

Annual Review of Linguistics Music and Language

David Temperley

Eastman School of Music, University of Rochester, Rochester, New York, USA; email: dtemperley@esm.rochester.edu

Annu. Rev. Linguist. 2022. 8:153-70

The Annual Review of Linguistics is online at linguistics.annualreviews.org

https://doi.org/10.1146/annurev-linguistics-031220-121126

Copyright © 2022 by Annual Reviews. All rights reserved

ANNUAL CONNECT

- www.annualreviews.org
- Download figures
- Navigate cited references
- Keyword search
- Explore related articles
- Share via email or social media

Keywords

music-language connections, rhythm, stress, meter, syntax, repetition

Abstract

This review presents a highly selective survey of connections between music and language. I begin by considering some fundamental differences between music and language and some nonspecific similarities that may arise out of more general characteristics of human cognition and communication. I then discuss an important, specific interaction between music and language: the connection between linguistic stress and musical meter. Next, I consider several possible connections that have been widely studied but remain controversial: cross-cultural correlations between linguistic and musical rhythm, effects of musical training on linguistic abilities, and connections in cognitive processing between music and linguistic syntax. Finally, I discuss some parallels regarding the use of repetition in music and language, which until now has been a little-explored topic. "What the music I love expresses to me is not too indefinite to be put into words, but rather, too definite."

-Felix Mendelssohn¹

1. INTRODUCTION

For hundreds of years, parallels have been drawn between music and language. Most often, these parallels have been drawn from the music side: Scholars of music have looked to language for ideas about how music is (or should be) composed, performed, and perceived. From the late Middle Ages to the eighteenth century, classical rhetoric was an enduring source of fascination for music theorists; Johann Mattheson's influential *Der vollkommene Capellmeister* (1739) presented a complete theory of musical organization based on rhetorical categories such as *exordium, confutatio*, and *peroratio*. The nineteenth-century philosopher Herbert Spencer (1890) suggested that music's emotional effect arises from its resemblance to speech, an argument that still carries weight today. In recent decades, ideas about syntax and prosody have found their way into music research, as well as concepts and methods from psycholinguistics and computational linguistics.

In this review, I discuss some well-established connections between music and language, survey some recent research in this area, and add some thoughts of my own. My treatment of this potentially vast topic is necessarily highly selective. I begin by considering some fundamental differences between music and language (Section 2) and some nonspecific similarities that may arise out of more general characteristics of human cognition and communication (Section 3). In Section 4, I discuss an important, specific interaction between music and language: the connection between linguistic stress and musical meter. In Section 5, I consider several recent proposals that have been widely studied but remain controversial. Finally, in Section 6, I discuss some parallels regarding the use of repetition in music and language, which until now has been a little-explored topic. With regard to music, my focus is somewhat biased toward Western music (both classical and popular), though many of my observations apply also to non-Western traditional and popular musical styles.

In examining possible connections between music and language, I find it useful to distinguish between similarities and interactions. Similarities are exactly what the term implies and require no further general comment. Interactions are causal connections between music and language.² These could be of a historical nature (e.g., ways that a culture's language affects its music); there could also be interactions across individuals (effects of linguistic training on musical ability, or vice versa) or in cognitive processes (interference or facilitation between simultaneous linguistic and musical processing). Similarities could exist without interactions; music and language could, in principle, share certain features without there being any interaction between them at all. Likewise, interactions could occur between aspects of music and language that are quite dissimilar.

I say little in this review about neuroscientific research on music–language connections, though this has become an extremely active area (for a recent review, see Besson et al. 2017). This omission is partly due to my limited expertise in neuroscience; in addition, my sense is that this work is not yet able to contribute much to the issues of structure and mental representation that are my focus here.

¹"Das, was mir eine Musik ausspricht, die ich liebe, sind mir nicht zu unbestimmte Gedanken, um sie in Worte zu fassen, sondern zu bestimmte" [Mendelssohn 1878 (1842), p. 221].

²I use the term "interaction" here in its everyday sense, in contrast to its more technical scientific meaning (referring to a nonadditive effect of two variables on a third variable).

2. SOME BASIC DIFFERENCES

In many respects, the differences between music and language are so great that it is difficult to know how to compare them. Both are systems of human communication, and both are universal across cultures (Brown 1991, Savage et al. 2015). Beyond this, it is not obvious how to proceed. The traditional division of language into domains of phonology, syntax, and semantics (some would add morphology and pragmatics) does not apply to music in any obvious way; while several ways of doing this mapping have been proposed (Bernstein 1976, Swain 1997, Patel 2008), none has been widely accepted in musical discourse. In particular, the concept of meaning in language-in the conventional sense of propositions referring to objects, events, and actions-has no natural counterpart in music (with occasional exceptions, such as leitmotifs in Wagner's operas). Music does communicate something, but by most accounts (and there are many), that "something" is fundamentally different from linguistic meaning: It refers to other things inside itself (such as an occurrence of a theme referring to an earlier occurrence, or a dissonant note implying a subsequent resolution) (Meyer 1956); it conveys emotions (Kivy 1980), or perhaps gives rise to emotions (Huron 2006); it conveys abstract structures of inherent beauty [Hanslick 1986 (1891)], or journeys in a space of chords and keys (Lerdahl 2001b), or metaphorical depictions of embodied experience (Cox 2016), or abstract narratives whose agents shift constantly from one kind of thing to another-instruments, pitches, keys, themes (Monahan 2013). Whatever music communicates, it seems clear that it is not easily translated into language, or vice versa-not necessarily due to any vagueness or indeterminacy of meaning on music's part (as Mendelssohn reminds us in my opening quote). On the other hand, the ubiquitous practice of singing with words suggests that music's communicative content can complement and support that of language, or vice versa.

It might seem more fruitful to shift from what is communicated to <u>how</u> it is communicated; but here, too, we encounter vast differences. Music relies on contextual frameworks—meter and key—that are inferred by the listener from events and then guide the interpretation of subsequent events; nothing of the kind is present in language. (The relationship of musical meter to language is considered in more detail below.) Another distinctive aspect of music (most Western music anyway) is that it is constructed from multiple notes occurring simultaneously, usually grouping together to form multiple lines—what we call polyphony or counterpoint. This is most obvious in pieces like fugues, but it is evident to some extent in most kinds of Western music; consider a rock song composed of a vocal melody, a bass line, and a guitar part. The extent to which ordinary listeners perceive the components of polyphonic musical textures is somewhat unclear, but in advanced musical study and discourse, much of our attention is devoted to the way musical lines combine and interact. Great traditions of non-Western music, such as West African drumming and Javanese gamelan, have their own forms of polyphony, though in these cases lines are distinguished primarily by timbre rather than by pitch. This enormously important aspect of music is completely without parallel in language.

Another difference between music and language that has received some attention, but perhaps not as much as it deserves, relates to the idea of arbitrariness (Hockett 1960, Swain 1997, Temperley 2001, Fitch 2006). The functional properties of a word—its syntactic category (or features) and its meaning—cannot be inferred from the phonological form of the word using general principles: They must be learned by the language user and then stored in memory for future use. (If the word is composed of multiple morphemes, the properties of the word can be inferred from the properties of its morphemes; but then arbitrariness applies to the morphemes.) There are well-known exceptions to this, such as onomatopoeia and systematicity—for example, the fact that certain vowel sounds tend to be associated cross-linguistically with large or small objects (Dingemanse et al. 2015)—but no one has suggested that word meanings can be inferred entirely from sound-meaning correlations of this kind. (If that were true, there would be no need to <u>learn</u> languages!) From a cognitive point of view, arbitrariness implies the existence of a lexicon, a library of mappings from sound patterns to functional properties; a typical human language requires thousands of such mappings. We could define the arbitrariness of a communicative system in terms of the size of its lexicon—the number of form-to-function mappings needed to use it. For example, suppose (contrary to fact) that the meanings of English words could be derived from their phonemes: In the case of the word *kill*, for example, suppose /k/ meant CAUSE, /I/ meant TO BECOME, and /l/ meant DEAD. In that case, no lexicon of words would be needed. A much smaller lexicon would still be required, containing (presumably arbitrary) mappings of phonemes to meanings; thus there would still be some arbitrariness in the system, but much less.

