines all aforementioned patterns of association between the brief PROMS and validity measures, suggests that the brief PROMS conforms remarkably well to an ideal pattern of discriminant, criterion and convergent validity. None of the individual subtests performs poorly either. One can conclude that the brief PROMS shows good overall construct validity.

One of the PROMS subtests, the tuning subtest, which asks participants whether two chords are tuned in the same way, is noteworthy. It shows a remarkably good pattern of discriminant, convergent, criterion, and overall construct validity. It is not associated with working memory and still correlates significantly with musical training years, musicianship status, self-rated musical talent and musical task accuracy. This subtest's unusual performance might not just result from the short stimuli. One could speculate that by asking for a tuning judgment, participants could compare each chord to a tuning standard held in long term memory. As a result, they do not really compare the standard with the comparison stimulus, but instead simply classify each as well tuned or mistuned. With this strategy a matching classification (e.g., both chords well tuned) suggests matching standard and comparison stimuli. If this hypothesis is true, the tuning subtest depends on a culturally specific representation of tuning held in long term memory and shared among listeners of Western music. This suggests that such a subtest is a valuable part of any evaluation of music ability in Western listeners but perhaps not in non-Western listeners.

Does the promising outcome of the psychometric evaluation of the brief PROMS which we present here mean music cognition researchers should adopt this measure? We believe that the answer depends on the research question. Different musical aptitude measures offer the interested researcher different advantages. We have shown that one advantage of the brief PROMS is its very good pattern of discriminant, convergent and criterion validity. However, other considerations could also play a role. For example, some measures of musical skills are claimed to be better suited for younger (Peretz et al., 2013) or musically impaired samples (Peretz et al., 2003). While others require shorter testing times of less than 20 minutes (Ullén et al., 2014; Wallentin et al., 2010). Yet others claim to measure a somewhat broader concept of musical sophistication which goes beyond music skills trained in formal instrument lessons (Müllensiefen et al., 2014; Schaal et al., 2014). Moreover, if the aim is to compare any results to previous studies using classical musical aptitude tests, the interested researcher is probably well advised to opt for these classical tests instead (Gordon, 1989; Seashore, Lewis, & Saetveit, 1960). We hope that the results presented here help in determining which music ability measure to choose. We believe that the brief PROMS should be on the list of musical skill tests to consider.

Chapter 7

General discussion and conclusions

7|1 Summary

The research presented in this thesis evaluated the shared syntactic integration resource hypothesis (Patel, 2003). This model proposes shared music-language resources located in frontal brain areas, probably in the inferior frontal gyrus according to the MUC model which has a very similar component (Hagoort, 2005, 2013). These resources are involved in syntactic integration, specifically the activation of structurally unexpected items stored in memory. Memory storage is thought to be domain-specific and these resources are hypothesizes to lie in posterior brain regions, probably the middle temporal gyrus in the case of language according to the MUC model which has a very similar component.

Chapter 2 presented two experiments evaluating the model prediction that easy or difficult structural processing of visually presented sentences should influence concurrent music processing. This was indeed the case: participants found it difficult to hold a harmonic key online when reading a syntactically unexpected word, compared to an expected word. Moreover, these experiments ruled out a number of alternative accounts of interference effects between music and language. First, none of the stimuli contained errors, suggesting that shared error processing is not at the heart of music-language interactions (Rogalsky et al., 2011). Second, given that an arithmetic difficulty manipulation or a language semantics manipulation did not influence music perception, a general attention mechanism is also unlikely to cause music-language interference in this paradigm. Third, the interference effect of language syntax on music harmony perception was not influenced by participants' working memory resources, suggesting that verbal working memory is not the cognitive mechanism mediating shared music and language processing. By focusing on music behaviour, chapter 2 also, for the first time, elucidated the role of shared music-language resources in the music network. Rather than being involved in combining chords to infer a harmonic key, shared resources are involved in holding an existing key online as an integration site.

Having established that shared music-language resources have behavioural consequences, and, thus, that they must in some way be represented in the brain, chapters 3 and 4 investigated the neural basis of shared music-language resources with fMRI and MEG respectively. Chapter 3 evaluated the prediction that shared music-language resources reside in frontal brain areas. Indeed, a region of interest analysis revealed that Broca's area displays a typical interference effect of music on language syntax processing. Given that Broca's area is associated with structural integration in many domains, this supports the idea that shared music-language resources are indeed involved in structural integration. However, elucidating the temporal dynamics of these shared resources failed in chapter 4, likely because of methodological short comings. Future research will have to test the time-course of music-language interactions.

Chapter 5 switched the focus from shared resources to non-shared resources hypothesized by the SSIRH (Patel, 2003). As predicted, a temporal brain region hypothesized to store language specific representations is not involved in music processing. However, finding a similar music-specific memory storage area of harmonic elements remains work to be done.

Chapter 6 offers a first glimpse for how to extend the SSIRH in the future. It presents a psychometric evaluation of a musical aptitude measure and concludes that the evaluated profile of music perception skills (PROMS; Law & Zentner, 2012) has excellent discriminant, criterion, convergent, and overall construct validity. Should researchers in the future want to investigate how inter-individual differences in musical aptitude relate to shared music-language resources, then the PROMS offers a good way to measure music skills. Such research could build on the finding that language syntax processing is altered according to individual differences in musical training (Fitzroy & Sanders, 2013; Jentschke & Koelsch, 2009), and that these musical training differences in turn relate to brain structural differences in Broca's area, the supposed location of shared music-language resources (Bermudez et al., 2009; Gaser & Schlaug, 2003; James et al., 2014; Sluming et al., 2002).

However, the careful reader will have noticed some apparent contradictions between chapters. Chapters 2, 3, and 4 all employed an interference paradigm in which music harmony processing is supposed to influence language behaviour (chapters 3 and 4) or vice versa language processing is supposed to influence music behaviour (chapter 2). Do the behavioural null effects observed in chapters 3 and 4 reduce the credibility of the behavioural music-language interference effects reported in chapter 2? I do not believe so for two reasons. First, the primary aim of the behavioural measures in chapters 3 and 4 was to induce optimal neural activity at the cost of a sensitive behavioural measure. This was not the case with chapter 2 which attempted to optimize the behavioural measure *per se*. It is not terribly surprising that the optimized behavioural measure of chapter 2 proved more sensitive to revealing shared music-language resources than the non-optimized one of chapters 3 and 4. Second, in retrospect, I believe that the design of chapter 2 is superior to that of chapters 3 and 4 because only in chapter 2 the point of syntactic integration difficulty occurs at the optimal time in order to interact with music processing.

Nonetheless, one might wonder why chapter 3 observed shared music-language resources in Broca's area while chapters 4 and 5 did not. I believe that the reasons are specific to each chapter. Chapter 3 might have *benefitted* from the poor temporal resolution of fMRI compared to the MEG measurements in chapter 4. This way, variable moments of processing the structural integration difficulty of object-extracted relative clause sentences are still summed in the fMRI BOLD signal, but not in the MEG signal. Given how difficult the manipulation was for Dutch native speakers, it is not unreasonable to think that they sometimes noticed the object-extracted relative clause 'a bit late'. Moreover, the absence of fillers in chapter 3 improved behavioural performance compared to chapter 4 which included fillers. This renders stimuli more predictable in chapter 3 and perhaps facilitates appropriate processing of difficult sentences which are otherwise beyond the capability of the participant who did not see them coming.

This still does not explain the failure of observing shared music-language resources in frontal areas in chapter 5. A closer look reveals that the harmonic manipulation in chapter 3 was very different (in-key versus out-of-key tone) to the harmonic manipulation in chapter 5 (structural repetition of harmony or no repetition). The original article proposing the SSIRH (Patel, 2003) actually only predicts the first kind of manipulation to have an effect. The second kind of manipulation is based on an extension of the functional role of the harmonic integration mechanism of the SSIRH. Future targeted investigations of what exactly the shared syntactic integration resources do in the music network might elucidate why chapter 5 did not reveal these resources the way chapter 3 did.

In summary, despite the partly contradictory findings, the SSIRH generally receives support. Of course, this evaluation is subject to independent replication in order to establish the reliability of the findings (Kunert, 2016; Open Science Collaboration, 2015). Until then, I am inclined to conclude that the SSIRH and the related MUC model offer a scientifically useful description for how music and language comprehension relate to each other in the brain.

7|2 The future of the shared syntactic integration resource hypothesis

I believe that future investigations of the relation between music and language processing will benefit both from novel paradigms to investigate the SSIRH in its current form, old paradigms to investigate a possibly updated SSIRH, and progress in the computational modeling of music and language comprehension. What novel paradigms are available to test the SSIRH in its current form? First, the speech-to-song illusion in which a speech stimulus is interpreted as music after repetition, lends itself well to investigating the SSIRH (Deutsch et al., 2011; Falk et al., 2014; Vanden Bosch der Nederlanden et al., 2015). Presumably, the melodic part of the speech stimulus at first gets interpreted as prosody, i.e. without reference to music-harmonic rules. However, after repetition these rules are used to interpret an increasingly more musical percept. Does this mean that the harmonic properties increasingly interact with the syntactic properties of the stimulus? If so, the speech-to-song illusion would offer a powerful means to establish that music and language processing are not dependent on acoustic input characteristics (as the input remains the same during the repetitions which increasingly lead to a more musical percept), but instead on the use of internal combinatorial rule systems (syntax). This would falsify an alternative account of music-language interference effects which proposes that music and language interact through sensory systems and the psychoacoustic properties of difficult to integrate tones/chords (Tillmann & Bigand, 2015).

