26 Characterizing Rhythmic Regularity in Speech and Song

MEYHA CHHATWAL, SELINA SABETTI-FRANKLIN, AND CHRISTINA VANDEN BOSCH DER NEDERLANDEN

26.1 What Is Rhythmic about Speech and Song?

Rhythms surround us: from footsteps in the hallway to the blinking red of a traffic signal. Although rhythms are a pervasive part of daily life, they play an important role in vocal communication for humans and a range of species (Honing, 2012; Merchant and Honing, 2014; Patel, 2010; Ravignani and Madison, 2017). In humans, rhythms are present in two primary means of communication: speech and song. People perceive rhythms in speech from prosodic and linguistic stress patterns of syllables, and perceive rhythms in song from notes as they unfold in time (Patel, 2010). Although the definition of rhythm can vary considerably, here we adopt a music cognition perspective and describe *rhythm* in its broadest sense as a pattern of events unfolding in time (McAuley, 2010; Ravignani and Madison, 2017). As such, speech events are typically syllables and song events are notes. This definition is inclusive of many different types of repeating sounds, such as cricket chirps, waves crashing on the beach, chewing food, and even a dough hook smacking around the bowl of an electric mixer.

Speech and song are both rhythmic, but a key difference between the two domains is how events relate to one another. In song, and music more generally, rhythms are organized around perceptually strong events, called a beat or pulse (Povel and Essens, 1985). The alternation of strong and weak beats in music occurs at roughly equal intervals, which gives rise to the perception of *isochrony* (i.e., "*iso*" meaning "same" and "*chronos*" meaning "time") despite considerable variation in note lengths or timing of note onsets relative to the beat. In many cases, a physical event is not always present on the beat and yet listeners perceive a strong pulse in that location (London, 2004; Nave et al., 2023; Parncutt, 1994; Snyder and Krumhansl, 2001; Temperley, 2004). When rhythms are structured around the beat, note durations are related by integer multiples with small integer ratios. That means that a long note in a piece of music is one, two, three, or four times the duration of a short note in the sequence (e.g., quarter versus half note; Jacoby and McDermott, 2017; Roeske et al., 2020). In contrast, speech rhythms are dictated by the word length, syntactic structure, and prosodic emphasis of the specific utterance (Patel, 2010; Selkirk, 1980; Turk and Shattuck-Hufnagel, 2013). For

https://doi.org/10.1017/9781009295888.031 Published online by Cambridge University Press

instance, stressed syllables are often longer than unstressed syllables (Cutler and Foss, 1977; Fry, 1955, 1958; Seidl et al., 2014). Articles, such as *a*, *an*, and *the*, occur in fixed positions relative to content words in English, and are unstressed, monosyllabic, and shorter than many content words, creating an unstressed-stressed rhythm (Andreou et al., 2011; Hayes, 1985). Words at the end of a phrase are lengthened (Klatt, 1975), words that receive prosodic emphasis for effect are also held out longer (Selkirk, 1995), and different languages have different stress patterns or a lack of stress at all (Cutler and Clifton, 1984; Peperkamp et al., 2010). The rhythmic patterns of speech can be predictable and repeating, sometimes even alternating between stressed and unstressed syllables much like the definition of a beat above (Breen and Clifton Jr, 2011; Cutler and Foss, 1977; Levey and Raphael, 2002; Patel, 2010). However, syllable durations do not relate to one another by consistent integer relationships, which means that there is no *regular* pulse or beat in speech.

We present two additional terms, periodicity and rhythmic regularity, to carefully differentiate how rhythmic events relate to each other in speech and song. *Periodicity* is defined as a pattern repeating in time (Patel, 2010) and is traditionally related to a fixed interval or period with which the pattern repeats. However, speech research has long used the term periodicity to mean repeating patterns, such as alternations of stressed and unstressed syllables, that do not have a fixed period (Beier and Ferreira, 2018). So here we use this term in that sense of a repeating pattern that does not repeat at a fixed interval, but instead at some variable period (e.g., like the meaning of *periodic*). To illustrate this notion, imagine an auditory notification played to alert travelers that a subway train is about to arrive (a quick sequential ascending major chord: G-B-D-G) and a new train comes every four to six minutes (see Figure 26.1A). This musical alert is periodic – it occurs repeatedly whenever a train arrives – and is also rhythmic – the four-note chord has a short-short-short-long pattern of note events – but it does not give rise to the perception of a regular pulse. *Rhythmic regularity* is when an alternation between strong and weak events gives rise to a beat (i.e., a series of perceptually strong events; Large, 2008) and the intervals between strong events occur at roughly equal intervals (Ravignani and Madison, 2017; Ravignani and Norton, 2017). Think again of the train arrival notification. If the train arrived every 2,400 ms and the musical alert itself was composed of three 200 ms intervals followed by one 600 ms interval, then it would give rise to a regular pulse, or a beat in music. Specifically, the first and last notes would become strong beats because the (implausible) 2,400 ms train arrival window would suggest two 600 ms silent beats, making the whole sequence a cycle of four beats (see Figure 26.1B). In everyday English speech, an individual phrase provides a rhythm of stressed and unstressed syllables, and, when we speak in longer utterances, new rhythmic patterns with similar properties repeat at irregularly spaced phrasal and prosodic intervals. As such, speech is considered periodic: It is a repeating rhythmic pattern of events, but it is not rhythmically regular. In contrast, song is rhythmically regular, where a repeating pattern occurs at intervals of the same duration and gives rise to the perception of an underlying pulse or beat. In music, these beats are grouped into a hierarchy of strong and weak beats (e.g., groups of three or four beats per measure; Jones, 1985; Lerdahl and Jackendoff, 1983) called *meter*. Although speech has also been described as having meter (e.g., Selkirk, 1984), metricality in speech is the alternation of strong and weak syllables without a fixed integer-related interval relative to previous syllables. Speech,

https://doi.org/10.1017/9781009295888.031 Published online by Cambridge University Press