What role does arbitrariness play in music? Here it is helpful to think about the properties and functions of musical elements quite generally, without getting bogged down in notions of meaning or syntax. Consider pitches. A pitch has height (frequency); this plays a role in its expressive functions (higher pitch implies higher arousal; more on this below). No lexicon is needed here, just the general principle relating pitch height to arousal. The height of a pitch also determines its relationship to the key (its scale degree), which in turn gives it a certain degree of stability or fit. At most, this requires a tiny lexicon mapping the 12 scale degrees (in Western music at least) onto levels of stability. Some have suggested that even these relationships are not arbitrary but arise out of general psychoacoustic principles (Terhardt 1974, Parncutt 1989). Pitches also participate in forming harmonies; this follows very simple, general (and again, acoustically motivated) principles (e.g., every pitch is a member of the major triad whose root is a major third below). Moving to higher levels, chords have levels of stability depending on their relationship to the key, and fall into larger "functional" categories (tonic, dominant, pre-dominant) that determine their normative use in harmonic progressions. Here again, a very small lexicon might be needed, though in large part the functional properties of chords follow from general principles; for example, two chords that share two pitches tend to be similar in function. (Again, psychoacoustic principles may also be involved.) Constraints on the combination of melodic lines can also be described in terms of general rules; indeed, it has been argued that rules of counterpoint themselves arise from still more general principles of auditory perception (Huron 2016). Repeated patterns of pitches-themes and motives-might serve certain structural or expressive functions in a piece, but any lexicon of these would be unique to a particular piece and would contain only a handful of items. In short, the level of arbitrariness in music-the size of the lexicon required to understand it-is microscopic compared with that of language.

One might wonder if there is anything in music comparable to words—perhaps small groups of pitches or chords that are used frequently across pieces. (This is distinct from motives—patterns that recur within a piece of music and are unique to that piece.) Several scholars have made proposals along these lines, relating to styles such as Renaissance music (Schubert & Cumming 2015), mid-eighteenth-century classical music (Gjerdingen 2007), jazz (Rolland 1999), and Indian classical music (Lipiczky 1985). But none of these authors has said much about the communicative functions of these patterns—what they convey to the listener, if anything, and why it is necessary or even helpful for the listener to identify them. It is also hard to know whether these patterns are genuine schemata in the minds of composers and listeners or whether they simply arise by chance as clusters of highly probable events.³

³Some researchers have explored this issue using concepts borrowed from information theory and computational linguistics, such as entropy and *n*-gram models (Rolland 1999, Conklin & Anagnostopoulou 2001, Pearce & Wiggins 2004). Such studies show that a note in a melody can be predicted better from the previous

3. NONSPECIFIC SIMILARITIES

It is often said that music and language are both "hierarchical" in that they involve larger (or higher-level) elements dominating smaller (or lower-level) ones. In fact, both music and language involve multiple hierarchies: syntactic and prosodic structures in language; and meter, phrase structure, and pitch organization in music (more on this below). But all kinds of things are hierarchical in this broad sense-often in some natural way, or at least, in the way we mentally represent them; think of a human body, a baseball game, a wedding, or a university. So this might be better regarded as a general characteristic of human cognition than as a special feature of music and language. One rather close parallel between musical and linguistic hierarchies is between prosodic structure in language [the grouping of linguistic sounds into syllables, feet, and phrases (Shattuck-Hufnagel & Turk 1996)] and what is called grouping structure in music [the grouping of notes into motives, phrases, and sections (Lerdahl & Jackendoff 1983; see also Heffner & Slevc 2015 for discussion)]. In both cases, temporal proximity (pauses or lengthening at segment boundaries) is an important perceptual cue; this is not unique to music and language but reflects a general principle of gestalt perception (Bregman 1990). In music there are also other important cues to grouping-notably, similarity between sounds in pitch, timbre, and other parameters-that play less of a role in prosodic structure. Some other similarities and differences between musical and linguistic hierarchies are discussed below.

Other similarities between music and language may arise simply because both are systems of (primarily auditory) communication. One relates to what I have elsewhere called "communicative pressure" (Temperley 2004). For something to function as a system of communication, the information that it is designed to communicate has to be effectively conveyed (almost a tautology when stated in that way). This appears to be an important shaping force on both music and language. It may explain previously observed "trading relationships," such as the fact that languages in which grammatical case is not indicated by inflection tend to indicate it with fixed word order, or the fact that musical styles that allow flexibility in tempo (i.e., fluctuation in the speed of the underlying meter) tend to allow less syncopation (conflict between acoustic accents and metrical structure). [Swain (1997) has offered an insightful discussion of this issue.] We would expect communicative pressure to operate not only in music and language but in any communicative system-in animal signaling systems or in nonlinguistic human systems such as computer languages. Another example that comes to mind, much discussed in recent years, is statistical learning-people's ability to use probabilistic co-occurrences between sequential elements as a means of grouping them together. This has been observed not only in language (Saffran et al. 1996) and music (Saffran et al. 1999) but also in visual patterns (Kirkham et al. 2002); thus it seems to represent a basic aspect of human cognition.

One important similarity between music and language concerns the expression of emotion. In both music and language, pitch, loudness, and temporal density are used for expressive purposes, and they tend to be used in similar ways in the two domains. Evidence for this comes from numerous experimental studies in which people were asked to speak sentences or perform melodies in such a way as to convey a particular emotion (for a survey, see Juslin & Laukka 2003). In both music and language, higher levels of pitch, temporal density, and loudness convey emotions that are high in "energy" or "arousal" (Russell 1980), such as happiness or anger. Continuous

few notes than from just the preceding note. This may point to the existence of collocations in music, though there are other possible explanations for it as well; a sequence of notes may help to predict the following note simply because it provides information about more general features of the context, such as the style, key, or underlying harmony (I return to this issue in Section 5.3).

variations in loudness and tempo are used for expressive purposes in both speech and music. Pitch and temporal density are also manipulated in music in a categorical fashion, with similar expressive effects: Higher pitches and faster rhythmic values convey higher arousal (Gabrielsson & Lindström 2001).

We might ask why increases in temporal density, loudness, and pitch convey a connotation of energy and arousal; one reason may be that such increases actually require more energy on the part of the producer. In that sense, we might expect to find similar phenomena in other modes of communication; in dance, for example, we would expect more rapid, extreme, or strenuous motions to carry a connotation of increased energy (I am unable to find any research on this issue). We would also expect similar mappings between acoustic dimensions and communicative function in animal communication, as indeed has been observed (Briefer 2012, Filippi 2016). Thus, this parallel in emotional expression between music and language may simply arise out of their being auditory modes of communication—similar to parallels discussed earlier in this section.