Second, another methodological paradigm which has not been used in evaluating the SSIRH, yet, is transcranial magnetic stimulation (TMS). This technique can be used to interfere with the neural processing of a relatively limited patch of neocortex. According to the SSIRH, applying repetitive TMS to Broca's area in order to induce a so-called virtual temporal lesion should interfere with both music and language comprehension. A lesion to the posterior middle temporal gyrus identified as being specifically involved in language but not music in chapter 5, on the other hand, should impair language processing selectively without affecting music processing. The behavioural tasks used by Broca's aphasics could be used again in order to operationalise music and language processing (Patel et al., 2008). Such an investigation could provide causal (rather than just correlational) evidence for the distinction between music and language processing resources which are shared or not shared according to the SSIRH (Patel, 2003).

A third avenue to investigate the SSIRH relates to extending the stimulus material to include also non-Western musical systems. Whether the effects presented in this thesis and elsewhere will generalize to non-Western musical styles with similar in-key versus out-of-key contrasts (see for example recent unpublished work by Dr. Rachna Raman) is entirely unclear. Presumably, the shared syntactic resources between music and language are not dependent on the specific rules they need to implement, but instead on the operations which different combinatorial rules can have in common such as raising the activation level of an unexpected tone. Luckily, tonal harmony-like rules exist in many musical systems (Brown & Jordania, 2013), allowing future investigations to evaluate whether the SSIRH accounts well for how music and language comprehension relate to each other when culturally specific Western structural rules are applied or whether the SSIRH accounts well for how culturally general cognitive processes affect the relation between music and language comprehension.

However, I also see developments which call for an updated SSIRH, specifically regarding its shared music-language structural integration component. In terms of music, two experiments (chapter 2 and Van de Cavey & Hartsuiker, 2016) actually found evidence for a shared syntactic processing machinery whose role is different to what Patel (2003, 2008) predicted. To recall, the SSIRH hypothesizes that shared music-language resources rapidly and selectively increase the activation level of an unexpected element's (word or tone/chord) representation in long term memory up to threshold after which integration with the structural context can take place. However, chapter 2 revealed that the role of shared music-language resources in the music network is to hold an already established key online rather than quickly process unexpected chords in order to integrate them into a harmonic key. This is surprising since holding an established harmonic key online is exactly the opposite of dealing with unexpected chords.

Moreover, Van de Cavey & Hartsuiker (2016) found that the global structure of melodic sequences (with center-embedding like ABA versus without center embedding like ABB) can prime the interpretation of ambiguous sentence fragments of the form *The man sees the chairs of the room* which can be completed in an ABA like form with *…that are wide* (referring back to *the chairs* rather than *the room*) or in an ABB like form with *…that is spacious* (referring back to *the room* rather than *the chairs*). Apparently then, shared syntactic processing resources can hold a global structure in mind for re-use in another cognitive domain. Both chapter 2 and the results of Van de Cavey and Hartsuiker suggest that a syntactic working memory like component (Fiveash & Pammer, 2014) or perhaps the unification workspace proposed by Hagoort (2005) is also shared between music and language. Integrating such a similar component in the SSIRH will be an important future development of the SSIRH.

The SSIRH will also need updating in terms of the role of shared syntactic integration resources in language processing. For example, it is currently unclear whether the size of the to-be integrated elements is important for the integration mechanism. One could construct morpho-semantic garden-path sentences as in (7|1) which include a morphologically and semantically ambiguous word (**uitje**) which either means small onion and has the suffix '-tje' (**ui-tje**) or which means small day out and has the suffix '-je' (**uit-je**). After a local ambiguity, a disambiguating word (<u>gebakken</u> - baked) requires a reinterpretation of both the meaning and the morphological structure of the word. Should music harmony processing interact with language syntax processing (akin to chapter 2 and Slevc et al., 2009) but not with the processing of morpho-semantic garden-path sentences, then the integration mechanism requires a minimal element size for integration (syntactic word integration is included but smaller morphological part of word integration is not included in the shared processing mechanism).

Alternatively, if both syntactic and morpho-semantic garden-path sentences interact with music harmony processing, then the role of the shared structural integration mechanism would have to be extended to also include morphological processing. Either way, such an experiment would offer an interesting way to further specify the functional characteristics of shared music-language structural integration mechanisms.

(7|1)

a) morpho-semantic garden-path

Een **uitje** is alleen aangenaam als het goed <u>gebakken</u> is en nog vers ruikt.

Translation: *A small onion is only pleasant if it is well baked and still smells fresh.*

b) non-garden-path

Een **sjalotje** is alleen aangenaam als het goed <u>gebakken</u> is en nog vers ruikt.

Translation: *A small shallot is only pleasant i fit is well baked and still smells fresh.*

Future developments of the SSIRH might also make use of recent advances in the computational modeling of music and language. The application of computational models of note expectation to neuroscientific questions has recently been successful (Carrus et al., 2013; Pearce, Ruiz, Kapasi, Wiggins, & Bhattacharya, 2010). Similar developments have also taken place in the language domain (Frank, Otten, Galli, & Vigliocco, 2015; Willems, Frank, Nijhof, Hagoort, & Bosch, 2016). The chosen music and language models relate cognitive processing to the prediction of upcoming elements in the musical or linguistic input stream. Thus, combining both approaches allows for moving the SSIRH closer to the debate about the role of expectation and prediction in higher order cognition (Friston, 2005).

I hope that this short overview of possible future tests of the SSIRH shows that even after 13 years the SSIRH has lost none of its predictive power. It will be interesting to see how future studies develop the SSIRH further.

7|3 Beyond music and language

The SSIRH proposes a shared pre-frontal processing machinery for two different cognitive domains: music and language. Such a proposal is not unique for the link between music and language. Previously, the left inferior frontal gyrus, which chapter 3 suggests is the location of shared music-language resources, has been found to also be involved in action sequence processing (Fadiga, Craighero, & D'Ausilio, 2009; Fazio et al., 2009; Tettamanti & Weniger, 2006) and artificial grammar processing (Petersson, Forkstam, & Ingvar, 2004; Uddén & Bahlmann, 2012). Moreover, the processing of visual image sequences, i.e. comics (Cohn, Jackendoff, Holcomb, & Kuperberg, 2014; Cohn, Paczynski, Jackendoff, Holcomb, & Kuperberg, 2012), and even visual scenes (Võ & Wolfe, 2013) has been claimed to rely on separate semantic and structural/syntactic processing based on EEG evidence.

It will be interesting to see whether these cognitive domains display shared processing with music and language because none of them is tied to the auditory domain. As laid out in the introduction, acoustic signals can only convey one piece of information at a time, necessitating combinatorial rules in order to combine information bits (elements) into structured sequences. Visual signals are not limited in the same way. They allow for presenting more than one piece of information at the same time. Whether the processing resources specialized for auditory structured sequences (and their translation in written script) are so general as to also process purely visual information without an obvious auditory equivalent will be an interesting question for the future. As a first step, one might present the visual image sequences of Cohn et al. (2014) with or without an unexpected blank image (which disrupts syntactic processing) and observe whether this manipulation interacts with music perception akin to the effect of language structural processing difficulties; see chapter 2.

7|4 Conclusion

What happens in your head when you hear music? What happens when you hear language? These questions have puzzled researchers for decades and this PhD thesis contributes but a small part to a satisfying, encompassing answer. The results suggest that the structural relations inherent in the music and the language stimuli are partly processed by the same neural circuitry in the inferior frontal gyrus. This processing machinery relies on representations in posterior brain areas (the middle temporal gyrus in the case of language) in order to form higher order sequences (sentences or melodies/harmonic sequences) out of identified elements (words or tones/chords). The time-course of this process is largely unclear. It is also unclear how musical aptitude affects these processes. However, at least one novel musical aptitude measure offers a promising avenue for future research.

What is clear is that the account of how music and language resources relate to each other has consequences for behaviour. When you sit at the breakfast table reading the newspaper and you listen to songs on the radio, these two activities probably influence each other. I hope that such small mutual influences between music and language processing can be used in the future in order to affect everyday life in positive ways. If so, this PhD thesis might be used as one source of inspiration for how music and language comprehension can be optimized.

References

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. http://doi.org/10.1093/cercor/bhq198

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. http://doi.org/10.1016/S0022-0965(02)00124-8

Bastiaansen, M. C. M., & Hagoort, P. (2006). Oscillatory neuronal dynamics during language comprehension. In C. N. and W. Klimesch (Ed.), *Progress in Brain Research* (Vol. 159, pp. 179–196). Elsevier. Retrieved from http://www.sciencedirect.com/science/article/pii/S0079612306590120

Bastiaansen, M. C. M., van Berkum, J. J. A., & Hagoort, P. (2002). Syntactic Processing Modulates the θ Rhythm of the Human EEG. *NeuroImage*, *17*(3), 1479–1492. http://doi.org/10.1006/nimg.2002.1275

Bermudez, P., Lerch, J. P., Evans, A. C., & Zatorre, R. J. (2009). Neuroanatomical correlates of musicianship as revealed by cortical thickness and voxel-based morphometry. *Cerebral Cortex*, *19*(7), 1583–1596. http://doi.org/10.1093/cercor/bhn196

Besson, M., Faïta, F., Peretz, I., Bonnel, A.-M., & Requin, J. (1998). Singing in the brain: independence of lyrics and tunes. *Psychological Science*, *9*(6), 494–498. http://doi.org/10.1111/1467-9280.00091

Bigand, E., & Pineau, M. (1997). Global context effects on musical expectancy. *Perception & Psychophysics*, *59*(7), 1098–1107. http://doi.org/10.3758/BF03205524

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. http://doi.org/10.1016/j.cognition.2005.11.007

Bigand, E., Tillmann, B., Poulin, B., D’Adamo, D. A., & Madurell, F. (2001). The effect of harmonic context on phoneme monitoring in vocal music. *Cognition*, *81*(1), B11–B20. http://doi.org/10.1016/S0010-0277(01)00117-2

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. http://doi.org/10.1038/nn1198

Bookheimer, S. (2002). FUNCTIONAL MRI OF LANGUAGE: New approaches to understanding the cortical organization of semantic processing. *Annual Review of Neuroscience*, *25*(1), 151–188. http://doi.org/10.1146/annurev.neuro.25.112701.142946

Brett, M., Anton, J. L., Valabregue, R., & Poline, J. B. (2002). Region of interest analysis using an SPM toolbox. Presented at the 8th international conference on functional mapping of the human brain, Sendai, Japan.