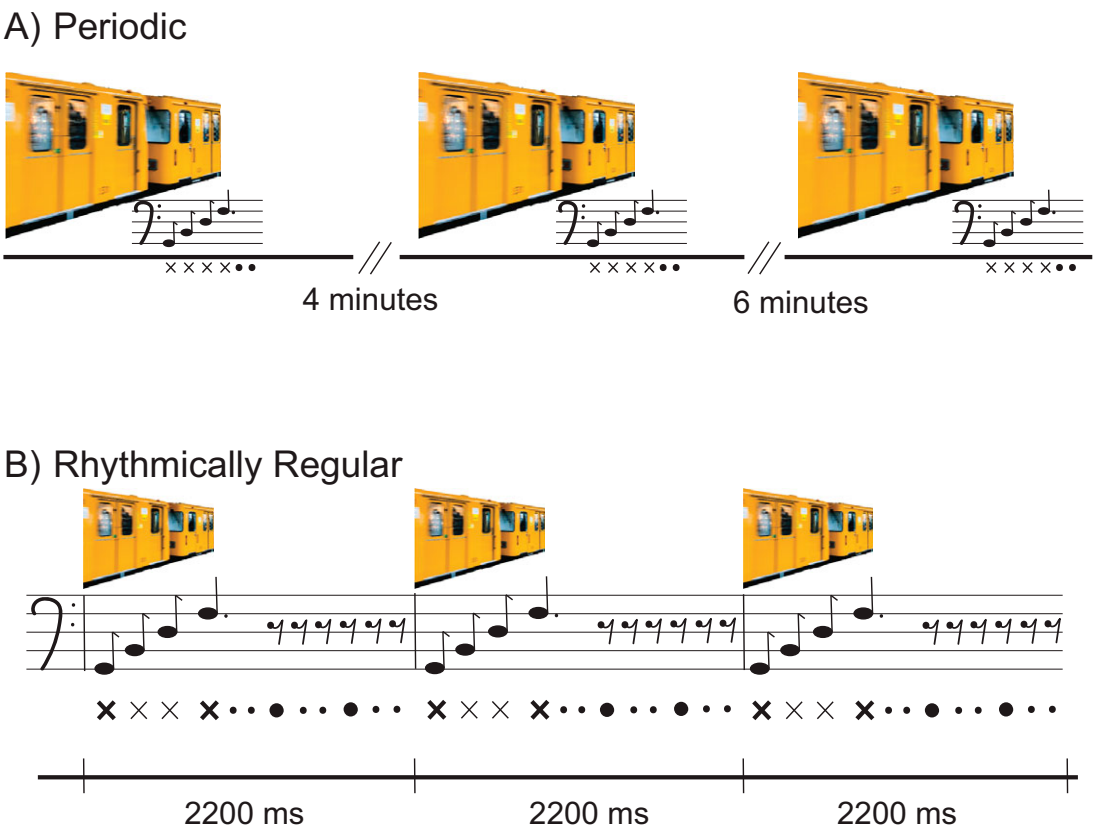


Figure 26.1 Illustrating rhythmic regularity as different from periodic.
Illustration of the difference between a signal that is A) periodic and rhythmic compared to one that is B) rhythmically regular and gives rise to the percept of a beat.

even at the metrical level, is more about the recurrence of a particular pattern (periodicity) instead of the presence of an underlying beat occurring at isochronous intervals.

The differences between periodicity and rhythmic regularity may seem subtle, but the ramifications of the presence or absence of rhythmic regularity are anything but subtle. Musical behaviors such as dancing, singing, and playing an instrument all depend on the presence of rhythmic regularity to allow an individual or multiple individuals to move or play together in an ensemble or to play with music (Hannon et al., 2018; Honing, 2012; Nave-Blodgett et al., 2021). Even the survival of the human species has been related back to rhythmic coordination and its importance for social bonding, especially through coordination with a beat as in music-making (Savage et al., 2021; Trainor and Cirelli, 2015).

Music is not the only stimulus that has rhythmic regularity, even if it is the strongest naturally occurring example of it. For instance, our heartbeat is perhaps the first regular beat we hear, with a weak-strong alternation in the lub-dub-lub-drub of a normal heartbeat cycle (De Meo et al., 2016; Ullal-Gupta et al., 2013). Even speech can have a beat, exemplified

https://doi.org/10.1017/9781009295888.031 Published online by Cambridge University Press

through the uncharacteristically regular alternation of unstressed and stressed syllables in iambic pentameter (e.g., "*the rose is bright and shines a beam of light*") and other forms of poetry. When speech has a beat, that type of speech arguably plays a qualitatively different functional role than speech without a beat. Instead of a speech versus song dichotomy based on the presence or absence of rhythmic regularity, it is more likely that a communicative continuum exists with speech and song as the endpoints and several different types of vocalizations falling between, including poetry, rap, sprechstimme, public speeches, emotional speech, acting, infant-directed speech, and more (Vanden Bosch der Nederlanden et al., 2022b). There are many acoustic features that move a stimulus along the speech-to-song continuum (Albouy et al., 2024; Brown, 2000; Fitch, 2006; Ozaki et al., 2024; Vanden Bosch der Nederlanden et al., 2022b), but rhythmic regularity is key to differentiating the endpoints of this continuum (see Chapters 11, 32, and 35).

26.2 Examining Acoustic Features of Rhythmic Regularity in Speech and Song

When adults describe acoustic differences between speech and song, they tend to focus on the different uses of pitch, while younger children (4–11 years of age) describe differences based on acoustic timing (Vanden Bosch der Nederlanden et al., 2022b). Our previous qualitative work simply asked children and adults to describe the difference between speech and song in their own words, and we reported themes from their responses binned across all temporal acoustic features and other features (melodic, spectral, loudness). Here, we provide a reanalysis of these themes (data here: https://osf.io/xar3c/) to examine how children and adults describe different temporal features, such as the presence of a beat, throughout development. As is clear from Table 26.1, temporal characteristics were mentioned infrequently for describing the differences between speech and song (5–30% of participants depending on the age group: N = 19, 4–7-year-olds; N = 16, 8–11-year-olds; N = 16, 12–17-year-olds; and N = 47, 18–24-year-olds). Children and adults mentioned tempo at similar rates (reporting that song is faster than speech for

Table 26.1 Spontaneous mentions of rhythm in children's differentiation of speech and song

Reanalysis of the qualitative data of Vanden Bosch der Nederlanden et al. (2022b) *illustrating that individual participants' spontaneous descriptions of the differences between speech and song infrequently describe beat or rhythm as a key differentiator compared to pitch or melody. Values are the proportion of participants endorsing each acoustic feature by age group.*