4. A SPECIFIC CONNECTION: SYLLABIC STRESS AND MUSICAL METER

When text is set to music (or when words are written for an existing melody), there is a strong preference for stressed syllables to be placed on metrically strong beats.⁴ Both "stressed syllable" and "strong beat" are to be understood in relative terms; the important point is that the relative stress of nearby syllables (especially adjacent ones) should be matched by the relative strength of their corresponding beats. An example is shown in **Figure 1***a*: *Take* and *out* are both more stressed than *me* and are on stronger metrical positions; *ball* is more stressed than *game* and is metrically stronger. (The metrical grids used in **Figure 1** to represent both stress and meter are discussed in more detail below.) At higher levels, the pressure for alignment between stress and meter is weaker; *out* is on a weaker beat than *take* despite being more stressed. The importance of stressmeter alignment can be demonstrated informally by a well-known trick: Sing "Take Me Out to the Ball Game" with the syllables shifted so that the second syllable is aligned with the first note, the third syllable with the second note, as so on, as shown in **Figure 1***b*. The awkwardness of this misalignment is immediately evident.

The need to align stress and meter has long been recognized as a basic principle in the composition of vocal music; it has been formalized theoretically by Liberman (1975) and Halle & Lerdahl (1993). Corpus studies of vocal music have shown a strong correlation between the stress levels of syllables and their metrical strength (Palmer & Kelly 1992, Temperley & Temperley 2013). In recent popular music, a stressed syllable may be shifted to occur on the weak beat just before a strong one; however, this occurs only in restricted and quite systematic ways (Temperley 1999b, Tan et al. 2019). Even in French, a language that supposedly lacks lexical stress, content words tend to be on stronger beats than function words, and the final syllables of content words tend to be metrically stronger than nonfinal syllables, though these correlations are much weaker than in English (Temperley & Temperley 2013). There also is experimental evidence that the alignment of stress and meter facilitates performance, perception, and memory. Hayes & Kaun (1996) found that when people are given a phrase and told to chant it in a rhythmically regular way, they tend to space the stressed syllables evenly; when a phrase of text is sung to a melody, both the words and the notes are sung more accurately when the two are metrically aligned (A. Reed, B. Maxwell & D. Temperley, manuscript in preparation). Children remember text phrases better when they are

⁴This discussion does not apply to languages that lack stress, such as Japanese and Korean.



Figure 1

(*a*) The first phrase of "Take Me Out to the Ball Game." Metrical grids are used to represent musical meter (above the musical notation) and syllabic stress (below the lyrics); the accentuation of a note or syllable is indicated by its height in the grid (the number of x's above or below it). In panel a (from the original song), meter and stress are aligned, meaning that syllables with higher stress are placed at metrically accented positions; in panel b, the syllables have been shifted so that the meter and stress are misaligned.

initially presented with good textsetting (with stressed syllables falling on strong beats) (Gingold & Abravanel 1987).

The tendency toward alignment between stressed syllables and strong beats is, in my view, the most important interaction between language and music; it represents a direct connection between musical and linguistic mental representations, specific to those two domains. We might say that stress in language and meter in music are cognitively coupled: There is pressure for them to align in specific situations. Functionally, both linguistic stress and musical meter serve to guide attention to certain points of time in the auditory input. In the case of music, the pitch of a note can be judged more accurately if it occurs in a metrically predictable position—that is, one that continues a previous pattern of temporal intervals (Jones et al. 2002). In language, phonemes are detected more quickly in stressed syllables than in unstressed ones, suggesting enhanced attention for stressed syllables (Cutler & Foss 1977).

Structurally, both syllabic stress and musical meter consist of multileveled patterns of accentuation, commonly represented with metrical grids as shown in **Figure 1** (Lerdahl & Jackendoff 1983, Hayes 1995). In language, the elements represented in the grid are syllables; in music, they are beats—subjectively accented points in time that are inferred from the music and then maintained in the mind of the listener. Beats often coincide with notes but need not do so. In music (most kinds of music anyway), there is a strong tendency toward equal spacing of beats at each level. In languages like English, by contrast, stress patterns within words are stored in the lexicon (though they are to some extent constrained by phonology) and are shaped at higher levels by syntactic and pragmatic structures; the metrical grid arises directly from these stress patterns and may be highly irregular. (In **Figure 1**, there is one completely unstressed syllable between *take* and *out*, two between *out* and *ball*, and zero between *ball* and *game*.) There is a slight tendency toward isochrony in language—equal spacing of stresses; in English, this is seen most clearly in cases where the addition of a suffix to a stem causes a shift in the stress pattern of the stem (e.g., *a-<u>dore</u> / <u>a-dor-a-tion</u>). In certain situations, the preference for isochrony may affect the way that words are perceived or produced (Liberman & Prince 1977, Kelly & Bock 1988, Dilley & McAuley 2008, Temperley 2009), but the tendency is far weaker than in music. Claims that languages reflect equal spacing of stresses or syllables have been largely discredited (Roach 1982, Dauer 1983).*

In pronunciation, stressed syllables are normally marked acoustically by increased loudness, duration, and pitch (Sluijter & Van Heuven 1996). Music also features acoustically based accents. arising from some of the same acoustic cues associated with linguistic stress-duration and loudness (as well as other factors, such as harmonic change); the status of pitch as a source of musical accent is controversial (Sloboda 1983, Huron & Royal 1996). Lerdahl & Jackendoff (1983) introduced the vital distinction between these "phenomenal accents" and "metrical accents" arising from a note's position in the metrical grid. It is from the phenomenal accents that the meter of a piece is initially inferred; but once established, the phenomenal accents need not constantly reinforce the meter and indeed may conflict with it—this is the phenomenon of syncopation, which has no counterpart in language.⁵ Crucially, it is the metrical accents, not the phenomenal ones, that must align with stressed syllables: A trochaic (stressed-unstressed) syllabic pattern will seem incorrect when set to a metrically weak-strong note pattern, no matter how loud or long the first note is. (In Figure 1b, singing out louder than to does not reduce the awkwardness.) The metrical placement of a note greatly affects how it is mentally represented. A pattern of notes that is repeated in an acoustically identical way but in a different metrical context can seem like a completely different melody (Sloboda 1983, Povel & Essens 1985, Acevedo et al. 2014)-a phenomenon that, again, has no linguistic parallel.

5. CONTROVERSIAL PROPOSALS

In this section I discuss several recent proposals for connections between music and language. All of them concern interactions of various kinds: across cultures (Section 5.1), across individuals (Section 5.2), or in cognitive processing (Section 5.3). All are interesting and deserve serious consideration, but in all cases, the evidence is mixed.

5.1. Rhythmic Variability in Music and Language

It is well known that languages differ in variability with regard to syllable length; for example, English has higher variability in this regard than French (Grabe & Low 2002, Ramus 2002). This is often measured using the normalized pairwise variability index (nPVI), which measures variability between pairs of adjacent syllables. Using the nPVI, Patel & Daniele (2003) found that English instrumental melodies from the years around 1900 tend to have higher durational variability than French ones, and suggested that linguistic influence might be responsible for this difference. A number of subsequent studies have been done in this area, with mixed results. McGowan & Levitt (2011) examined three dialects of English (Scottish, Irish, and Kentucky) and found parallels between their syllabic and melodic nPVIs; Lee et al. (2017) found a similar parallel in Southern British English and Multicultural London English; Hannon (2009) confirmed Patel & Daniele's

⁵A middle ground between music and language in this respect is poetry, where (in some poetic traditions at least) there is a regular pattern of stressed and unstressed syllables; this may sometimes allow something analogous to syncopation, where syllabic stress levels momentarily conflict with the underlying metrical pattern (Lerdahl 2001a).