Brown, S., & Jordania, J. (2013). Universals in the world's musics. *Psychology of Music*, *41*(2), 229–248. http://doi.org/10.1177/0305735611425896

Carroll, L. (1872). *Through the Looking-Glass*. Mineola, N.Y: Dover Publications.

Carrus, E., Koelsch, S., & Bhattacharya, J. (2011). Shadows of music–language interaction on low frequency brain oscillatory patterns. *Brain and Language*, *119*(1), 50–57.

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. http://doi.org/10.1016/j.cortex.2012.08.024

Chomsky, N. (1956). Three models for the description of language. *IRE Transactions on Information Theory*, *2*(3), 113–124. http://doi.org/10.1109/TIT.1956.1056813

Chumbley, J. R., & Friston, K. J. (2009). False discovery rate revisited: FDR and topological inference using Gaussian random fields. *NeuroImage*, *44*(1), 62–70. http://doi.org/10.1016/j.neuroimage.2008.05.021

Chumbley, J. R., Worsley, K., Flandin, G., & Friston, K. (2010). Topological FDR for neuroimaging. *NeuroImage*, *49*(4), 3057–3064. http://doi.org/10.1016/j.neuroimage.2009.10.090

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155.

Cohn, N., Jackendoff, R., Holcomb, P. J., & Kuperberg, G. R. (2014). The grammar of visual narrative: Neural evidence for constituent structure in sequential image comprehension. *Neuropsychologia*, *64*, 63–70. http://doi.org/10.1016/j.neuropsychologia.2014.09.018

Cohn, N., Paczynski, M., Jackendoff, R., Holcomb, P. J., & Kuperberg, G. R. (2012). (Pea)nuts and bolts of visual narrative: Structure and meaning in sequential image comprehension. *Cognitive Psychology*, *65*(1), 1–38. http://doi.org/10.1016/j.cogpsych.2012.01.003

Copland, D. A. (2006). Meaning Selection and the Subcortex: Evidence of Reduced Lexical Ambiguity Repetition Effects Following Subcortical Lesions. *Journal of Psycholinguistic Research*, *35*(1), 51–66. http://doi.org/10.1007/s10936-005-9003-6

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*(3), 201–215. http://doi.org/10.1038/nrn755

Deutsch, D., Henthorn, T., & Lapidis, R. (2011). Illusory transformation from speech to song. *The Journal of the Acoustical Society of America*, *129*(4), 2245–2252. http://doi.org/10.1121/1.3562174

Drai, D., & Grodzinsky, Y. (2006). A new empirical angle on the variability debate: Quantitative neurosyntactic analyses of a large data set from Broca's Aphasia. *Brain and Language*, *96*(2), 117–128. http://doi.org/10.1016/j.bandl.2004.10.016

Eerola, T., Himberg, T., Toiviainen, P., & Louhivuori, J. (2006). Perceived complexity of western and African folk melodies by western and African listeners. *Psychology of Music*, *34*(3), 337–371. http://doi.org/10.1177/0305735606064842

Embick, D., Marantz, A., Miyashita, Y., O'Neil, W., & Sakai, K. L. (2000). A syntactic specialization for Broca's area. *Proceedings of the National Academy of Sciences*, *97*(11), 6150–6154. http://doi.org/10.1073/pnas.100098897

Ericsson, A. K., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 363–406. http://doi.org/10.1037/0033-295X.100.3.363

Escoffier, N., & Tillmann, B. (2008). The tonal function of a task-irrelevant chord modulates speed of visual processing. *Cognition*, *107*(3), 1070–1083. http://doi.org/10.1016/j.cognition.2007.10.007

Fadiga, L., Craighero, L., & D'Ausilio, A. (2009). Broca's area in language, action, and music. *Annals of the New York Academy of Sciences*, *1169*(1), 448–458. http://doi.org/10.1111/j.1749-6632.2009.04582.x

Falk, S., Rathcke, T., & Dalla Bella, S. (2014). When speech sounds like music. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(4), 1491–1506. http://doi.org/10.1037/a0036858

Fazio, P., Cantagallo, A., Craighero, L., D'Ausilio, A., Roy, A. C., Pozzo, T., … Fadiga, L. (2009). Encoding of human action in Broca's area. *Brain*, *132*(7), 1980–1988. http://doi.org/10.1093/brain/awp118

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. http://doi.org/10.1371/journal.pone.0076600

Fedorenko, E., Behr, M. K., & Kanwisher, N. (2011). Functional specificity for high-level linguistic processing in the human brain. *Proceedings of the National Academy of Sciences*, *108*(39), 16428–16433. http://doi.org/10.1073/pnas.1112937108

Fedorenko, E., McDermott, J. H., Norman-Haignere, S., & Kanwisher, N. (2012). Sensitivity to musical structure in the human brain. *Journal of Neurophysiology*, *108*(12), 3289–3300. http://doi.org/10.1152/jn.00209.2012

Fedorenko, E., Patel, A., Casasanto, D., Winawer, J., & Gibson, E. (2009). Structural integration in language and music: Evidence for a shared system. *Memory & Cognition*, *37*(1), 1–9. http://doi.org/10.3758/MC.37.1.1

Fedorenko, E., & Varley, R. (2016). Language and thought are not the same thing: evidence from neuroimaging and neurological patients. *Annals of the New York Academy of Sciences*, *1369*(1), 132–153. http://doi.org/10.1111/nyas.13046

Ferreira, F., Bailey, K. G. D., & Ferraro, V. (2002). Good-Enough Representations in Language Comprehension. *Current Directions in Psychological Science*, *11*(1), 11–15. http://doi.org/10.1111/1467-8721.00158

Fiebach, C. j., Schlesewsky, M., Lohmann, G., von Cramon, D. y., & Friederici, A. d. (2005). Revisiting the role of Broca's area in sentence processing: Syntactic integration versus syntactic working memory. *Human Brain Mapping*, *24*(2), 79–91. http://doi.org/10.1002/hbm.20070

Fitzroy, A. B., & Sanders, L. D. (2013). Musical expertise modulates early processing of syntactic violations in language. *Frontiers in Psychology*, *3*. http://doi.org/10.3389/fpsyg.2012.00603

Fiveash, A., & Pammer, K. (2014). Music and language: Do they draw on similar syntactic working memory resources? *Psychology of Music*, *42*(2), 190–209. http://doi.org/10.1177/0305735612463949

Fox, M. D., Corbetta, M., Snyder, A. Z., Vincent, J. L., & Raichle, M. E. (2006). Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. *Proceedings of the National Academy of Sciences*, *103*(26), 10046–10051. http://doi.org/10.1073/pnas.0604187103

Frank, S. L., Otten, L. J., Galli, G., & Vigliocco, G. (2015). The ERP response to the amount of information conveyed by words in sentences. *Brain and Language*, *140*, 1–11. http://doi.org/10.1016/j.bandl.2014.10.006

Frazier, L. (1987). Syntactic processing: Evidence from dutch. *Natural Language & Linguistic Theory*, *5*(4), 519–559. http://doi.org/10.1007/BF00138988

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. http://doi.org/10.1002/1097-0193(200009)11:1<1::AID-HBM10>3.0.CO;2-B

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *360*(1456), 815–836. http://doi.org/10.1098/rstb.2005.1622

Garza Villarreal, E. A., Brattico, E., Leino, S., Østergaard, L., & Vuust, P. (2011). Distinct neural responses to chord violations: A multiple source analysis study. *Brain Research*, *1389*, 103–114. http://doi.org/10.1016/j.brainres.2011.02.089

Gaser, C., & Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *The Journal of Neuroscience*, *23*(27), 9240–9245.

Gibson, E. (1998). Linguistic complexity: locality of syntactic dependencies. *Cognition*, *68*(1), 1–76. http://doi.org/10.1016/S0010-0277(98)00034-1

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.

Gingras, B., Honing, H., Peretz, I., Trainor, L. J., & Fisher, S. E. (2015). Defining the biological bases of individual differences in musicality. *Phil. Trans. R. Soc. B*, *370*(1664), 20140092. http://doi.org/10.1098/rstb.2014.0092

Gordon, E. E. (1989). *Advanced measures of music audiation: kit. CD. Advanced measures of music audiation*. GIA Publ.

Griffiths, T. D., Rees, A., Witton, C., Cross, P. M., Shakir, R. A., & Green, G. G. (1997). Spatial and temporal auditory processing deficits following right hemisphere infarction. A psychophysical study. *Brain*, *120*(5), 785–794. http://doi.org/10.1093/brain/120.5.785

Groth-Marnat, G. (2001). The Wechsler intelligence scales. *Specific Learning Disabilities and Difficulties in Children and Adolescent. Psychological Assessment and Evaluation*, 29–51.