Age group	*Beat/rhythm*	*Tempo*	*Flow*
4–7	0.11	0.11	0.05
8–11	0.13	0.13	0.19
12–17	0.06	0.13	0.19
18–24	0.23	0.09	0.30

https://doi.org/10.1017/9781009295888.031 Published online by Cambridge University Press

children, but that speech is faster than song for adults); smoothness or flow (song is more smooth than speech) – which could be related to temporal and/or spectral features – increased with age; and adults mentioned beat or rhythm (lumped together because children's definitions did not necessarily mention beat but talked about rhythm in a beat-like manner) more often than children overall (reporting more beat/rhythm in song than speech). Although these qualitative data suggest that rhythmic regularity is not the first thing people focus on in their verbal or written descriptions, perceptual judgments reported by our group (Vanden Bosch der Nederlanden et al., 2022b) suggest otherwise: Adults rank the presence of a beat and rhythmic regularity in the top three characteristics that distinguish speech from song, and children of all ages agree that song has a beat.

Since rhythmic regularity is important for distinguishing speech from song, how is the degree of rhythmic regularity measured across these domains? Many studies characterize the strength of a beat in music (Alluri and Toiviainen, 2010; Burger et al., 2013, 2014; De Gregorio et al., 2021; Henry et al., 2017; Lartillot et al., 2008; Matthews et al., 2019; Roeske et al., 2020) and the degree of durational contrastiveness in speech (i.e., short–long) (Arvaniti, 2012; Grabe and Low, 2002; Jadoul et al., 2016; Nolan and Jeon, 2014; White and Mattys, 2007; Wiget et al., 2010), but few compare the degree of rhythmic regularity across speech and song. Our recent study (Yu et al., 2023) illustrated that subjective ratings of regularity reliably differentiated speech from song, when regularity was operationalized as how easy it would be to tap or clap along with the stimulus. Regularity ratings were also predicted by the acoustic features of spectral flux and syllable duration, which meant that less moment-to-moment change in the spectrum (i.e., spectral flux) and longer durations were associated with higher regularity ratings (Yu et al., 2023). This finding held true even after controlling for modality (speech and song), suggesting that these features predicted regularity ratings *for both* speech and song. Other studies that examine what features listeners use to differentiate speech and song can also shed light on key temporal characteristics. Speech and song are differentiated by their temporal modulation rates (e.g., song is slower than speech; Ding et al., 2017), spectro-temporal modulation rates (co-occurrence of high-frequency spectral modulation with low-frequency temporal modulation for song relative to speech; Albouy et al., 2024), pulse clarity (Hilton et al., 2022), tempo (Ozaki et al., 2024), and rhythmic regularity (Ozaki et al., 2024; see their exploratory analyses). These features, including spectral flux and duration mentioned above, are certainly dependent on temporal dynamics, but nearly all fail to capture the integer-multiple relationships that characterize song's alignment with a beat. It is still an open question what features listeners use to perceive rhythmic regularity in speech and song and whether the features reported here are merely a byproduct of the integer relationships in sounds with a beat.

One issue with the practice of relating acoustic features in speech and song to subjective ratings of regularity is the influence human perception has in altering the incoming signal. For instance, naturally occurring music does not have strictly equal durations between beats because of motor output delays, micro-timing deviations, or expressive timing (Appen et al., 2015; Danielsen et al., 2015; Datseris et al., 2019; Drake and Palmer, 1993; Rasch, 1988; Senn et al., 2016). People perceptually regularize, or quantize, the signal to make it sound more rhythmically regular than the physical signal would suggest. It is also likely that people perceive the moment of the beat in natural stimuli later than the sound's physical onset. In musical timbres, people report that the

https://doi.org/10.1017/9781009295888.031 Published online by Cambridge University Press

perceptual occurrence of an event, or p-center, is delayed relative to the physical onset of the sound (see Chapter 11). Fast attack time sounds (e.g., percussion) have p-centers with shorter delays relative to the onset than slow attack time sounds (e.g., violin), and longer notes have p-centers with longer delays relative to onsets compared to shorter notes (Danielsen et al., 2019). The timing of the p-center is also dependent on the experience or expertise of the listener, which leads researchers to describe p-centers in music as "beat bins," or windows of time, in which people are most likely to hear a beat (Speich et al., 2023).

The concept of the p-center originates from speech studies (Morton et al., 1976), but this literature is quite mixed as to the acoustic correlates of the p-center (Villing, 2010). The most consistent finding relates the p-center to a window of time near the vowel onset (Fox and Lehiste, 1987; Marcus, 1981). Like music, the p-center in speech is dependent on the "attack" of the consonants preceding the vowel onset (Cooper et al., 1986; Pompino-Marschall, 1989). More recent studies have used sensorimotor synchronization tasks such as tapping to looped speech to characterize that the vowel onset is a strong attractor of taps in speech and music (Lidji et al., 2011; Rathcke et al., 2021). More work is needed to characterize what acoustic factors contribute to the perceptual location of a beat in speech, song, and a range of natural environmental stimuli. If beat bins can be used to calculate integer-multiple relationships from stimulus features in complex stimuli such as speech and song, then beat bins could also be used to calculate integer-multiple relationships in a range of naturally occurring stimuli.

In summary, rhythmic regularity is an important feature for differentiating speech and song, children are sensitive to temporal features, but beat/rhythm are not critical for differentiating speech and song until later in development (see Section 6 for more about development). Finally, subjective ratings of rhythmic regularity reliably differentiate speech from song, although acoustic features (i.e., spectral flux, duration) that predict regularity ratings do not capture the integer relationships inherent to stimuli with a hierarchical beat structure. As outlined above, next steps include determining if psychologically informed metrics of rhythmic regularity are stronger predictors of regularity than easily extracted acoustic features, such as spectral flux or syllable duration. The long-term goal of this work is to characterize the degree of regularity in all sounds, including those beyond human communicative contexts, such as dog barks, typing, bird calls, water droplets, and sneezes. As such, acoustic metrics of regularity should be general enough to be applied in non-communicative contexts.