(2003) findings in a study of French and English children's songs. However, VanHandel & Song (2010) found no difference in nPVI between German and French vocal melodies (though German is higher than French in syllabic nPVI); Temperley (2017) found higher nPVI in French and Italian vocal melodies (both low-nPVI languages) than in German and English ones (both high-nPVI languages); and Vukovics & Shanahan (2020) examined instrumental themes in the same four nations (England, France, Germany, and Italy), and again found slightly higher melodic nPVI for the low-nPVI languages. Efforts to tie historical changes in melodic nPVI to the rise of nationalism, on the reasoning that language may have greater effects on music in more nationalistic periods, have also had discouraging results (Daniele & Patel 2013, VanHandel 2017, Vukovics & Shanahan 2020). On balance, these studies do not suggest any causal relationship between syllabic and melodic durational variability. In a slightly different vein, Temperley & Temperley (2011) examined the Scotch Snap, a rhythmic pattern featuring a short note on a strong beat followed by a longer off-beat note, which is common in English and Scottish melodies but rare in German and Italian ones; the authors hypothesized that this might be due to the high frequency of very short stressed syllables in English compared with German and Italian (which the authors verified using corpus data). Like Patel & Daniele's (2003) original finding, this is a suggestive result, but inconclusive with regard to the causal relationship between music and language. In all studies of this type, each nation really represents only a single data point; we should be cautious about drawing strong conclusions about general relationships from only a few data points.

5.2. Interactions Between Music and Language in Training and Ability

An issue of great interest in recent music psychology has been the effect of musical training on other cognitive abilities. Some of this research has specifically examined these effects with regard to verbal (language-related) skills. With studies of this kind, of course, one must distinguish correlation from causation: If children with musical training show better verbal abilities, this may be due to environmental (or even genetic) factors that affect both variables rather than to a causal connection between them. Gordon et al. (2015) presented a meta-analysis of 12 studies examining the effects of musical training on verbal skills; they included only controlled studies, in which one group had been given musical training and another group had not. They also considered only studies with behavioral measures of verbal ability, since neurological measures are difficult to interpret in this regard. Overall, the meta-analysis found no significant effect of musical training on reading ability. They also examined the effect of musical training on phonological skills; there was a significant benefit on rhyming ability but not on other phonological skills. On balance, then, existing research suggests there may be a small and rather specific effect of musical training on verbal ability.

Numerous studies have suggested that musical training enhances the perception of speech in noisy environments (e.g., Parbery-Clark et al. 2009). This has been attributed to effects of higherlevel brain centers on the representation of the sound signal in the brain stem. However, other studies have failed to replicate these results. Madsen et al. (2019) suggested that positive effects of musical training obtained in some "speech-in-noise" studies may be due to conditions that lack ecological validity, with regard to the speech materials used (words chosen from a limited set), the relative positions of the speaker and listener, and the lack of reverberation. Using more ecologically valid conditions, Madsen et al. (2019) found no difference between musicians and nonmusicians in speech perception ability.

A possible interaction between music and language in the other direction concerns tone languages. In tone languages, pitch is used in a categorical fashion to distinguish phonemes and syllables, similar to the categorical treatment of pitch in music. Given this similarity, the question arises as to whether knowledge of a tone language influences musical ability. Deutsch (2006) has hypothesized that the prevalence of tone languages in East Asia might explain why people in that region possess absolute pitch (the ability to name pitches that one hears, e.g., C or D-flat) at much higher rates than in other regions. However, further research has cast doubt on that finding, suggesting that this difference has more to do with early musical training than with linguistic background (Schellenberg & Trehub 2008). Other studies suggest that a tone-language background may confer more general advantages in pitch perception and production; tone-language speakers can more accurately imitate singing, and they show finer pitch discrimination and better memory for melodies (Pfordresher & Brown 2009, Bidelman et al. 2013).

5.3. Syntactic Structure and Music

An interesting and controversial issue is whether there is any musical counterpart to syntactic structure in language. There have been several proposals along these lines. In the influential theory of Schenker [1979 (1935)], notes elaborate other notes in a recursive fashion, forming an enormous hierarchical structure perhaps spanning many minutes. Schenker's theory is more prescriptive than descriptive in character; even advocates of the theory mostly regard it as an idealized form of hearing to which we should aspire (for discussion, see Temperley 1999a). Lerdahl & Jackendoff (1983) put forth a more formalized, psychologically oriented version of Schenker's theory. Other syntax-like, hierarchical representations of music have been proposed by Rohrmeier (2011) and Granroth-Wilding & Steedman (2014). Inspired by the similarity between Lerdahl & Jackendoff's (1983) structures and linguistic dependency structures, Patel (2003, 2008) proposed a deep cognitive connection between the two: He suggested that linguistic and musical syntax involve independent mental representations but that shared cognitive resources are used to construct them. Evidence for Patel's theory comes from the fact that musical (e.g., harmonic) anomalies are more disruptive to syntactic than to semantic processing (Slevc et al. 2009) and that people with deficits in syntactic processing also show musical deficits (Patel 2005). More recent studies have cast doubt on this theory, suggesting that the disruptive effects of musical anomalies extend to semantic processing (Perruchet & Poulin-Charronnat 2013) and may reflect more general phenomena of cognitive interference (Slevc et al. 2013). Neurophysiological evidence for overlap between syntactic and musical brain areas has also been equivocal (Tillmann et al. 2006, Fedorenko et al. 2011).

My skepticism about the connection between syntactic and musical processing has less to do with experimental support for it than with its assumptions about music. While some rather technical musical matters arise here, the main issue is a simple one. In language, there is a powerful motivation for positing long-distance relationships between words. In the sentence The dog with the long tail is chasing the cat, we understand that the dog is doing the chasing, not the tail or the cat; and because there is just one dog, we use is rather than are. There are different ways of capturing these long-distance relationships-with dependencies, constituents, or some combination-but everyone agrees that they are real and need to be represented in some way. In music, there is simply nothing comparable to this. In general, constraints on notes by other notes are of an extremely local nature; in particular, there is a strong tendency in a melody for each note to be close in pitch to the previous note. Nonlocal constraints come not from other notes but rather from higher-level structures. The key constrains the choice of notes and their interpretation; the harmony at a given point constrains the notes in a more local fashion (notes must be part of the chord or resolved by step-i.e., to an adjacent note in the scale); the phrase structure calls for certain melodic/harmonic patterns (cadences) at specific points; the overall range of the melody imposes another kind of constraint (discouraging notes that are far from the middle of the range);

motivic patterns impose yet another kind, favoring note patterns that repeat earlier patterns. While there are hierarchies involved here, they are nonrecursive hierarchies between elements of different types, quite different from the recursive hierarchies present in language. I am hard-pressed to think of an example of nonlocal dependencies between notes, analogous to the *dog-is* relationship in the sentence proposed above—where the function or grammaticality of a note depends on another note several notes earlier.⁶ Similarly, constraints between chords are primarily local in nature (the V chord usually goes to I, II usually goes to V, and so on); nonlocal constraints are imposed mainly by the key and by repeated patterns. Experimental efforts to establish the psychological reality of recursive pitch structure in music have been inconclusive (Serafine et al. 1989, Bigand & Parncutt 1999). Some results that have been claimed to show evidence of recursive structure such as fluctuations in tension (Lerdahl & Krumhansl 2007) or apparent dependencies between chords over long distances (Koelsch et al. 2013)—might also be attributed to effects of higherlevel structures, especially key: Chords outside the key tend to be high in tension, and we expect a chord progression to end in the same key it started in.

Thus the motivation for positing a recursive hierarchical structure, so obvious in language, is absent in music. If I am right, pitch organization in music and syntax in language have virtually nothing in common. While both are hierarchical, the hierarchies involved are no more similar to each other than they are to that of (say) a baseball game. Of course, this does not rule out the possibility that there could be a cognitive interaction between them; perhaps future research will resolve this issue more conclusively. But I see little reason to expect such an interaction.