Hagoort, P. (2005). On Broca, brain, and binding: a new framework. *Trends in Cognitive Sciences*, *9*(9), 416–423. http://doi.org/10.1016/j.tics.2005.07.004

Hagoort, P. (2013). MUC (Memory, Unification, Control) and beyond. *Frontiers in Psychology*, *4*. http://doi.org/10.3389/fpsyg.2013.00416

Hansen, M., Wallentin, M., & Vuust, P. (2013). Working memory and musical competence of musicians and non-musicians. *Psychology of Music*, *41*(6), 779–793. http://doi.org/10.1177/0305735612452186

Herdener, M., Humbel, T., Esposito, F., Habermeyer, B., Cattapan-Ludewig, K., & Seifritz, E. (2014). Jazz Drummers Recruit Language-Specific Areas for the Processing of Rhythmic Structure. *Cerebral Cortex*, *24*(3), 836–843. http://doi.org/10.1093/cercor/bhs367

Hillebrand, A., & Barnes, G. R. (2002). A Quantitative Assessment of the Sensitivity of Whole-Head MEG to Activity in the Adult Human Cortex. *NeuroImage*, *16*(3, Part A), 638–650. http://doi.org/10.1006/nimg.2002.1102

Hoch, L., Poulin-Charronnat, B., & Tillmann, B. (2011). The influence of task-irrelevant music on language processing: syntactic and semantic structures. *Frontiers in Psychology*, *2*. http://doi.org/10.3389/fpsyg.2011.00112

Hoeks, J. C. J., Hendriks, P., Vonk, W., Brown, C. M., & Hagoort, P. (2006). Processing the noun phrase versus sentence coordination ambiguity: Thematic information does not completely eliminate processing difficulty. *The Quarterly Journal of Experimental Psychology*, *59*(9), 1581–1599. http://doi.org/10.1080/17470210500268982

Hoeks, J. C. J., Vonk, W., & Schriefers, H. (2002). Processing coordinated structures in context: The effect of topic-structure on ambiguity resolution. *Journal of Memory and Language*, *46*(1), 99–119. http://doi.org/10.1006/jmla.2001.2800

Indefrey, P., Hagoort, P., Herzog, H., Seitz, R. J., & Brown, C. M. (2001). Syntactic processing in left prefrontal cortex is independent of lexical meaning. *NeuroImage*, *14*(3), 546–555. http://doi.org/10.1006/nimg.2001.0867

Jackendoff, R. (2009). Parallels and Nonparallels between Language and Music. *Music Perception: An Interdisciplinary Journal*, *26*(3), 195–204. http://doi.org/10.1525/mp.2009.26.3.195

James, C. E., Oechslin, M. S., Ville, D. V. D., Hauert, C.-A., Descloux, C., & Lazeyras, F. (2014). Musical training intensity yields opposite effects on grey matter density in cognitive versus sensorimotor networks. *Brain Structure and Function*, *219*(1), 353–366. http://doi.org/10.1007/s00429-013-0504-z

Jeffreys, H. (1961). *Theory of probability*. New York: Oxford University Press.

Jentschke, S., & Koelsch, S. (2009). Musical training modulates the development of syntax processing in children. *Neuroimage*, *47*(2), 735–744. http://doi.org/10.1016/j.neuroimage.2009.04.090

Jentschke, S., Koelsch, S., Sallat, S., & Friederici, A. D. (2008). Children with Specific Language Impairment Also Show Impairment of Music-syntactic Processing. *Journal of Cognitive Neuroscience*, *20*(11), 1940–1951. http://doi.org/10.1162/jocn.2008.20135

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*, *96*(3), 459–491. http://doi.org/10.1037//0033-295X.96.3.459

Kerkhofs, R., Vonk, W., Schriefers, H., & Chwilla, D. J. (2008). Sentence processing in the visual and auditory modality: Do comma and prosodic break have parallel functions? *Brain Research*, *1224*, 102–118. http://doi.org/10.1016/j.brainres.2008.05.034

Koelsch, S. (2011). Towards a neural basis of processing musical semantics. *Physics of Life Reviews*, *8*(2), 89–105. http://doi.org/10.1016/j.plrev.2011.04.004

Koelsch, S., Fritz, T., Schulze, K., Alsop, D., & Schlaug, G. (2005). Adults and children processing music: An fMRI study. *NeuroImage*, *25*(4), 1068–1076. http://doi.org/10.1016/j.neuroimage.2004.12.050

Koelsch, S., Gunter, T. C., von Cramon, D. Y., Zysset, S., Lohmann, G., & Friederici, A. D. (2002). Bach speaks: A cortical 'language-network' serves the processing of music. *Neuroimage*, *17*(2), 956–966. http://doi.org/10.1016/S1053-8119(02)91154-7

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. http://doi.org/10.1162/089892905774597290

Koelsch, S., Gunter, T., Friederici, A. D., & Schröger, E. (2000). Brain indices of music processing: 'nonmusicians' are musical. *Journal of Cognitive Neuroscience*, *12*(3), 520–541. http://doi.org/10.1162/089892900562183

Koelsch, S., Rohrmeier, M., Torrecuso, R., & Jentschke, S. (2013). Processing of hierarchical syntactic structure in music. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(38), 15443–15448. http://doi.org/10.1073/pnas.1300272110

Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S. F., & Baker, C. I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nature Neuroscience*, *12*(5), 535–540. http://doi.org/10.1038/nn.2303

Krumhansl, C. L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology*, *11*(3), 346–374. http://doi.org/10.1016/0010-0285(79)90016-1

Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, *89*(4), 334–368. http://doi.org/10.1037/0033-295X.89.4.334

Kruschke, J. K. (2013). Bayesian Estimation Supersedes the t Test. *Journal of Experimental Psychology: General*, *142*(2), 573–603. http://doi.org/DOI: 10.1037/a0029146

Kunert, R. (2016). Internal conceptual replications do not increase independent replication success. *Psychonomic Bulletin & Review*, 1–8. http://doi.org/10.3758/s13423-016-1030-9

Kunert, R., & Slevc, L. R. (2015). A Commentary on: 'Neural overlap in processing music and speech'. *Frontiers in Human Neuroscience*, *9*, 330. http://doi.org/10.3389/fnhum.2015.00330

Kunert, R., Willems, R. M., Casasanto, D., Patel, A. D., & Hagoort, P. (2015). Music and Language Syntax Interact in Broca's Area: An fMRI Study. *PLoS ONE*, *10*(11), e0141069. http://doi.org/10.1371/journal.pone.0141069

Kunert, R., Willems, R. M., & Hagoort, P. (2016). Language influences music harmony perception: effects of shared syntactic integration resources beyond attention. *Royal Society Open Science*, *3*(2), 150685. http://doi.org/10.1098/rsos.150685

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, *106*(1), 119–159. http://doi.org/10.1037/0033-295X.106.1.119

Law, L. N. C., & Zentner, M. (2012). Assessing musical abilities objectively: Construction and validation of the Profile of Music Perception Skills. *PLoS ONE*, *7*(12), e52508. http://doi.org/10.1371/journal.pone.0052508

Lerdahl, F. (2001). *Tonal Pitch Space*. Oxford University Press.

Lerdahl, F. (2013). Tension and expectation in a Schubert song. In L. F. Bernstein & A. Rozins (Eds.), *Musical Implications: Essays in Honor of Eugene Narmour* (pp. 255–274). Hillsdale, NY: Pendragon Press.

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

Lund, T. E., Nørgaard, M. D., Rostrup, E., Rowe, J. B., & Paulson, O. B. (2005). Motion or activity: their role in intra- and inter-subject variation in fMRI. *NeuroImage*, *26*(3), 960–964. http://doi.org/10.1016/j.neuroimage.2005.02.021

Macnamara, B. N., Hambrick, D. Z., & Oswald, F. L. (2014). Deliberate Practice and Performance in Music, Games, Sports, Education, and Professions A Meta-Analysis. *Psychological Science*, *25*(8), 1608–1618. http://doi.org/10.1177/0956797614535810

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. http://doi.org/10.1038/87502

Maidhof, C., & Koelsch, S. (2011). Effects of selective attention on syntax processing in music and language. *Journal of Cognitive Neuroscience*, *23*(9), 2252–2267. http://doi.org/10.1162/jocn.2010.21542

Makuuchi, M., Bahlmann, J., Anwander, A., & Friederici, A. D. (2009). Segregating the core computational faculty of human language from working memory. *Proceedings of the National Academy of Sciences*, *106*(20), 8362–8367. http://doi.org/10.1073/pnas.0810928106

Manly, B. F. J. (2006). *Randomization, bootstrap and Monte Carlo methods in Biology, Third Edition*. CRC Press.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. http://doi.org/10.1016/j.jneumeth.2007.03.024

Martin, R. C. (1993). Short-term memory and sentence processing: Evidence from neuropsychology. *Memory & Cognition*, *21*(2), 176–183. http://doi.org/10.3758/BF03202730

Menenti, L., Gierhan, S. M. E., Segaert, K., & Hagoort, P. (2011). Shared Language Overlap and Segregation of the Neuronal Infrastructure for Speaking and Listening Revealed by Functional MRI. *Psychological Science*, *22*(9), 1173–1182. http://doi.org/10.1177/0956797611418347

Meredith, M., & Kruschke, J. K. (2015). Package 'BEST' (Version 0.4.0). Retrieved from https://cran.r-project.org/web/packages/BEST/index.html

Miyazaki, K. (1988). Musical pitch identification by absolute pitch possessors. *Perception & Psychophysics*, *44*(6), 501–512. http://doi.org/10.3758/BF03207484

Morey, R. D., Rouder, J. N., & Jamil, T. (2015). Package 'BayesFactor' (Version 0.9.12-2). Retrieved from http://bayesfactorpcl.r-forge.r-project.org/

Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). The Musicality of Non-Musicians: An Index for Assessing Musical Sophistication in the General Population. *PLOS ONE*, *9*(2), e89642. http://doi.org/10.1371/journal.pone.0089642

Musso, M., Weiller, C., Horn, A., Glauche, V., Umarova, R., Hennig, J., … Rijntjes, M. (2015). A single dual-stream framework for syntactic computations in music and language. *NeuroImage*, *117*, 267–283. http://doi.org/10.1016/j.neuroimage.2015.05.020

Nieto-Castañón, A., & Fedorenko, E. (2012). Subject-specific functional localizers increase sensitivity and functional resolution of multi-subject analyses. *NeuroImage*, *63*(3), 1646–1669. http://doi.org/10.1016/j.neuroimage.2012.06.065

Noppeney, U., & Price, C. J. (2004). An fMRI Study of Syntactic Adaptation. *Journal of Cognitive Neuroscience*, *16*(4), 702–713. http://doi.org/10.1162/089892904323057399

Norman-Haignere, S., Kanwisher, N. G., & McDermott, J. H. (2015). Distinct Cortical Pathways for Music and Speech Revealed by Hypothesis-Free Voxel Decomposition. *Neuron*, *88*(6), 1281–1296. http://doi.org/10.1016/j.neuron.2015.11.035

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*, e156869. http://doi.org/10.1155/2011/156869

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716. http://doi.org/10.1126/science.aac4716

Pasinski, A. C., Hannon, E. E., & Snyder, J. S. (2016). How musical are music video game players? *Psychonomic Bulletin & Review*, 1–6. http://doi.org/10.3758/s13423-015-0998-x

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, *6*(7), 674–681. http://doi.org/10.1038/nn1082

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

Patel, A. D. (2013). Sharing and nonsharing of brain resources for language and music. In M. Arbib (Ed.), *Language, Music, and the Brain: A Mysterious Relationship* (pp. 329–355). Cambridge, MA: MIT Press.