So why is it important to characterize the degree of regularity in speech, song, and environmental sounds? We now turn to the existing literature on neural and perceptual dynamics of rhythm to provide rationale for the important functions rhythmic regularity plays in attention, learning, and memory.

26.3 Attention Is Rhythmic

Already 60 years ago, Mari Reiss Jones outlined her work on dynamic attending theory (DAT) (Jones, 1976), which proposed that attention is not constant but cyclical, with alternations between moments of focused attention followed by inattention (Large and Jones, 1999). In this theory, external stimulation, such as music, can guide attention toward moments in time when

https://doi.org/10.1017/9781009295888.031 Published online by Cambridge University Press

important information or events have a high likelihood of occurrence. In a musical sequence, attention would be allocated to the strong beats, when important melodic or metrical information is likely to occur, and attention would wane on weaker or off-beat positions. If attention is allocated more often to the strong than to the weak beats, this should be evident as better processing and sequencing of information occurring on the beat than off. Indeed, people are better at remembering faces when they are presented on the beat of a musical rhythm compared to when they are presented off the beat or in silence (Johndro et al., 2019).

A stronger test of rhythmic facilitation through DAT is whether an external rhythm can start a cycle of focused attention and inattention at the rate of the rhythmic stimulus that continues even when the rhythm is no longer present. Several studies have shown that after a rhythmically regular prime, participants have better memory for information played after the prime, such as speech (Canette et al., 2020a; Cason and Schön, 2012; Cason et al., 2015; Chern et al., 2018; Przybylski et al., 2013) and visual information (Plancher et al., 2018), compared to an irregular prime or silence. Rhythmic attention may improve working memory performance by freeing up attentional resources by processing fewer, but highly relevant, elements in an ongoing stream of sensory input (Johndro et al., 2019). Music and speech signals themselves are designed to provide key details at rhythmically salient moments in time. Speech stress typically highlights semantically meaningful elements in a sentence (Mattys, 1997; Pitt and Samuel, 1990), and strong beats in music are more often populated with pitches that are crucial to the tonality of a piece of music (Prince and Schmuckler, 2014; Prince et al., 2020). Thus, not only do perceptual and attentional systems sample sensory input in a rhythmic manner, but the stimulus streams themselves are optimized for the listener. It remains to be seen whether human communicative sounds are uniquely structured to capitalize on information transfer at rhythmic moments in time or whether other animal vocalizations or environmental sounds are structured similarly (but see De Gregorio et al., 2021; Roeske et al., 2020).

26.4 Neural Correlates of Rhythmic Regularity

The type of rhythmic attention described in DAT is predicted by the cyclical excitatory-inhibitory oscillations for populations of neurons in the brain (Lakatos et al., 2008; Large and Jones, 1999). A period of heightened sensitivity at the peak of an excitatory phase of a neural oscillation would explain better perceptual or memory outcomes described in the studies reported above (see also Henry and Herrmann, 2014). Indeed, change sensitivity is related to the phase of oscillations aligned with the external rhythmic sensory input (e.g., Henry and Obleser, 2012). Not all neural tracking of sensory input can be described as neural entrainment (Haegens, 2018; Haegens and Golumbic, 2018; Kösem et al., 2018; Obleser and Kayser, 2019; Zoefel et al., 2018), as neural tracking of stimulus rhythms could be evidence of true neural entrainment (e.g., oscillations continue even in the absence of stimulation) or stimulus-driven responses to sound onsets (e.g., N1-P2 onset responses). Regardless, the degree of phase alignment between brain and stimulus rhythms predicts behavioral performance, including speech comprehension (Peelle, 2012), attention (Zion-Golumbic and Schroeder, 2012), and expertise (Doelling and Poeppel, 2015; Harding et al., 2019). For more on this topic, see Chapters 3 and 5.

https://doi.org/10.1017/9781009295888.031 Published online by Cambridge University Press

Neural oscillations are described as inherently rhythmically regular, ascribing to a single frequency of oscillation (e.g., Zoefel et al., 2018; but see Haas and Kubin, 1998, for nonlinear multi-frequency oscillators capable of entraining to the irregularities of speech). If oscillators are biased toward isochrony, then the alignment of ongoing neural oscillations should be better for a stimulus that is also rhythmically regular (e.g., less phase resetting of ongoing oscillations to match the incoming stimulus). Since song is more rhythmically regular than speech (Yu et al., 2023), it follows that song should garner greater alignment with ongoing neural activity – greater neural tracking – than speech. Consistent with this assertion, neural tracking is reduced when speech rhythms are made less regular by inserting pauses (delta band; Kayser et al., 2015), and neural tracking is increased when words are sung over spoken (theta band; Vanden Bosch der Nederlanden et al., 2020, 2022a). Thus, it is possible that neural tracking increases linearly with the degree of regularity (see Section 6 chapters for more about rhythm and learning).

Here, we report a reanalysis that characterizes how subjective rhythmic regularity relates to neural tracking. We reanalyzed magnetoencephalography (MEG) data from the Vanden Bosch der Nederlanden et al. (2022a) dataset (available here: https://data.donders.ru.nl/) and related it with the regularity ratings of the same stimulus set of Yu et al. (2023) (available here: https://osf.io/hnw5t/) to elucidate whether the degree of regularity predicts neural tracking at the syllable rate across both spoken and sung utterances. The MEG data was collected as part of the melody familiarity study, where participants listened to 96 stimuli (presented four times) and rated whether the melody of the spoken or sung stimulus was one they had learned during their training sessions (they learned half of the sung melodies without lyrics). We used a jackknifing technique (Richter et al., 2015) to extract single-trial estimates of neural tracking for each stimulus (96 stimuli, four presentations) presented to each of their 32 participants and related these stimulus-level data with the average regularity rating (see Section 26.2 – subjective ratings of ease of clapping or tapping to the stimulus) provided for each stimulus by 51 participants. Jackknifing estimates use a leave-one-out approach, so if a particular stimulus contributed strongly to the overall estimate of neural tracking, there would be a large reduction in the jackknifed neural tracking estimate when the stimulus was left out. If that same stimulus was also rated high in rhythmic regularity, then a significant negative correlation between neural tracking and rhythmic regularity would mean that greater neural tracking is related to increases in subjective experiences of rhythmic regularity.