6. REPETITION IN MUSIC AND LANGUAGE

An essential aspect of music—most kinds of music, anyway—is repetition. (I refer here to repetition of elements and patterns within a piece of music, not between pieces.) Meter itself can be regarded as a form of repetition: the repetition of time intervals between pairs of adjacent beats. Repetition is pervasive, too, at higher levels of musical structure (Margulis 2014). Consider a pop song: We find repetition of short pitch patterns (e.g., a riff on a guitar or synthesizer); within the verse or chorus, melodic phrases are repeated, along with harmonic patterns; and whole sections (verses and choruses) are repeated to create the song's large-scale form. Much the same could be said of music in other Western styles, serious and popular.

With regard to language, the concept of repetition has been construed and explored in various ways (for surveys, see Fischer 1994, Johnstone 1994). An area of particular focus has been the role of repetition in conversation. This includes the tendency, generally unintentional and unnoticed, to repeat recently heard syntactic patterns (such as the passive construction), a phenomenon known as syntactic priming (Pickering & Ferreira 2008); it also includes repetition of words and phrases, which can serve functions such as questioning or affirming someone else's statement (Tannen 1989). Repetition may even be grammaticalized; in some languages, for example, a noun may be converted into an adjective by repeating it (Finkbeiner & Freywald 2018). Here I wish to focus on some other uses of repetition that I believe are more related to music. I have

⁶One might point to implied counterpoint (also called pseudopolyphony). This is where a single sequence of pitches is actually composed of multiple interleaved melodic lines; Bach uses this to great effect in his solo violin and cello suites. This is indeed a type of dependency between nonadjacent pitches. In this case, however, the pitch sequence really consists of multiple independent (or at least hierarchically equal) lines; there is no recursive dependency structure that connects them all, as there is in a sentence. One could also say that the lines are all dependent on the underlying harmony, making them indirectly interdependent. Again, I see no linguistic parallel here.

in mind in particular the repetition of linguistic patterns within a sentence, or between nearby sentences, in a way that is intentional and emphasized (meaning that it is meant to be noticed). Often, such uses of repetition are treated as devices of rhetoric. Classical rhetoricians discussed repetition at length, focusing especially on the repetition of words and phrases, and proposed various subtypes of this strategy (Vickers 1994). One device is the repeated use of a word or phrase at the beginning of a series of sentences or clauses. Examples of this readily come to mind from famous speeches; think of Churchill's "we shall fight on the beaches, we shall fight on the landing grounds" or Martin Luther King Jr.'s "I have a dream" speech. A recurrent theme in discussions of rhetorical repetition is that it helps to establish the emotional commitment and authenticity of the speaker and to arouse emotions in the listener; Vickers (1994) has offered numerous examples.

Also of interest is the repetition of syntactic structures. In a recent study (Temperley & Gildea 2015), Dan Gildea and I examined syntactic repetition in coordinate NP phrases in English, sometimes known as syntactic parallelism. Previous studies have shown that, in coordinate phrases, the syntactic structure of the second coordinate tends to match that of the first, and that comprehension is sensitive to this (Frazier et al. 2000, Dubey et al. 2008). We verified this phenomenon using a large corpus of written English and made three further predictions based on the theory of Uniform Information Density (Fenk-Oczlon 1989, Levy & Jaeger 2007). (Information is the negative log of probability; as probability decreases, information increases.) First we predicted that, in "matching" NP coordinates (i.e., those sharing the same syntactic structure), the second coordinate would tend to have less probable words and word combinations than the first, the reason being that the second coordinate is more syntactically predictable than the first and therefore is able to be lower in lexical probability. Thus we predicted example 1a to be more typical than example 1b. We also predicted that the second coordinate would be lower in lexical probability when it matched the first coordinate (making it syntactically predictable) than when it did not; thus, we expected second conjuncts with low lexical probability (such as an inefficient corkscrew) to show an especially strong tendency to occur after a syntactically matching phrase (as in example 1a) rather than a nonmatching one (as in example 1c). Both predictions were confirmed (examples are from Temperley & Gildea 2015, p. 1805).

- (1a) A tug-of-war between an old bottle and an inefficient corkscrew may do as much harm as a week at sea.
- (1b) A tug-of-war between *an inefficient corkscrew* and *an old bottle* may do as much harm as a week at sea.
- (1c) A tug-of-war between *a bottle that is several years old* and *an inefficient corkscrew* may do as much harm as a week at sea.

Our third prediction—especially important for my argument here—was that rare syntactic patterns would show an especially strong bias toward repeated usage; in this case, the syntactic pattern itself creates a spike in information that can be softened by repeating it. This prediction, too, was confirmed; rare syntactic NP expansions like determiner + adjective show an especially strong tendency to be repeated, as in the following sentences (example 2a is from Temperley & Gildea 2015, p. 1805).

- (2a) Its uneasy mixture of *the serious* and *the comic* is no doubt one reason why it is very much in vogue with directors just now.
- (2b) But above all, this is the national headquarters for boiler-room operators, those slick-talking snake-oil salesmen who use the telephone to extract money from *the gullible* and *the greedy* and then vanish.

In a subsequent study (Temperley 2019), I proposed musical analogs to these phenomena. In music, it is very common for patterns of melodic intervals to be repeated but shifted along the scale. (The opening of Beethoven's Fifth Symphony is an example; the pattern of the first four notes is repeated in the second four, shifted down by a step.) When a pattern of musical intervals is repeated, then, the second occurrence is generally low in information (highly predictable). Corpus analysis shows that when an interval pattern is repeated but with a single pitch interval changed, the interval tends to be increased in size rather than decreased. This is as we would predict from an informational viewpoint; large intervals are rarer than small intervals, so increasing the size of an interval in a repetition of a pattern makes it less probable, balancing the low information of the rest of the pattern (analogous to using less common words in the second conjunct of a coordinate phrase). As for our third prediction (Temperley & Gildea 2015), it has been observed that certain rare melodic devices, such as a nonchord tone that does not resolve by step, have an especially strong tendency to be used in repetitive contexts; this is analogous to the increased tendency for repetition in rare NP expansions.

Further study of linguistic corpus data has made me aware that the tendency toward repetition of rare syntactic constructions extends well beyond NPs. Consider the following sentences from a large corpus of late-1980s *Wall Street Journal* text. Each one contains a highly unusual syntactic construction repeated in a coordinate phrase (I describe each construction in parentheses).

- (3) Large and lavish, "Traviata" is another addition to the Met's growing stock of cast-proof productions.... (Adjectives without complement phrases, used as a premodifying adjunct phrase)
- (4) [Mr. X] has come to work this day wearing sports shirt, slacks and cap because he plans to play golf in the afternoon. (Singular count nouns—sports shirt and cap—without determiners; this type of construction was not found in Temperley & Gildea's (2015) study, since the nouns in such constructions are not annotated in the corpus as separate NPs.)
- (5) [Mr. Y says that] the NFL is telling college players to "go ahead and rob, steal, cheat and kill." (Normally transitive verbs—rob, steal, and kill—without direct objects)
- (6) Such comparisons hurt because despite its unpopularity, President Chun Doo Hwan's government raised Koreans' ambitions—socially, economically and politically. (Postverbal manner adverbs after an em dash)

The use of repetition here licenses syntactic constructions that would seem odd, if not incorrect, without repetition—we might call them "edgy" constructions. Compare the sentences above with the versions below in which the repetitions have been removed. Examples 7 and 8 seem downright ungrammatical; examples 9 and 10 are, at best, awkward.

- (7) Lavish, "Traviata" is another addition to the Met's growing stock of cast-proof productions....
- (8) [Mr. X] has come to work this day wearing sports shirt because he plans to play golf in the afternoon.
- (9) [Mr. Y says that] the NFL is telling college players to "go ahead and rob."
- (10) Such comparisons hurt because despite its unpopularity, President Chun Doo Hwan's government raised Koreans' ambitions—economically.