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. http://doi.org/10.1162/089892998563121

Patel, A. D., Iversen, J. R., Wassenaar, M., & Hagoort, P. (2008). Musical syntactic processing in agrammatic Broca’s aphasia. *Aphasiology*, *22*(7–8), 776–789. http://doi.org/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. http://doi.org/10.1016/j.neuroimage.2009.12.019

Peretz, I. (1993). Auditory atonalia for melodies. *Cognitive Neuropsychology*, *10*(1), 21–56.

Peretz, I., Belleville, S., & Fontaine, S. (1997). Dissociations entre musique et langage après atteinte cérébrale: Un nouveau cas d’amusie sans aphasie. [Dissociations between music and language after cerebral damage: A new case of music deficits without aphasia.]. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *51*(4), 354–368. http://doi.org/10.1037/1196-1961.51.4.354

Peretz, I., Champod, A. S., & Hyde, K. (2003). Varieties of musical disorders: the Montreal battery of evaluation of amusia. *Annals of the New York Academy of Sciences*, *999*(1), 58–75. http://doi.org/10.1196/annals.1284.006

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, *6*(7), 688–691. http://doi.org/10.1038/nn1083

Peretz, I., Gosselin, N., Nan, Y., Caron-Caplette, E., Trehub, S. E., & Béland, R. (2013). A novel tool for evaluating children's musical abilities across age and culture. *Frontiers in Systems Neuroscience*, *7*, 30. http://doi.org/10.3389/fnsys.2013.00030

Peretz, I., Kolinsky, R., Tramo, M., Labrecque, R., Hublet, C., Demeurisse, G., & Belleville, S. (1994). Functional dissociations following bilateral lesions of auditory cortex. *Brain*, *117*(6), 1283–1301. http://doi.org/10.1093/brain/117.6.1283

Peretz, I., Vuvan, D., Lagrois, M.-É., & Armony, J. L. (2015a). Neural overlap in processing music and speech. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *370*(1664), 20140090. http://doi.org/10.1098/rstb.2014.0090

Peretz, I., Vuvan, D., Lagrois, M.-É., & Armony, J. L. (2015b). Neural overlap in processing music and speech. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1664), 20140090. http://doi.org/10.1098/rstb.2014.0090

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. http://doi.org/10.3758/s13423-012-0344-5

Petersson, K. M., Forkstam, C., & Ingvar, M. (2004). Artificial syntactic violations activate Broca's region. *Cognitive Science*, *28*(3), 383–407. http://doi.org/10.1016/j.cogsci.2003.12.003

Poser, B. A., Versluis, M. J., Hoogduin, J. M., & Norris, D. G. (2006). BOLD contrast sensitivity enhancement and artifact reduction with multiecho EPI: Parallel-acquired inhomogeneity-desensitized fMRI. *Magnetic Resonance in Medicine*, *55*(6), 1227–1235. http://doi.org/10.1002/mrm.20900

Poulin-Charronnat, B., Bigand, E., Madurell, F., & Peereman, R. (2005). Musical structure modulates semantic priming in vocal music. *Cognition*, *94*(3), B67–B78. http://doi.org/10.1016/j.cognition.2004.05.003

Rogalsky, C., Rong, F., Saberi, K., & Hickok, G. (2011). Functional anatomy of language and music perception: Temporal and structural factors investigated using functional Magnetic Resonance Imaging. *Journal of Neuroscience*, *31*(10), 3843–3852. http://doi.org/10.1523/JNEUROSCI.4515-10.2011

Rohrmeier, M. (2011). A generative grammar approach to tonal harmony. *J Math Music*, *5*(1), 35–53.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*(2), 225–237. http://doi.org/10.3758/PBR.16.2.225

Sacks, O. (1995). Musical ability. *Science*, *268*(5211), 621–622. http://doi.org/10.1126/science.7732360

Sammler, D., Baird, A., Valabrègue, R., Clément, S., Dupont, S., Belin, P., & Samson, S. (2010). The Relationship of Lyrics and Tunes in the Processing of Unfamiliar Songs: A Functional Magnetic Resonance Adaptation Study. *The Journal of Neuroscience*, *30*(10), 3572–3578. http://doi.org/10.1523/JNEUROSCI.2751-09.2010

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. http://doi.org/10.1016/j.cortex.2010.04.007

Santi, A., & Grodzinsky, Y. (2010). fMRI adaptation dissociates syntactic complexity dimensions. *NeuroImage*, *51*(4), 1285–1293. http://doi.org/10.1016/j.neuroimage.2010.03.034

Schaal, N. K., Bauer, A.-K. R., & Müllensiefen, D. (2014). Der Gold-MSI: Replikation und Validierung eines Fragebogeninstrumentes zur Messung Musikalischer Erfahrenheit anhand einer deutschen Stichprobe. *Musicae Scientiae*, *18*(4), 423–447. http://doi.org/10.1177/1029864914541851

Schachner, A., Brady, T. F., Pepperberg, I. M., & Hauser, M. D. (2009). Spontaneous Motor Entrainment to Music in Multiple Vocal Mimicking Species. *Current Biology*, *19*(10), 831–836. http://doi.org/10.1016/j.cub.2009.03.061

Schoot, L., Menenti, L., Hagoort, P., & Segaert, K. (2014). A little more conversation – the influence of communicative context on syntactic priming in brain and behavior. *Language Sciences*, *5*, 208. http://doi.org/10.3389/fpsyg.2014.00208

Seashore, C. E., Lewis, D., & Saetveit, J. G. (1960). *Measures of Musical Talents (test)*. Psychological Corporation.

Segaert, K., Kempen, G., Petersson, K. M., & Hagoort, P. (2013). Syntactic priming and the lexical boost effect during sentence production and sentence comprehension: An fMRI study. *Brain and Language*, *124*(2), 174–183. http://doi.org/10.1016/j.bandl.2012.12.003

Segaert, K., Menenti, L., Weber, K., Petersson, K. M., & Hagoort, P. (2012). Shared syntax in language production and language comprehension—an fMRI study. *Cerebral Cortex*, *22*(7), 1662–1670. http://doi.org/10.1093/cercor/bhr249

Segaert, K., Weber, K., de Lange, F. P., Petersson, K. M., & Hagoort, P. (2013). The suppression of repetition enhancement: A review of fMRI studies. *Neuropsychologia*, *51*(1), 59–66. http://doi.org/10.1016/j.neuropsychologia.2012.11.006

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. http://doi.org/10.1162/jocn_a_00371

Service, E., Helenius, P., Maury, S., & Salmelin, R. (2007). Localization of Syntactic and Semantic Brain Responses using Magnetoencephalography. *Journal of Cognitive Neuroscience*, *19*(7), 1193–1205. http://doi.org/10.1162/jocn.2007.19.7.1193

Slevc, L. R., & Miyake, A. (2006). Individual differences in second-language proficiency: Does musical ability matter? *Psychological Science*, *17*(8), 675–681.

Slevc, L. R., & Okada, B. M. (2014). Processing structure in language and music: a case for shared reliance on cognitive control. *Psychonomic Bulletin & Review*, 1–16. http://doi.org/10.3758/s13423-014-0712-4

Slevc, L. R., & Patel, A. D. (2011). Meaning in music and language: Three key differences: Comment on 'Towards a neural basis of processing musical semantics' by Stefan Koelsch. *Physics of Life Reviews*, *8*(2), 110–111.

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. http://doi.org/10.3758/16.2.374

Sluming, V., Barrick, T., Howard, M., Cezayirli, E., Mayes, A., & Roberts, N. (2002). Voxel-based morphometry reveals increased gray matter density in Broca's area in male symphony orchestra musicians. *NeuroImage*, *17*(3), 1613–1622. http://doi.org/10.1006/nimg.2002.1288

Snijders, T. M., Vosse, T., Kempen, G., Berkum, J. J. A. V., Petersson, K. M., & Hagoort, P. (2009). Retrieval and unification of syntactic structure in sentence comprehension: an fMRI study using word-category ambiguity. *Cerebral Cortex*, *19*(7), 1493–1503. http://doi.org/10.1093/cercor/bhn187

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. http://doi.org/10.1093/cercor/bhm149

Stolk, A., Todorovic, A., Schoffelen, J.-M., & Oostenveld, R. (2013). Online and offline tools for head movement compensation in MEG. *NeuroImage*, *68*, 39–48. http://doi.org/10.1016/j.neuroimage.2012.11.047

Tekman, H. G., & Bharucha, J. J. (1998). Implicit knowledge versus psychoacoustic similarity in priming of chords. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(1), 252–260. http://doi.org/10.1037/0096-1523.24.1.252

Tettamanti, M., & Weniger, D. (2006). Broca's area: a supramodal hierarchical processor? *Cortex*, *42*(4), 491–494. http://doi.org/10.1016/S0010-9452(08)70384-8

Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: A self-organizing approach. *Psychological Review*, *107*(4), 885–913. http://doi.org/10.1037//0033-295X.107.4.885