We used a linear mixed-effects model (R, lmer package; stimulus and participant as random effects) to predict neural tracking based on acoustic features and subjective regularity ratings described in Yu et al. (2023). We found that after controlling for utterance type (speech versus song), neural tracking was predicted by subjective rhythmic regularity ratings and pulse clarity (minimum value, MIR Toolbox; see Figure 26.2). Interestingly, although spectral flux is a stronger predictor of neural tracking than amplitude envelope measures (Weineck et al., 2022), it did not explain unique variance beyond rhythmic regularity ratings (see Table 26.2). This reanalysis suggests that the degree of rhythmic regularity in a stimulus is a strong driver of neural tracking, even beyond acoustic features that are highly predictive of rhythmic regularity (i.e., spectral flux).

These results may lend a neural explanation for recent studies showing that attention is biased toward rhythmic regularity, in both auditory (Andreou et al., 2011) and visual domains

https://doi.org/10.1017/9781009295888.031 Published online by Cambridge University Press

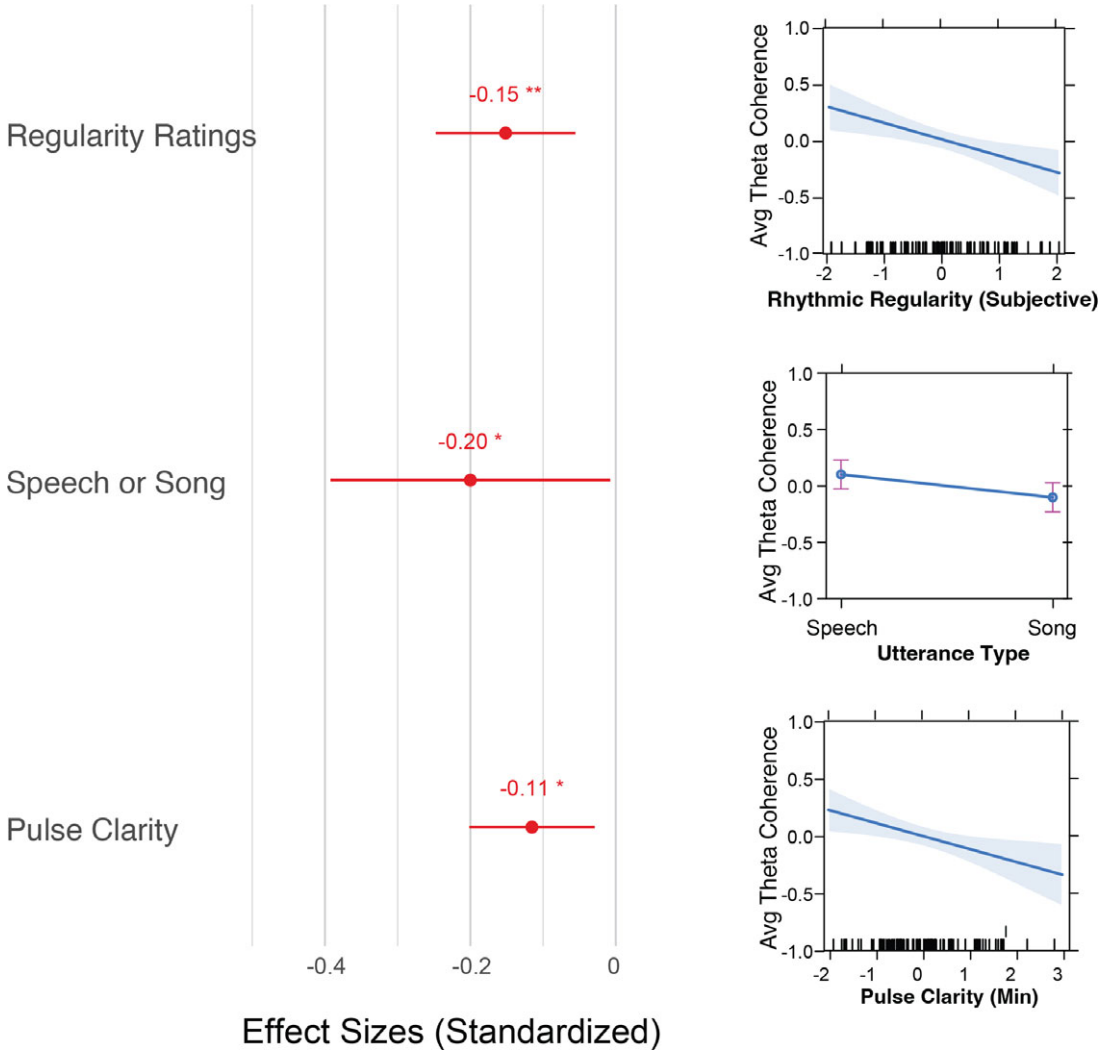


Figure 26.2 Linear mixed-effects regression results for neural and acoustic data.
Neural tracking (speech-brain coherence in the theta band – 4–8 Hz) is related to rhythmic regularity and pulse clarity, even after controlling for utterance type.

(Zhao et al., 2013), even without explicit awareness (Zhao et al., 2013). Perhaps, stimuli that capture attention are those with greater rhythmic regularity because they are easier to process neurally than less regular or irregular stimuli. Further, if regular stimuli capture attention, then the regularity of the stimulus would also make it more likely that important moments in time align with moments of high neural excitability, resulting in better memory or comprehension. Although speech is not rhythmically regular, like music, it likely has a higher degree of regularity than other real-world sounds, such as a babbling brook. While regularity may move stimuli from one side of the speech–song continuum to the other, regularity may be critical for a broader range of

https://doi.org/10.1017/9781009295888.031 Published online by Cambridge University Press

Table 26.2 Linear mixed-effects models for neural and acoustic data

Results from the linear mixed-effects regression models predicting average theta-band cerebro-acoustic phase coherence (raw values), with stimulus and participant as random effects. Statistically significant variables in each model are highlighted in bold.