The passage of text in example 11 below contains two sentence fragments, similar in their internal lexical and syntactic structure. Again, this normally ungrammatical construction (an NP with no predicate) seems to be licensed by its repetition. Using just a single sentence fragment (as in example 12) seems stylistically odd; my eyes stumble on it as if it were an error.

- (11) In his lifetime, 47 years, the Communists brought electricity to his village and piped in drinking water from the reservoir. No phones. No gas. "We wanted them to build a road here," he says.
- (12) In his lifetime, 47 years, the Communists brought electricity to his village and piped in drinking water from the reservoir. No phones. "We wanted them to build a road here," he says.

A related phenomenon, though somewhat different, is illustrated by sentence 13a. The sentence features a subject NP modified by three conjoined appositive NPs. In this case (unlike in the sentences above), including just one of the phrases, as in sentence 13b, would be syntactically unproblematic. What is unusual, and normally ungrammatical, is the use of conjoined NPs without a conjunction; this seems to be licensed, or at least improved, by the lexical repetition <u>within</u> the NPs. Without this repetition (sentence 13c), the sentence seems clumsy, though not ungrammatical.

- (13a) Colombia alone—its government, its people, its newspapers—does not have the capacity to fight this battle successfully.
- (13b) Colombia alone—its government—does not have the capacity to fight this battle successfully.
- (13c) Colombia alone—its government, the people, media outlets—does not have the capacity to fight this battle successfully.

Repetition is also used in prosody. Perhaps the prime example is listing intonation, in which a repeated contour is used to highlight a series of syntactically parallel phrases (a rising contour, in English anyway), with a contrasting pattern on the final phrase (a falling contour): for instance, *I bought milk, eggs, cheese, and butter* (Cauldwell & Hewings 1996). Consider also **Figure 2**; I sometimes notice people (including myself) saying things like this to their children. Typically, in such tirades, the same prosodic contour is used in every sentence. This is unlike typical listing intonation because the final item in the list has the same intonation as the previous ones (the implication being, perhaps, that the list could go on). What is especially curious about a case like this (and again, unlike typical listing intonation) is that the prosodic contour itself is somewhat variable; all that seems to matter is that it be repeated. **Figure 2** shows three contours that seem quite possible. Again, repetition seems crucial here; applying a different contour to each phrase would seem strange. Using just a single phrase can also sound odd with certain contours (e.g., contour B in **Figure 2**)—like an attempted tirade that never got off the ground.

I have discussed three uses of repetition in language that might seem quite different: the rhetorical repetition of words and phrases; the tendency to confine edgy syntactic constructions to



Figure 2

A series of sentences, with three possible intonational contours.

repetitive, syntactically parallel contexts; and repetitive prosodic patterns. In both the syntactic and prosodic cases, repetition appears to license linguistic patterns that might otherwise seem odd or incorrect; I have argued that this phenomenon has analogies in music. But there is a further commonality between these linguistic phenomena, one that also ties them to music. All of them—by my intuition, anyway—are associated with what are sometimes called persuasive rather than informative uses of language (Kamalski et al. 2008): They are designed to persuade, not simply to convey information in an objective manner. Going along with that, they tend to involve an overt emotional engagement on the part of the producer. In the rhetorical and prosodic cases, this seems clear. In the syntactic case, it is perhaps not so obvious, but it seems to me that patterns like those in examples 2–6 above are much more likely to occur in an arts review, an editorial, or a human-interest news story than in-for example-a hard news article, a scientific journal, or a legal document. In addition, at least in the rhetorical and syntactic cases (not so much in the prosodic one), there is an aesthetic, artful aspect to these linguistic usages: a reveling in linguistic play, in the pleasure of linguistic sounds and forms for their own sake. In the syntactic case, the edginess of the phenomena-stretching the usual rules, and thus posing a mild intellectual challenge for the perceiver-may relate to this as well.

Repetitive uses of language, designed to provoke an emotional response in the perceiver, and with a playful, aesthetic aspect: The parallels with music should be clear. Of course, all of this is also true (even more so) of poetry, whose connections with music have often been noted. Perhaps the lesson of all this is that "musical" uses of language are not confined to poetry but sometimes extend into prose as well.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

LITERATURE CITED

- Acevedo S, Temperley D, Pfordresher P. 2014. Effects of metrical encoding on melody recognition. Music Percept. 31:372–86
- Bernstein L. 1976. The Unanswered Question: Six Talks at Harvard. Cambridge, MA: Harvard Univ. Press
- Besson M, Barbaroux M, Dittinger E. 2017. Music in the brain: music and language processing. In *The Routledge Companion to Music Cognition*, ed. R Ashley, R Timmers, pp. 37–48. Abingdon, UK: Taylor & Francis
- Bidelman GM, Hutka S, Moreno S. 2013. Tone language speakers and musicians share enhanced perceptual and cognitive abilities for musical pitch: evidence for bidirectionality between the domains of language and music. PLOS ONE 8:e60676
- Bigand E, Parncutt R. 1999. Perceiving musical tension in long chord sequences. Psychol. Res. 62:237-54

Bregman AS. 1990. Auditory Scene Analysis: The Perceptual Organization of Sound. Cambridge, MA: MIT Press

- Briefer EF. 2012. Vocal expression of emotions in mammals: mechanisms of production and evidence. J. Zool. 288:1–20
- Brown DE. 1991. Human Universals. Philadelphia: Temple Univ. Press
- Cauldwell R, Hewings M. 1996. Intonation rules in ELT textbooks. ELT J. 50:327-34
- Conklin D, Anagnostopoulou C. 2001. Representation and discovery of multiple viewpoint patterns. In Proceedings of the 2001 International Computer Music Conference, pp. 479–85. Ann Arbor: Michigan Publ.
- Cox A. 2016. Music and Embodied Cognition: Listening, Moving, Feeling, and Thinking. Bloomington: Indiana Univ. Press
- Cutler A, Foss DJ. 1977. On the role of sentence stress in sentence processing. Lang. Speech 20:1-10

Daniele JR, Patel AD. 2013. An empirical study of historical patterns in musical rhythm: analysis of German & Italian classical music using the nPVI equation. *Music Percept*. 31:10–18

Dauer RM. 1983. Stress-timing and syllable-timing reanalyzed. *7. Phon.* 11:51–62

Deutsch D. 2006. The enigma of absolute pitch. Acoust. Today 2:11-19

- Dilley LC, McAuley JD. 2008. Distal prosodic context affects word segmentation and lexical processing. J. Mem. Lang. 59:294–311
- Dingemanse M, Blasi DE, Lupyan G, Christiansen MH, Monaghan P. 2015. Arbitrariness, iconicity, and systematicity in language. *Trends Cogn. Sci.* 19:603–15
- Dubey A, Keller F, Sturt P. 2008. A probabilistic corpus-based model of syntactic parallelism. *Cognition* 109:326-44
- Fedorenko E, Behr M, Kanwisher N. 2011. Functional specificity for high-level linguistic processing in the human brain. PNAS 108:16428–33
- Fenk-Oczlon G. 1989. Word order and frequency in freezes. Linguistics 27:517-56

Filippi P. 2016. Emotional and interactional prosody across animal communication systems: a comparative approach to the emergence of language. *Front. Psychol.* 7:1393

- Finkbeiner R, Freywald U. 2018. Exact Repetition in Grammar and Discourse. Berlin: Walter de Gruyter
- Fischer A, ed. 1994. SPELL: Swiss Papers in English Language and Literature, Vol. 7: Repetition. Tübingen, Ger.: Gunter Narr