Tillmann, B., & Bigand, E. (2004). The Relative Importance of Local and Global Structures in Music Perception. *The Journal of Aesthetics and Art Criticism*, *62*(2), 211–222. http://doi.org/10.1111/j.1540-594X.2004.00153.x

Tillmann, B., & Bigand, E. (2015). Response: A commentary on: 'Neural overlap in processing music and speech'. *Frontiers in Human Neuroscience*, 491. http://doi.org/10.3389/fnhum.2015.00491

Tillmann, B., Bigand, E., Escoffier, N., & Lalitte, P. (2006). The influence of musical relatedness on timbre discrimination. *European Journal of Cognitive Psychology*, *18*(3), 343–358. http://doi.org/10.1080/09541440500269548

Tillmann, B., Janata, P., & Bharucha, J. J. (2003). Activation of the inferior frontal cortex in musical priming. *Cognitive Brain Research*, *16*(2), 145–161. http://doi.org/10.1016/S0926-6410(02)00245-8

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*(4), 1771–1782. http://doi.org/10.1016/j.neuroimage.2006.02.028

Tillmann, B., & Lebrun-Guillaud, G. (2005). Influence of tonal and temporal expectations on chord processing and on completion judgments of chord sequences. *Psychological Research*, *70*(5), 345–358. http://doi.org/10.1007/s00426-005-0222-0

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., … Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, *15*(1), 273–289. http://doi.org/10.1006/nimg.2001.0978

Uddén, J., & Bahlmann, J. (2012). A rostro-caudal gradient of structured sequence processing in the left inferior frontal gyrus. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1598), 2023–2032. http://doi.org/10.1098/rstb.2012.0009

Ullén, F., Mosing, M. A., Holm, L., Eriksson, H., & Madison, G. (2014). Psychometric properties and heritability of a new online test for musicality, the Swedish Musical Discrimination Test. *Personality and Individual Differences*, *63*, 87–93. http://doi.org/10.1016/j.paid.2014.01.057

Van de Cavey, J., & Hartsuiker, R. J. (2016). Is there a domain-general cognitive structuring system? Evidence from structural priming across music, math, action descriptions, and language. *Cognition*, *146*, 172–184. http://doi.org/10.1016/j.cognition.2015.09.013

Vanden Bosch der Nederlanden, C. M., Hannon, E. E., & Snyder, J. S. (2015). Everyday musical experience is sufficient to perceive the speech-to-song illusion. *Journal of Experimental Psychology: General*, *144*(2), e43–e49. http://doi.org/10.1037/xge0000056

Võ, M. L.-H., & Wolfe, J. M. (2013). Differential Electrophysiological Signatures of Semantic and Syntactic Scene Processing. *Psychological Science*, *24*(9), 1816–1823. http://doi.org/10.1177/0956797613476955

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. http://doi.org/10.1016/j.lindif.2010.02.004

Waters, G., Caplan, D., Alpert, N., & Stanczak, L. (2003). Individual differences in rCBF correlates of syntactic processing in sentence comprehension: effects of working memory and speed of processing. *NeuroImage*, *19*(1), 101–112. http://doi.org/10.1016/S1053-8119(03)00007-7

Weber, K., & Indefrey, P. (2009). Syntactic priming in German–English bilinguals during sentence comprehension. *NeuroImage*, *46*(4), 1164–1172. http://doi.org/10.1016/j.neuroimage.2009.03.040

Westen, D., & Rosenthal, R. (2003). Quantifying construct validity: Two simple measures. *Journal of Personality and Social Psychology*, *84*(3), 608–618. http://doi.org/10.1037/0022-3514.84.3.608

Willems, R. M., Frank, S. L., Nijhof, A. D., Hagoort, P., & Bosch, A. van den. (2016). Prediction During Natural Language Comprehension. *Cerebral Cortex*, *26*(6), 2506–2516. http://doi.org/10.1093/cercor/bhv075

Xiong, J., Rao, S., Jerabek, P., Zamarripa, F., Woldorff, M., Lancaster, J., & Fox, P. T. (2000). Intersubject variability in cortical activations during a complex language task. *NeuroImage*, *12*(3), 326–339. http://doi.org/10.1006/nimg.2000.0621

Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends in Cognitive Sciences*, *6*(1), 37–46. http://doi.org/10.1016/S1364-6613(00)01816-7

Samenvatting

Wat gebeurt er in de hersenen als we taal lezen of horen en muziek horen? Dat is de centrale vraag in mijn PhD thesis. Volgens de zogenaamde Shared Syntactic Integration Resource Hypothesis wordt de structuur in taal en muziek stimuli gedeeltelijk door dezelfde hersensystemen verwerkt. In hoofdstuk 2 onderzocht ik met gedragsmaten of er inderdaad dit soort gedeelde hersensystemen bestaan. Nadat proefpersonen muziek hadden gehoord en tegelijkertijd een zin hadden gelezen, werd hen gevraagd om muziekstimuli te beoordelen op hun harmonisch einde. Wat bleek: syntactisch (structureel) ingewikkelde zinnen konden de harmoniebeoordelingen veranderen, alsof dit soort zinnen de gedeelde hersensystemen uitdaagden waardoor er minder goede muziekverwerking mogelijk was. In hoofdstuk 3 onderzocht ik of de gedeelde systemen die muziek- en taalstructuur verwerken in een prefrontale hersenregio bestaan zoals voorspeld door de Shared Syntactic Integration Resource Hypothesis. Dat leek inderdaad zo te zijn. De temporele eigenschappen van de processen in dezelfde hersenregio blijven onduidelijk (hoofdstuk 4). In hoofdstuk 5 keek ik zowel naar hersenregio's die gedeeld worden door taal en muziek en naar niet gedeelde hersenregio's. Zoals voorspeld door de Shared Syntactic Integration Resource Hypothesis werd een temporale hersenregio gevonden die wel bij taal betrokken is maar niet bij muziek. In hoofdstuk 6 werd een nieuw musicaliteitstoets geëvalueerd als een mogelijke eerste stap om verschillen te onderzoeken in hoe mensen muziek verwerken. De toets toont goede psychometrische eigenschappen. Ook na mijn onderzoek blijven er natuurlijk open vragen over hoe muziek en taal verwerkt worden. Maar het lijkt duidelijk dat taal en muziek niet helemaal gescheiden worden gehouden in de hersenen.

Zusammenfassung

Was geschieht im Gehirn wenn wir Sprache und Musik lesen oder hören? Das ist die zentrale Frage meiner Dissertation. Laut der sogenannten Shared Syntactic Integration Resource Hypothesis wird die Struktur von Sprach- und Musikstimuli teilweise durch die gleichen neuronalen Systeme verarbeitet. In Kapitel 2 beschäftigte ich mich daher mit der Frage, ob diese gemeinsamen Sprach-Musik-Systeme tatsächlich im Gehirn realisiert sind. Um dies zu untersuchen, sollten die Studienteilnehmer Sätze lesen während sie Musik hörten. Ihre Aufgabe bestand darin, die musikalischen Stimuli auf ihr harmonisches Ende hin zu bewerten. Das Ergebnis zeigt, dass syntaktisch (strukturell) schwierige Sätze die Verarbeitung musikalischer Harmonie verändert, so als ob syntaktisch schwierigee Sätze die gemeinsamen Hirnsysteme überladen, wodurch die Musik schlechter verarbeitet wird. Die Shared Syntactic Integration Resource Hypothesis prädiziert, dass diese gleichzeitige Verarbeitung von sowohl Musik- als auch Sprachstruktur in präfrontalen Hirnregionen geschieht. In Kapitel 3 präsentiere ich Evidenz für diese Hypothese, jedoch bleiben die zeitlichen Eigenschaften der neuronalen Prozesse dieser Gehirnregion undeutlich (Kapitel 4). In Kapitel 5 untersuchte ich sowohl die Hirnregionen, die *gemeinsam* Sprache und Musik verarbeiten, als auch jene, die eine Präferenz für nur eine Modalität zeigen. Wie durch die Shared Syntactic Integration Resource Hypothesis vorhergesagt, konnten wir eine Region im Temporallappen identifizieren, die lediglich bei der Sprachverarbeitung involviert ist, nicht jedoch bei der Musikverarbeitung. In Kapitel 6 wurde ein neuer Musikalitätstest evaluiert, um inter-individuelle Unterschiede in der Musikverarbeitung untersuchen zu können. Trotz weiterer offener Fragen zur Musik- und Sprachverarbeitung, zeigen die Ergebnisse dieser Dissertation, dass Sprache und Musik nicht völlig getrennt voneinander verarbeitet werden.

Acknowledgements

As opposed to every experimental chapter in this thesis, the thesis itself is only authored by a single person: me. This should by no means be taken to mean that I could have done the work described herein without the generous support of a great number of people, some of whom will be mentioned below.

My great thanks to **Peter Hagoort** and **Roel Willems** for supervising me during more than four years of research in a field for which there was very little local expertise. Looking back, it seems quite incredible that you trusted me to succeed in this project even though I had no formal expertise in music and there was none available in Nijmegen. That we nonetheless pulled it off is in no small part thanks to our international collaborators **Ani Patel** and **Jason Rosenberg**. Moreover, when expertise in singing and in electro-physiological brain data analysis was needed, **Jan-Mathijs Schoffelen** was very generous with his time and energy. Many thanks to my co-authors.