Model	*Variable*	*Estimate*	*t-value*	*p*
Model 1	**Regularity**	**−0.191**	**−4.145**	**<.001**
X^2(5, N=3072) = 16.113, p<.001, AIC = 8159.3 *(compared to random intercept model)*				
Model 2	**Regularity**	**−0.142**	**−2.821**	**0.006**
	Song	**−0.220**	**−2.183**	**0.032**
X^2(6, N=3072) = 4.798, p=.0285, AIC = 8156.5 *(compared to Model 1)*				
Model 3	**Regularity**	**−0.138**	**−2.703**	**0.036**
	Song	**−0.365**	**−2.496**	**0.008**
	Spectral flux	0.052	0.693	0.490
	F0 stability	−0.098	−1.251	0.215
	Pulse clarity	**−0.116**	**−2.528**	**0.013**
	Syllable duration	0.031	0.507	0.613
	Prop. small integer ratio	0.034	0.729	0.468
	Consonant PVI	−0.034	−0.641	0.523
	% vocalic	0.083	1.468	0.146
X^2(13, N=3072) = 11.662, p=.1122 AIC = 8158.8 *(compared to Model 2)*				
Model 4	**Regularity**	**−0.146**	**−2.974**	**0.004**
	Song	**−0.203**	**−2.077**	**0.041**
	Pulse clarity	**−0.113**	**−2.568**	**0.012**
X^2(7, N=3072) = 6.648, p=.0099, AIC = 8151.8 *(compared to Model 2)*				

noncommunicative stimuli forming a much larger acoustic continuum. The periodicity of speech may also underlie why people of all ages are nearly perfect at detecting when a human voice changes in a complex scene compared to musical instrument sounds, environmental sounds, or animal vocalizations (Vanden Bosch der Nederlanden et al., 2016, 2018; Vanden Bosch der Nederlanden and Vouloumanos et al., 2021). Future work should examine whether attentional biases toward real-world sounds could be predicted based on the degree of regularity in the signal, such that the stimulus with the highest degree of regularity is the one that captures attention. Together, this review and reanalysis suggest that the human brain is endowed with neural mechanisms specialized to process rhythmic regularities across multiple domains.

26.5 Leveraging Rhythmic Regularity in Music

A growing number of studies show relationships between musical rhythmic processing and language (e.g., Gordon et al., 2015; Kertész and Honbolygó, 2023; Ozernov-Palchik and Patel, 2018; Woodruff Carr et al., 2014). They show, for instance, that kindergarteners' rhythmic discrimination skills and letter-sound knowledge – an early predictor of reading development – are highly

https://doi.org/10.1017/9781009295888.031 Published online by Cambridge University Press

correlated (Ozernov-Palchik and Patel, 2018), that drumming consistency relates to language outcomes (Woodruff Carr et al., 2014), and that rhythm perception in elementary school-aged children predicts expressive grammar skills (Nitin et al., 2023). Deficits in rhythmic processing abilities have also been connected to a wide range of developmental disorders, including dyslexia, ADHD, autism, motor coordination disorders, and stuttering (Fiveash et al., 2021; Garnett et al., 2022; Hannon et al., 2018; Ladanyi et al., 2020; Lense et al., 2021). Compared to typically developing children, those with reading disorders and language impairment struggle to synchronize their movements to a metronome or musical rhythm (Corriveau and Goswami, 2009). The ability to extract a stable pulse or perceive rhythmic regularity may be crucial to successful language processing (Bekius et al., 2016; Centanni et al., 2018, 2022).

Given that language deficits are related to musical beat processing, music-based interventions that help people pick up on rhythmic regularity can complement existing therapies to increase the positive impact on language outcomes. As described above, rhythmic priming studies successfully use musical primes to entrain neural activity and facilitate the extraction of relevant information from subsequently presented speech stimuli (Canette et al., 2020b; Chern et al., 2018). Extra scaffolding and priming of rhythmic attention through music may be particularly beneficial for children with developmental disorders who struggle to create internal representations of salient moments in ongoing speech streams. This is consistent with reports that rhythm-based musical training can benefit reading and phonological processes in dyslexic children (Flaugnacco et al., 2015) and speech processing for children with cochlear implants (Bedoin et al., 2018). Beyond the potential benefits in language processing, rhythm in music is also tightly related to the experience of pleasure and reward (Fiveash et al., 2023; Todd and Lee, 2015; Zatorre, 2015). This makes musical interventions focused on rhythmic regularity especially fruitful for improving perceptual as well as socio-emotional outcomes across the lifespan. Musical features such as rhythmic regularity can be leveraged to improve neural tracking and downstream behavioral outcomes, such as comprehension or engagement.

Box 26.1 Chapter Overview

Summary

Rhythmic regularity is a key differentiator of speech and song, but how it is measured in the acoustic signal is not trivial. Listeners subjectively report rhythmic regularity in song, but the acoustic features that are consistent with regularity do not correspond with the small integer ratios of events that align with a beat. Rhythmic regularity is a significant modulator of neural tracking in speech and song, suggesting regularity likely plays a crucial role in guiding other behavioral outcomes related to neural tracking, such as attention, memory, engagement, and even comprehension.

Implications

Our work suggests that regularity drives neural tracking and is a key factor differentiating speech and song. Given the neural tracking findings, would regularizing the speech

https://doi.org/10.1017/9781009295888.031 Published online by Cambridge University Press

Box 26.1 (cont.)

signal to be more music-like improve intelligibility and comprehension compared to typical, irregular, speech rhythms? If this is the case, then why did humans evolve two distinct systems of communication that differ in the degree of regularity and their functional goals: speech for transaction of meaning, and song for conveying or regulating emotion? That is, why do we not exclusively sing to one another? Future work should characterize the role of rhythmic regularity across multiple domains, including speech, song, and other everyday sounds.

Gains

This chapter presents a first step toward understanding the role of regularity across communicative modalities and characterizing its impact on neural processing. We highlight the necessity for a good metric of the degree of regularity that can characterize a range of naturally occurring stimuli. Such a metric has the potential to have broad impact, for example on the best practices for teaching and learning in the classroom to improve retention, or in high-risk situations to grab listeners' attention, as in hospital-alert settings.

REFERENCES

Albouy, P., Mehr, S. A., Hoyer, R. S., Ginzburg, J., and Zatorre, R. J. (2024). Spectro-temporal acoustical markers differentiate speech from song across cultures. *Nature Communications*, 15, 4835. https://doi.org/10.1038/s41467-024-49040-3

Alluri, V., and Toiviainen, P. (2010). Exploring perceptual and acoustical correlates of polyphonic timbre. *Music Perception*, 27(3), 223–242.