Fitch W. 2006. The biology and evolution of music: a comparative perspective. Cognition 100:173-215

- Frazier L, Munn A, Clifton C. 2000. Processing coordinate structures. J. Psycholinguist. Res. 29:343-70
- Gabrielsson A, Lindström E. 2001. The influence of musical structure on emotional expression. In Music and Emotion: Theory and Research, ed. P Juslin, J Sloboda, pp. 223–48. Oxford, UK: Oxford Univ. Press
- Gingold H, Abravanel E. 1987. Music as a mnemonic: the effects of good- and bad-music settings on verbatim recall of short passages by young children. *Psychomusicology* 7:25–39
- Gjerdingen R. 2007. Music in the Galant Style. New York: Oxford Univ. Press

Gordon RL, Fehd HM, McCandliss BD. 2015. Does music training enhance literacy skills? A meta-analysis. Front. Psychol. 6:1777

Grabe E, Low E. 2002. Durational variability in speech and the rhythm class hypothesis. In *Laboratory Phonology* 7, ed. C Gussenhoven, N Warner, pp. 515–46. Berlin: De Gruyter Mouton

- Granroth-Wilding M, Steedman M. 2014. A robust parser-interpreter for jazz chord sequences. J. Music Res. 43:355–74
- Halle J, Lerdahl F. 1993. A generative textsetting model. Curr. Musicol. 55:3-23
- Hannon E. 2009. Perceiving speech rhythm in music: Listeners classify instrumental songs according to language of origin. *Cognition* 111:403–9
- Hanslick E. 1986 (1891). On the Musically Beautiful: A Contribution Towards the Revision of the Aesthetics of Music, trans. G Payzant. Indianapolis: Hackett
- Hayes B. 1995. Metrical Stress Theory: Principles and Case Studies. Chicago: Univ. Chicago Press
- Hayes B, Kaun A. 1996. The role of phonological phrasing in sung and chanted verse. Linguist. Rev. 13:243–304
- Heffner CC, Slevc LR. 2015. Prosodic structure as a parallel to musical structure. Front. Psychol. 6:1962
- Hockett CF. 1960. The origin of speech. Sci. Am. 203:5–12
- Huron D. 2006. Sweet Anticipation. Cambridge, MA: MIT Press
- Huron D. 2016. Voice Leading: The Science Behind a Musical Art. Cambridge, MA: MIT Press
- Huron D, Royal M. 1996. What is melodic accent? Converging evidence from musical practice. *Music Percept*. 13:489–516
- Johnstone B, ed. 1994. Repetition in Discourse: Interdisciplinary Perspectives. Norwood, NJ: Ablex
- Jones MR, Moynihan H, MacKenzie N, Puente J. 2002. Temporal aspects of stimulus-driven attending in dynamic arrays. Psychol. Sci. 13:313–19
- Juslin PN, Laukka P. 2003. Communication of emotions in vocal expression and music performance: different channels, same code? *Psychol. Bull.* 129:770–814
- Kamalski J, Sanders T, Lentz L. 2008. Coherence marking, prior knowledge, and comprehension of informative and persuasive texts: sorting things out. *Discourse Process*. 45:323–45
- Kelly MH, Bock JK. 1988. Stress in time. J. Exp. Psychol.: Hum. Percept. Perform. 14:389-403
- Kirkham NZ, Slemmer JA, Johnson SP. 2002. Visual statistical learning in infancy: evidence for a domain general learning mechanism. *Cognition* 83:B35–42

Kivy P. 1980. The Corded Shell. Philadelphia: Temple Univ. Press

- Koelsch S, Rohrmeier M, Torrecuso R, Jentschke S. 2013. Processing of hierarchical syntactic structure in music. PNAS 110:15443–48
- Lee C, Brown L, Müllensiefen D. 2017. The musical impact of Multicultural London English (MLE) speech rhythm. *Music Percept.* 34:452–81
- Lerdahl F. 2001a. The sounds of poetry viewed as music. Ann. N.Y. Acad. Sci. 930:337-54

- Lerdahl F, Jackendoff R. 1983. A Generative Theory of Tonal Music. Cambridge, MA: MIT Press
- Lerdahl F, Krumhansl CL. 2007. Modeling tonal tension. Music Percept. 24:329-66
- Levy R, Jaeger F. 2007. Speakers optimize information density through syntactic reduction. In Advances in Neural Information Processing Systems, ed. B Schölhopf, J Platt, T Hofmann, pp. 849–56. Cambridge, MA: MIT Press
- Liberman M. 1975. The intonational system of English. PhD Thesis, MIT, Cambridge, MA
- Liberman M, Prince A. 1977. On stress and linguistic rhythm. *Linguist. Ing.* 8:249–336
- Lipiczky T. 1985. Tihai formulas and the fusion of 'composition' and 'improvisation' in North Indian music. Music. Q. 71:157–71
- Madsen SM, Marschall M, Dau T, Oxenham AJ. 2019. Speech perception is similar for musicians and nonmusicians across a wide range of conditions. Sci. Rep. 9:10404
- Margulis EH. 2014. On Repeat: How Music Plays the Mind. New York: Oxford Univ. Press
- Mattheson J. 1739. Der vollkommene Kapellmeister. Hamburg, Ger.: C. Herold
- McGowan RW, Levitt AG. 2011. A comparison of rhythm in English dialects and music. *Music Percept*. 28:307–14
- Mendelssohn F. 1878 (1842). Briefe aus den Jahren 1830 bis 1847 von Felix Mendelssohn Bartholdy, ed. C. Mendelssohn. Leipzig, Ger.: Hermann Mendelssohn
- Meyer L. 1956. Emotion and Meaning in Music. Chicago: Chicago Univ. Press
- Monahan S. 2013. Action and agency revisited. J. Music Theory 57:321-71
- Palmer C, Kelly MH. 1992. Linguistic prosody and musical meter in song. J. Mem. Lang. 31:525-42
- Parbery-Clark A, Skoe E, Lam C, Kraus N. 2009. Musician enhancement for speech-in-noise. Ear Hear. 30:653–61
- Parncutt R. 1989. Harmony: A Psychoacoustical Approach. Berlin: Springer
- Patel AD. 2003. Language, music, syntax and the brain. Nat. Neurosci. 6:674-81
- Patel AD. 2005. The relationship of music to the melody of speech and to syntactic processing disorders in aphasia. In The Neurosciences and Music II: From Perception to Performance, ed. G Avanzini, L Lopez, S Koelsch, M Manjno, pp. 59–70. New York: N.Y. Acad. Sci.
- Patel AD. 2008. Music, Language, and the Brain. New York: Oxford Univ. Press
- Patel AD, Daniele JR. 2003. An empirical comparison of rhythm in language and music. Cognition 87:B35-45
- Pearce M, Wiggins G. 2004. Improved methods for statistical modelling of monophonic music. J. New Music Res. 33:367–85
- Perruchet P, Poulin-Charronnat B. 2013. Challenging prior evidence for a shared syntactic processor for language and music. *Psychon. Bull. Rev.* 20:310–17
- Pfordresher PQ, Brown S. 2009. Enhanced production and perception of musical pitch in tone language speakers. Atten. Percept. Psychophys. 71:1385–98
- Pickering MJ, Ferreira VS. 2008. Structural priming: a critical review. Psychol. Bull. 134:427-59
- Povel D, Essens P. 1985. Perception of temporal patterns. Music Percept. 2:411-40
- Ramus F. 2002. Acoustic correlates of linguistic rhythm: perspectives. In Proceedings of Speech Prosody, ed. B Bell, I Marlien, pp. 115–20. Aix-en-Provence, Fr.: Lab. Parole Lang.
- Roach P. 1982. On the distinction between 'stress-timed' and 'syllable-timed' languages. In *Linguistic Contro*versies, ed. D Crystal, pp. 73–79. London: Edward Arnold
- Rohrmeier M. 2011. Towards a generative syntax of tonal harmony. J. Math. Music 5:35-53
- Rolland P. 1999. Discovering patterns in musical sequences. J. New Music Res. 28:334-50
- Russell JA. 1980. A circumplex model of affect. 7. Pers. Soc. Psychol. 39:1161-78
- Saffran JR, Aslin RN, Newport EL. 1996. Statistical learning by 8-month-old infants. Science 274:1926–28
- Saffran JR, Johnson E, Aslin RN, Newport EL. 1999. Statistical learning of tone sequences by human infants and adults. *Cognition* 70:27–52