A second, far more numerous, group of people contributed to the research reported in this thesis by generating experimental data. A total of more than 450 participants engaged in the music and language tasks of chapters 2 to 6. Their anonymity should not be confused with lack of importance. Of course, the path to a PhD begins before the first participant is tested. I am grateful to **Nicolas Battich**, **Kacper Wieczorek**, **Jamie Hillis**, **Lars Muckli**, and especially **Christoph Scheepers** for supporting my first steps on the path to becoming a researcher while I was still in Glasgow. Once in the Netherlands, **Jelle Zuidema**, **Raquel Fernández**, **Richard Ridderinkhof**, and **Dilene van Campen** provided the supervision necessary for developing my skills. My Master's degree at the University of Amsterdam was an intellectually very dynamic time thanks to an extraordinary group of students put together by **Silke van Beekum** which included more people than I can name with among them **Tom Pfeffer**, **Wouter Boekel**, **Katerina Georgopoulou**, **Katja Abramova**, **Hielke Prins**, **Anna Gerlicher**, **Johanna Kostka**, **Rudy van den Brink**, **William Steel**, **Dimka Drewczynski**, **Jakob Andree**, **Stella Björk**, **Kasper Kok**, **Rony Halevy**, **Thijs van Vuure**, **Bettina Gro Sørensen**, **Christine Wilhelm**, **William Johnson**, and many others. I would like to single out **Jonas van Nijnatten** in this group. It was your generous support of my initially very poor Dutch language skills which facilitated my psycholinguistic research in particular and my absorption in Dutch culture in general. I am greatly indebted to you.

Afterwards, the Donders Institute in Nijmegen provided a social network facilitating excellent research papers which I am very grateful for. This network would not exist without the social initiative and energy of **Sean Fallon**. I greatly enjoyed my time as a member of the Donders Wonders blogging team who improved my writing (**Susanne Vogel, Jeanette Mostert**, **Ruud Berkers**, **Winke Francx**, **Piet Schipper**, **Alina Lartseva**, **Romy Bakker**, **Julian Tramper**, **Jeroen van Baar**, **Annelies van Nuland**, and especially **Lieneke Janssen**). Moreover, I am grateful to have met **Tom Marshall**, **Brónagh McCoy**, **Matthias Ekman**, **Ruben van Bergen**, **Iske Bakker**, **Lorijn Zaadnoordijk**, **Tim van Mourik**, **Jeff Martin**, **Linda de Voogd**, **Stephen Whitmarsh**, **Johanna Zumer**, **Maarten Leenders**, **Eelke Spaak**, **Jolien Fancken**, **Anke Marit Albers**, **Roemer van der Meij**, **Larry O'Dwyer**, **David Neville**, **Ili Ma**, **Jim Herring**, **Robert Oostenveld**, **Remco Bastiaannet**, **Pim Mostert**, **Miriam de Boer**, **Simon Jan Hazenberg**, **Daan van Rooij**, **Loek Brinkman**, and many others through the Donders Institute. The administrative and technical support staff at the Donders are a surprisingly effective and agreeable bunch and deserve more recognition as science facilitators.

I am furthermore thankful to the Max Planck Institute for Psycholinguistics for providing me with the financial means to do my research and for getting me in touch with **Suzanne Jongman**, **Lisa Morano**, **Elliott Hoey**, **Rick Janssen**, **Amie Fairs**, **Annika Nijveld**, **Gerardo Ortega**, **Manu Schuetze**, **Elliot Sollis**, **Mark Dingemanse**, **Nicco Usberti**, **Will Schuerman**, and many others. Peter Hagoort's Neurobiology of Language group was also a great audience for my many presentations. Thanks to **Irina Simanova**, **Franziska Hartung**, **Gwilym Lockwood**, **Annika Hultén**, **Ashley Lewis**, **Julia Uddén**, **Monique Flecken**, **Evelien Heyselaar**, **Dan Sharoh**, **Matthias Franken**, **David Peeters**, **Anne Kösem**, **Diana Dimitrova**, **Dan Acheson**, **Hartmut Fitz**, **Tineke Snijders**, **Kirsten Weber**, **Atsuko Takashima**, **Salomi Asaridou**, **Katrien Segaert**, **Lotte Schoot**, **Nietzsche Lam**, **Flora Vanlangendonck**, **Alexis Adelman**, **Geeertje van Bergen**, **Ina Grevel**, **Tildie Stijns**, and many others.

While it is certainly true that scientific skills, honed and supported by the aforementioned people, are necessary to complete a PhD, the last four years were also a time of great emotional challenges relating to the daily pursuit of scientific truths against a backdrop of disbelief, doubt, and frustration. Without the emotional support of my family and friends, I could not have mustered the will to finish this PhD trajectory. Therefore, special thanks goes out to my family (**Elke Lampka**, **Matthias Kunert**, **the late Wolfgang Lampka**, **Ulrike Huke**, **Sebastian Kunert and family**, **Juliane Kunert and partner**) who unwaveringly believed in me completing the PhD. However, it was the great people I met in Nijmegen who provided the daily social scaffold which would pick me up when the PhD would get me down. I would like to particularly mention the **Nijmeegs Studentenkoor Alphons Diepenbrock** who not only reminded me of the relevance of my research but who most of all allowed me to regain my focus when things looked insurmountable. Moreover, my friends and flatmates (**Paul Bickmann**, **Elisabeth Wiersma**, **Jan Hermsen**, **Laura Stoop**, **Irene van Ommering**, **Annemieke ter Telgte**, **Edwin Mensink**, **Felix Frehé**) never seemed to mind when I vented my frustrations or needed mindless diversions for which I am immensely grateful. There are more, perhaps innumerable people without whose friendship I could not have finished this PhD. It is impossible to name them all, but the names which come to mind spontaneously are **Annabel Nijhof**, **Dalya Samur**, **Bastien Boutonnet, Keil MacLachlan**, **Bartosz Majchrzycki**, **Nelly Van Oost**, **Aurélie Hiez**, **Simon Englebert**, and **Henry Gifford**.

I could not have done it without you all, including many unnamed people. Please also feel proud of what I achieved with your generous help. While I am the sole PhD author and I alone take responsibility of its content, you all made it possible. Great thanks for that.

Curriculum Vitae

Richard Kunert (1985, Halberstadt, Germany) attended school in Halberstadt (Gymnasium Martineum) and Arras (Lycée Gambetta, France). He obtained his high school diploma (Abitur) in 2005, receiving the highest overall grade (1,0). Following national service in West Lothian (United Kingdom), he started an undergraduate degree in Psychology at the University of Glasgow (United Kingdom) in 2006. The German National Academic Foundation (Studienstiftung des Deutschen Volkes) supported him from 2009 onwards. The Wellcome Trust supported his thesis research under the supervision of Christoph Scheepers through a Vacation Scholarship for Summer Research.

Having graduated in Psychology with the highest grade (Honours of the first class), he pursued a Master's degree in Brain and Cognitive Science at the University of Amsterdam (The Netherlands). The Dutch Organisation for Internationalisation in Education (Nuffic) supported him from 2011 onwards through a Huygens scholarship. Richard received a Master of Science degree with the highest distinction (cum laude) in 2013 and immediately started working at the Max Planck Institute for Psycholinguistics in Nijmegen (The Netherlands) and the Radboud University (Donders Institute) in Nijmegen (The Netherlands), funded by an MPI PhD grant, under the supervision of Roel Willems and Peter Hagoort. Together with Suzanne Jongman, he was awarded the Interdisciplinary Innovation Grant by the MPI for Psycholinguistics in 2014. In 2015, Richard worked as a junior policy officer at the Dutch Science Funding Organisation (NWO) in The Hague (The Netherlands) before returning to Nijmegen to finish his PhD.

Since 2012 he contributes to the communication of academic research to the general public via blogs (Brain's Idea, Donders Wonders), online articles (Die Zeit, aeon magazine), and presentations to lay people. Moreover, Richard attempted to engage audiences' neural music processing network by singing in the Nijmegen student choir Alphons Diepenbrock since 2012. He hopes to find harmony in the future.

Publications

Kunert, R., Roel M. Willems, Jason C. Rosenberg, Aniruddh D. Patel, Peter Hagoort (submitted). Structural processing of music and language: imaging domain-specific neural resources.

Campen, A.D. van, **Kunert, R.**, Wildenberg, W. van den, Ridderinkhof, K.R. (submitted). Repetitive TMS over IFC impairs suppression but not expression of action impulses during action conflict.

Kunert, R., Jongman, S.R. (accepted). Entrainment to an auditory signal: is attention involved? *Journal of Experimental Psychology: General*.

Kunert, R. (2016). Internal conceptual replications do not increase independent replication success. *Psychonomic Bulletin and Review, 23*, 1631. doi: 10.3758/s13423-016-1030-9

Kunert, R., Willems, R.M., Hagoort, P. (2016). An independent psychometric evaluation of the PROMS measure of music perception skills. *PLoS ONE, 11* (7), e0159103. doi:10.1371/journal.pone.0159103.

Kunert, R., Willems, R.M., Hagoort, P. (2016). Language influences music perception - effects of shared syntactic integration resources beyond attention. *Royal Society Open Science, 3*, 150685. doi: 10.1098/rsos.150685

Kunert, R., Willems, R.M., Casasanto, D., Patel, A.D., Hagoort, P. (2015). Music and language syntax interact in Broca's area: an fMRI study. *PLoS ONE, 10* (11), e0141069. doi:10.1371/journal.pone.0141069

Kunert, R., Slevc, L.R. (2015). A commentary on “Neural overlap in processing music and speech” (Peretz et al., 2015). *Frontiers in Human Neuroscience, 9*, 330. doi: 10.3389/fnhum.2015.00330

Kunert, R., Scheepers, C. (2014). Speed and accuracy of dyslexic versus typical word recognition: An eye-movement investigation. *Frontiers in Psychology, 5*, 1129. doi: 10.3389/fpsyg.2014.01129