Andreou, L. V., Kashino, M., and Chait, M. (2011). The role of temporal regularity in auditory segregation. *Hearing Research*, 280(1–2), 228–235.

Appen, R. V., Doehring, A., and Moore, A. F. (2015). *Song Interpretation in 21st-Century Pop Music* (first edition). Routledge. https://doi.org/10.4324/9781315609881

Arvaniti, A. (2012). The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics*, 40(3), 351–373.

Bedoin, N., Besombes, A. M., Escande, É., et al. (2018). Boosting syntax training with temporally regular musical primes in children with cochlear implants. *Annals of Physical and Rehabilitation Medicine*, 61(6), 365–371.

Beier, E. J., and Ferreira, F. (2018). The temporal prediction of stress in speech and its relation to musical beat perception. *Frontiers in Psychology*, 9, 431.

Bekius, A., Cope, T. E., and Grube, M. (2016). The beat to read: A cross-lingual link between rhythmic regularity perception and reading skill. *Frontiers in Human Neuroscience*, 10, 425

Breen, M., and Clifton Jr, C. (2011). Stress matters: Effects of anticipated lexical stress on silent reading. *Journal of Memory and Language*, 64(2), 153–170.

https://doi.org/10.1017/9781009295888.031 Published online by Cambridge University Press

Brown, S. (2000). The "musilanguage" model of musical evolution. In N. L. Wallin, B. Merker, and S. Brown (Eds.), *The Origins of Music*, pp. 271–300. MIT Press. https://doi.org/10.7551/mitpress/5190.001.0001

Burger, B., Thompson, M. R., Luck, G., Saarikallio, S., and Toiviainen, P. (2013). Influences of rhythm- and timbre-related musical features on characteristics of music-induced movement. *Frontiers in Psychology*, 4, 183.

Burger, B., Thompson, M. R., Luck, G., Saarikallio, S. H., and Toiviainen, P. (2014). Hunting for the beat in the body: On period and phase locking in music-induced movement. *Frontiers in Human Neuroscience*, 8, 903.

Canette, L. H., Fiveash, A., Krzonowski, J., et al. (2020a). Regular rhythmic primes boost P600 in grammatical error processing in dyslexic adults and matched controls. *Neuropsychologia*, 138, 107324.

Canette, L. H., Lalitte, P., Bedoin, N., et al. (2020b). Rhythmic and textural musical sequences differently influence syntax and semantic processing in children. *Journal of Experimental Child Psychology*, 191, 104711.

Cason, N., and Schön, D. (2012). Rhythmic priming enhances the phonological processing of speech. *Neuropsychologia*, 50(11), 2652–2658.

Cason, N., Astésano, C., and Schön, D. (2015). Bridging music and speech rhythm: Rhythmic priming and audio-motor training affect speech perception. *Acta Psychologica*, 155, 43–50.

Centanni, T. M., Beach, S. D., Ozernov-Palchik, O., et al. (2022). Categorical perception and influence of attention on neural consistency in response to speech sounds in adults with dyslexia. *Annals of Dyslexia*, 72, 56–78. https://doi.org/10.1007/s11881-021-00241-1

Centanni, T. M., Pantazis, D., Truong, D. T., et al. (2018). Increased variability of stimulus-driven cortical responses is associated with genetic variability in children with and without dyslexia. *Developmental Cognitive Neuroscience*, 34, 7–17.

Chern, A., Tillmann, B., Vaughan, C., and Gordon, R. L. (2018). New evidence of a rhythmic priming effect that enhances grammaticality judgments in children. *Journal of Experimental Child Psychology*, 173, 371–379.

Cooper, A. M., Whalen, D. H., and Fowler, C. A. (1986). P-centers are unaffected by phonetic categorization. *Perception & Psychophysics*, 39(3), 187–196.

Corriveau, K. H., and Goswami, U. (2009). Rhythmic motor entrainment in children with speech and language impairments: Tapping to the beat. *Cortex*, 45(1), 119–130.

Cutler, A., and Clifton Jr, C. E. (1984). The use of prosodic information in word recognition. In H. Bouma and D. G. Bouwhuis (Eds.), *Attention and Performance X: Control of Language Processes*, pp. 183–196. Erlbaum.

Cutler, A., and Foss, D. J. (1977). On the role of sentence stress in sentence processing. *Language and Speech*, 20(1), 1–10.

Danielsen, A., Haugen, M. R., and Jensenius, A. R. (2015). Moving to the beat: Studying entrainment to micro-rhythmic changes in pulse by motion capture. *Timing & Time Perception*, 3(1–2), 133–154.

Danielsen, A., Nymoen, K., Anderson, E., et al. (2019). Where is the beat in that note? Effects of attack, duration, and frequency on the perceived timing of musical and quasi-musical sounds. *Journal of Experimental Psychology: Human Perception and Performance*, 45(3), 402.

Datseris, G., Ziereis, A., Albrecht, T., et al. (2019). Microtiming deviations and swing feel in jazz. *Scientific Reports*, 9(1), 19824.

De Gregorio, C., Valente, D., Raimondi, T., et al. (2021). Categorical rhythms in a singing primate. *Current Biology*, 31(20), R1379–R1380.

De Meo, R., Matusz, P. J., Knebel, J. F., et al. (2016). What makes medical students better listeners? *Current Biology*, 26(13), R519–R520.

https://doi.org/10.1017/9781009295888.031 Published online by Cambridge University Press

Ding, N., Patel, A. D., Chen, L., et al. (2017). Temporal modulations in speech and music. *Neuroscience & Biobehavioral Reviews*, 81, 181–187.

Doelling, K. B., and Poeppel, D. (2015). Cortical entrainment to music and its modulation by expertise. *Proceedings of the National Academy of Sciences*, 112(45), E6233–E6242.

Drake, C., and Palmer, C. (1993). Accent structures in music performance. *Music Perception*, 10(3), 343–378.

Fitch, W. T. (2006). The biology and evolution of music: A comparative perspective. *Cognition*, 100(1), 173–215

Fiveash, A., Bedoin, N., Gordon, R. L., and Tillmann, B. (2021). Processing rhythm in speech and music: Shared mechanisms and implications for developmental speech and language disorders. *Neuropsychology*, 35(8), 771.