Lerdahl F. 2001b. Tonal Pitch Space. New York: Oxford Univ. Press

Savage PE, Brown S, Sakai E, Currie TE. 2015. Statistical universals reveal the structures and functions of human music. PNAS 112:8987–92

Schellenberg EG, Trehub SE. 2008. Is there an Asian advantage for pitch memory? *Music Percept*. 25:241–52 Schenker H. 1979 (1935). *Free Composition*, trans. E Oster. New York: Longman

- Schubert P, Cumming J. 2015. Another lesson from Lassus: using computers to analyse counterpoint. Early Music 43:577–86
- Serafine ML, Glassman N, Overbeeke C. 1989. The cognitive reality of hierarchic structure in music. Music Percept. 6:397–430
- Shattuck-Hufnagel S, Turk AE. 1996. A prosody tutorial for investigators of auditory sentence processing. J. Psycholinguist. Res. 25:193–247
- Slevc LR, Reitman J, Okada B. 2013. Syntax in music and language: the role of cognitive control. Proc. Annu. Meet. Cogn. Sci. Soc. 35:3414–19
- Slevc LR, Rosenberg JC, Patel AD. 2009. Making psycholinguistics musical: self-paced reading time evidence for shared processing of linguistic and musical syntax. *Psychon. Bull. Rev.* 16:374–81
- Sloboda J. 1983. The communication of musical metre in piano performance. Q. 7. Exp. Psychol. A 35:377-96
- Sluijter AM, Van Heuven VJ. 1996. Spectral balance as an acoustic correlate of linguistic stress. J. Acoust. Soc. Am. 100:2471–85
- Spencer H. 1890. The origin of music. Mind 15:449-68
- Swain J. 1997. Musical Languages. New York: W.W. Norton
- Tan I, Lustig E, Temperley D. 2019. Anticipatory syncopation in rock: a corpus study. Music Percept. 36:353-70
- Tannen D. 1989. Talking Voices: Repetition, Dialogue, and Imagery in Conversational Discourse. Cambridge, UK: Cambridge Univ. Press
- Temperley D. 1999a. The question of purpose in music theory: description, suggestion, and explanation. *Curr: Musicol.* 66:66–85
- Temperley D. 1999b. Syncopation in rock: a perceptual perspective. Pop. Music 18:19-40
- Temperley D. 2001. The Cognition of Basic Musical Structures. Cambridge, MA: MIT Press
- Temperley D. 2004. Communicative pressure and the evolution of musical styles. Music Percept. 21:313–37
- Temperley D. 2009. Distributional stress regularity: a corpus study. 7. Psycholinguist. Res. 38:75-92
- Temperley D. 2017. Rhythmic variability in European vocal music. Music Percept. 35:193-99
- Temperley D. 2019. Uniform information density in music. Music Theory Online 25(2):5
- Temperley D, Gildea D. 2015. Information density and syntactic repetition. Cogn. Sci. 39:1802-23
- Temperley N, Temperley D. 2011. Music-language correlations and the 'Scotch Snap.' Music Percept. 29:51-63
- Temperley N, Temperley D. 2013. Stress-meter alignment in French vocal music. J. Acoust. Soc. Am. 134:520– 27
- Terhardt E. 1974. Pitch, consonance, and harmony. J. Acoust. Soc. Am. 55:1061-69
- Tillmann B, Koelsch S, Escoffier N, Bigand E, Lalitte P, et al. 2006. Cognitive priming in sung and instrumental music: activation of inferior frontal cortex. *NeuroImage* 31:1771–82
- VanHandel L. 2017. The war of the Romantics: an alternate hypothesis using nPVI for the quantitative anthropology of music. *Empir. Musicol. Rev.* 11:234–42
- VanHandel L, Song T. 2010. The role of meter in compositional style in 19th-century French and German art song. J. New Music Res. 39:1–11
- Vickers B. 1994. Repetition and emphasis in rhetoric: theory and practice. See Fischer 1994, pp. 85-114
- Vukovics K, Shanahan D. 2020. Rhythmic variability, language, and style: a replication and extension of nPVI findings with the RISM dataset. J. New Music Res. 49:285–97

Contents

Annual Review of Linguistics

Volume 8, 2022

How I Got Here and Where I'm Going Next <i>Sarah G. Thomason</i>
Argument Structure in Sign Languages Vadim Kimmelman
Advances in Morphological Theory: Construction Morphology and Relational Morphology Jenny Audring
Perspective Shift Across Modalities Emar Maier and Markus Steinbach
Learning Through Processing: Toward an Integrated Approach to Early Word Learning <i>Stephan C. Meylan and Elika Bergelson</i> 77
The Probabilistic Turn in Semantics and Pragmatics <i>Katrin Erk</i>
Coherence Establishment as a Source of Explanation in Linguistic Theory Andrew Kehler
When Do Children Lose the Language Instinct? A Critical Review of the Critical Periods Literature <i>Joshua K. Hartshorne</i>
Music and Language David Temperley
Crosslinguistic Corpus Studies in Linguistic Typology Stefan Schnell and Nils Norman Schiborr
On the Acquisition of Attitude Verbs Valentine Hacquard and Jeffrey Lidz
Meaning and Alternatives Nicole Gotzner and Jacopo Romoli
Second Language Sentence Processing Holger Hopp

Nominalization and Natural Language Ontology Scott Grimm and Louise McNally
Perfects Across Languages Östen Dahl
Speech and Language Outcomes in Adults and Children with Cochlear Implants <i>Terrin N. Tamati, David B. Pisoni, and Aaron C. Moberly</i>
Glue Semantics Ash Asudeh
Intonation and Prosody in Creole Languages: An Evolving Ecology Shelome Gooden 34.
Navigating Accent Variation: A Developmental Perspective Elizabeth K. Johnson, Marieke van Heugten, and Helen Buckler
Reverse Engineering Language Acquisition with Child-Centered Long-Form Recordings Marvin Lavechin, Maureen de Seyssel, Lucas Gautheron, Emmanuel Dupoux, and Alejandrina Cristia
Stance and Stancetaking Scott F. Kiesling
Neurocomputational Models of Language Processing John T. Hale, Luca Campanelli, Jixing Li, Shohini Bhattasali, Christophe Pallier, and Jonathan R. Brennan
Semantic Structure in Deep Learning <i>Ellie Pavlick</i>
Deriving the Wug-Shaped Curve: A Criterion for Assessing Formal Theories of Linguistic Variation <i>Bruce Hayes</i>
Structural, Functional, and Processing Perspectives on Linguistic Island Effects <i>Yingtong Liu, Elodie Winckel, Anne Abeillé, Barbara Hemforth,</i> <i>and Edward Gibson</i>

Errata

An online log of corrections to *Annual Review of Linguistics* articles may be found at http://www.annualreviews.org/errata/linguistics