MPI Series in Psycholinguistcs

1. The electrophysiology of speaking: Investigations on the time course of semantic, syntactic, and phonological processing. *Miranda van Turennout*
2. The role of the syllable in speech production: Evidence from lexical statistics, metalinguistics, masked priming, and electromagnetic midsagittal articulography. *Niels O. Schiller*
3. Lexical access in the production of ellipsis and pronouns. *Bernadette M. Schmitt*
4. The open-/closed-class distinction in spoken-word recognition. *Alette Haveman*
5. The acquisition of phonetic categories in young infants: A self-organising artificial neural network approach. *Kay Behnke*
6. Gesture and speech production. *Jan-Peter de Ruiter*
7. Comparative intonational phonology: English and German. *Esther Grabe*
8. Finiteness in adult and child German. *Ingeborg Lasser*
9. Language input for word discovery. *Joost van de Weijer*
10. Inherent complement verbs revisited: Towards an understanding of argument structure in Ewe. *James Essegbey*
11. Producing past and plural inflections. *Dirk Janssen*
12. Valence and transitivity in Saliba: An Oceanic language of Papua New Guinea. *Anna Margetts*
13. From speech to words. *Arie van der Lugt*
14. Simple and complex verbs in Jaminjung: A study of event categorisation in an Australian language. *Eva Schultze-Berndt*
15. Interpreting indefinites: An experimental study of children's language comprehension. *Irene Krämer*
16. Language-specific listening: The case of phonetic sequences. *Andrea Weber*
17. Moving eyes and naming objects. *Femke van der Meulen*
18. Analogy in morphology: The selection of linking elements in Dutch compounds. *Andrea Krott*
19. Morphology in speech comprehension. *Kerstin Mauth*
20. Morphological families in the mental lexicon. *Nivja H. de Jong*
21. Fixed expressions and the production of idioms. *Simone A. Sprenger*

22. The grammatical coding of postural semantics in Goemai (a West Chadic language of Nigeria). *Birgit Hellwig*
23. Paradigmatic structures in morphological processing: Computational and cross-linguistic experimental studies. *Fermín Moscoso del Prado Martín*
24. Contextual influences on spoken-word processing: An electrophysiological approach. *Daniëlle van den Brink*
25. Perceptual relevance of prevoicing in Dutch. *Petra M. van Alphen*
26. Syllables in speech production: Effects of syllable preparation and syllable frequency. *Joana Cholin*
27. Producing complex spoken numerals for time and space. *Marjolein Meeuwissen*
28. Morphology in auditory lexical processing: Sensitivity to fine phonetic detail and insensitivity to suffix reduction. *Rachèl J. J. K. Kemps*
29. At the same time...: The expression of simultaneity in learner varieties. *Barbara Schmiedtová*
30. A grammar of Jalonke argument structure. *Friederike Lüpke*
31. Agrammatic comprehension: An electrophysiological approach. *Marlies Wassenaar*
32. The structure and use of shape-based noun classes in Miraña (North West Amazon). *Frank Seifart*
33. Prosodically-conditioned detail in the recognition of spoken words. *Anne Pier Salverda*
34. Phonetic and lexical processing in a second language. *Mirjam Broersma*
35. Retrieving semantic and syntactic word properties. *Oliver Müller*
36. Lexically-guided perceptual learning in speech processing. *Frank Eisner*
37. Sensitivity to detailed acoustic information in word recognition. *Keren B. Shatzman*
38. The relationship between spoken word production and comprehension. *Rebecca Özdemir*
39. Disfluency: Interrupting speech and gesture. *Mandana Seyfeddinipur*
40. The acquisition of phonological structure: Distinguishing contrastive from non-contrastive variation. *Christiane Dietrich*
41. Cognitive cladistics and the relativity of spatial cognition. *Daniel B.M. Haun*
42. The acquisition of auditory categories. *Martijn Goudbeek*
43. Affix reduction in spoken Dutch. *Mark Pluymaekers*

44. Continuous-speech segmentation at the beginning of language acquisition: Electrophysiological evidence. *Valesca Kooijman*
45. Space and iconicity in German Sign Language (DGS). *Pamela Perniss*
46. On the production of morphologically complex words with special attention to effects of frequency. *Heidrun Bien*
47. Crosslinguistic influence in first and second languages: Convergence in speech and gesture. *Amanda Brown*
48. The acquisition of verb compounding in Mandarin Chinese. *Jidong Chen*
49. Phoneme inventories and patterns of speech sound perception. *Anita Wagner*
50. Lexical processing of morphologically complex words: An information-theoretical perspective. *Victor Kuperman*
51. A grammar of Savosavo, a Papuan language of the Solomon Islands. *Claudia Wegener*
52. Prosodic structure in speech production and perception. *Claudia Kuzla*
53. The acquisition of finiteness by Turkish learners of German and Turkish learners of French: Investigating knowledge of forms and functions in production and comprehension. *Sarah Schimke*
54. Studies on intonation and information structure in child and adult German. *Laura de Ruiter*
55. Processing the fine temporal structure of spoken words. *Eva Reinisch*
56. Semantics and (ir)regular inflection in morphological processing. *Wieke Tabak*
57. Processing strongly reduced forms in casual speech. *Susanne Brouwer*
58. Ambiguous pronoun resolution in L1 and L2 German and Dutch. *Miriam Ellert*
59. Lexical interactions in non-native speech comprehension: Evidence from electro-encephalography, eye-tracking, and functional magnetic resonance imaging. *Ian FitzPatrick*
60. Processing casual speech in native and non-native language. *Annelie Tuinman*
61. Split intransitivity in Rotokas, a Papuan language of Bougainville. *Stuart Robinson*
62. Evidentiality and intersubjectivity in Yurakaré: An interactional account. *Sonja Gipper*
63. The influence of information structure on language comprehension: A neurocognitive perspective. *Lin Wang*
64. The meaning and use of ideophones in Siwu. *Mark Dingemanse*

65. The role of acoustic detail and context in the comprehension of reduced pronunciation variants. *Marco van de Ven*
66. Speech reduction in spontaneous French and Spanish. *Francisco Torreira*
67. The relevance of early word recognition: Insights from the infant brain. *Caroline Junge*
68. Adjusting to different speakers: Extrinsic normalization in vowel perception. *Matthias J. Sjerps*
69. Structuring language. Contributions to the neurocognition of syntax. *Katrien R. Segaert*
70. Infants' appreciation of others' mental states in prelinguistic communication: A second person approach to mindreading. *Birgit Knudsen*
71. Gaze behavior in face-to-face interaction. *Federico Rossano*
72. Sign-spatiality in Kata Kolok: how a village sign language of Bali inscribes its signing space**.** *Conny de Vos*
73. Who is talking? Behavioural and neural evidence for norm-based coding in voice identity learning. *Attila Andics*
74. Lexical processing of foreign-accented speech: Rapid and flexible adaptation. *Marijt Witteman*
75. The use of deictic versus representational gestures in infancy. *Daniel Puccini*
76. Territories of knowledge in Japanese conversation. *Kaoru Hayano*
77. Family and neighbourhood relations in the mental lexicon: A cross-language perspective. *Kimberley Mulder*
78. Contributions of executive control to individual differences in word production. *Zeshu Shao*
79. Hearing speech and seeing speech: Perceptual adjustments in auditory-visual processing. *Patrick van der Zande*
80. High pitches and thick voices: The role of language in space-pitch associations. *Sarah Dolscheid*
81. Seeing what's next: Processing and anticipating language referring to objects. *Joost Rommers*
82. Mental representation and processing of reduced words in casual speech. *Iris Hanique*
83. The many ways listeners adapt to reductions in casual speech. *Katja Poellmann*
84. Contrasting opposite polarity in Germanic and Romance languages: Verum Focus and affirmative particles in native speakers and advanced L2 learners. *Giuseppina Turco*

85. Morphological processing in younger and older people: Evidence for flexible dual-route access. *Jana Reifegerste*
86. Semantic and syntactic constraints on the production of subject-verb agreement. *Alma Veenstra*
87. The acquisition of morphophonological alternations across languages. *Helen Buckler*
88. The evolutionary dynamics of motion event encoding. *Annemarie Verkerk*
89. Rediscovering a forgotten language. *Jiyoun Choi*
90. The road to native listening: Language-general perception, language-specific input. *Sho Tsuji*
91. Infants'understanding of communication as participants and observers. *Gudmundur Bjarki Thorgrímsson*
92. Information structure in Avatime. *Saskia van Putten*
93. Switch reference in Whitesands. *Jeremy Hammond*
94. Machine learning for gesture recognition from videos. *Binyam Gebrekidan Gebre*
95. Acquisition of spatial language by signing and speaking children: a comparison of Turkish sign language (TID) and Turkish. *Beyza Sümer*
96. An ear for pitch: on the effects of experience and aptitude in processing pitch in language and music. *Salomi Savvatia Asaridou*
97. Incrementality and Flexibility in Sentence Production. *Maartje van de Velde*
98. Social learning dynamics in chimpanzees: Reflections on (nonhuman) animal culture. *Edwin van Leeuwen*
99. The request system in Italian interaction. *Giovanni Rossi*
100. Timing turns in conversation: A temporal preparation account. *Lilla Magyari*
101. Assessing birth language memory in young adoptees**.** Wencui Zhou
102. A social and neurobiological approach to pointing in speech and gesture. *David Peeters*
103. Investigating the genetic basis of reading and language skills. *Alessandro Gialluisi*
104. Conversation Electrified: The Electrophysiology of Spoken Speech Act Recognition***.*** *Rósa Signý Gisladottir*

105. Modelling Multimodal Language Processing. *Alastair Smith*

106. Predicting language in different contexts: The nature and limits of mechanisms in anticipatory language processing. *Florian Hintz*

107. Situational variation in non-native communication. *Huib Kouwenhoven*

108. Sustained attention in language production. *Suzanne Jongman*

109. Acoustic reduction in spoken-word processing: Distributional, syntactic, morphosyntatic, and orthographic effects. *Malte Viebahn*

110. Nativeness, dominance, and the flexibility of listening to spoken language. *Laurence Bruggeman*

111. Semantic specificity of perception verbs in Maniq. *Ewelina Wnuk*

112. On the identification of FOXP2 gene enhancers and their role in brain development. *Martin Becker*

113. Events in language and thought: The case of serial verb constructions in Avatime. *Rebecca Defina*

114. Deciphering common and rare genetic effects on reading ability. *Amaia Carrión Castillo*

115. Music and language comprehension in the brain. *Richard Kunert*