Fiveash, A., Ferreri, L., Bouwer, F. L., et al. (2023). Can rhythm-mediated reward boost learning, memory, and social connection? Perspectives for future research. *Neuroscience & Biobehavioral Reviews*, 149, 105153.

Flaugnacco, E., Lopez, L., Terribili, C., et al. (2015). Music training increases phonological awareness and reading skills in developmental dyslexia: A randomized control trial. *PLoS ONE*, 10(9), e0138715.

Fox, R. A., and Lehiste, I. (1987). The effect of vowel quality variations on stress-beat location. *Journal of Phonetics*, 15(1), 1–13.

Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America*, 27(4), 765–768.

Fry, D. B. (1958). Experiments in the perception of stress. *Language and Speech*, 1(2), 126–152.

Garnett, E. O., Chow, H. M., Limb, S., Liu, Y., and Chang, S. E. (2022). Neural activity during solo and choral reading: A functional magnetic resonance imaging study of overt continuous speech production in adults who stutter. *Frontiers in Human Neuroscience*, 16, 894676.

Gordon, R. L., Shivers, C. M., Wieland, E. A., et al. (2015). Musical rhythm discrimination explains individual differences in grammar skills in children. *Developmental Science*, 18(4), 635–644.

Grabe, E., and Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. *Papers in Laboratory Phonology*, 7(1982), 515–546.

Haas, H., and Kubin, G. (1998). A multi-band nonlinear oscillator model for speech. Conference Record of Thirty-Second Asilomar Conference on Signals, Systems and Computers (Cat. No.98CH36284), 1, pp. 338–342. https://doi.org/10.1109/ACSSC.1998.750882

Haegens, S. (2018). Entrainment revisited: A commentary on Meyer, Sun & Martin (2020). *Language, Cognition and Neuroscience*, 35(9), 1119–1123. https://doi.org/10.1080/23273798.2020.1758335

Haegens, S., and Golumbic, E. Z. (2018). Rhythmic facilitation of sensory processing: A critical review. *Neuroscience & Biobehavioral Reviews*, 86, 150–165.

Hannon, E. E., Nave-Blodgett, J. E., and Nave, K. M. (2018). The developmental origins of the perception and production of musical rhythm. *Child Development Perspectives*, 12(3), 194–198.

Harding, E. E., Sammler, D., Henry, M. J., Large, E. W., and Kotz, S. A. (2019). Cortical tracking of rhythm in music and speech. *NeuroImage*, 185, 96–101.

Hayes, B. (1985, June). Iambic and trochaic rhythm in stress rules. Annual Meeting of the Berkeley Linguistics Society, Vol. 11, pp. 429–446.

Henry, M. J., and Herrmann, B. (2014). Low-frequency neural oscillations support dynamic attending in temporal context. *Timing & Time Perception*, 2(1), 62–86.

Henry, M. J., and Obleser, J. (2012). Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. *Proceedings of the National Academy of Sciences*, 109(49), 20095–20100.

Henry, M. J., Herrmann, B., and Grahn, J. A. (2017). What can we learn about beat perception by comparing brain signals and stimulus envelopes? *PLoS ONE*, 12(2), e0172454.

https://doi.org/10.1017/9781009295888.031 Published online by Cambridge University Press

Hilton, C. B., Moser, C. J., Bertolo, M., et al. (2022). Acoustic regularities in infant-directed speech and song across cultures. *Nature Human Behaviour*, 6(11), 1545–1556. https://doi.org/10.1038/s41562-022-01410-x

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

Jacoby, N., and McDermott, J. H. (2017). Integer ratio priors on musical rhythm revealed cross-culturally by iterated reproduction. *Current Biology*, 27(3), 359–370.

Jadoul, Y., Ravignani, A., Thompson, B., Filippi, P., and De Boer, B. (2016). Seeking temporal predictability in speech: Comparing statistical approaches on 18 world languages. *Frontiers in Human Neuroscience*, 10, 586.

Johndro, H., Jacobs, L., Patel, A. D., and Race, E. (2019). Temporal predictions provided by musical rhythm influence visual memory encoding. *Acta Psychologica*, 200, 102923.

Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, 83(5), 323.

Jones, R. (1985). A dialectic analysis of selected contradictions among definitions of meter in music. In *Bulletin of the Council for Research in Music Education, No. 83*, pp. 43–56. University of Illinois Press.

Kayser, S. J., Ince, R. A., Gross, J., and Kayser, C. (2015). Irregular speech rate dissociates auditory cortical entrainment, evoked responses, and frontal alpha. *Journal of Neuroscience*, 35(44), 14691–14701.

Kertész, C., and Honbolygó, F. (2023). First school year tapping predicts children's third-grade literacy skills. *Scientific Reports*, 13(1), 2298.

Klatt, D. H. (1975). Vowel lengthening is syntactically determined in a connected discourse. *Journal of Phonetics*, 3(3), 129–140.

Kösem, A., Bosker, H. R., Takashima, A., et al. (2018). Neural entrainment determines the words we hear. *Current Biology*, 28(18), 2867–2875.

Ladányi, E., Persici, V., Fiveash, A., Tillmann, B., and Gordon, R. L. (2020). Is atypical rhythm a risk factor for developmental speech and language disorders? *Wiley Interdisciplinary Reviews: Cognitive Science*, 11(5), e1528.

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., and Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science*, 320(5872), 110–113.

Large, E. W. (2008). Resonating to musical rhythm: Theory and experiment. In S. Grondin (Ed.), *The Psychology of Time*, pp. 189–231. Emerald Group Publishing.

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

Lartillot, O., Eerola, T., Toiviainen, P., and Fornari, J. (2008, September). Multi-feature modeling of pulse clarity: Design, validation and optimization. ISMIR 2008, Ninth International Conference on Music Information Retrieval, Drexel University, Philadelphia, PA, USA, September 14–18, 2008, pp. 521–526.

Lense, M. D., Ladányi, E., Rabinowitch, T. C., Trainor, L., and Gordon, R. (2021). Rhythm and timing as vulnerabilities in neurodevelopmental disorders. *Philosophical Transactions of the Royal Society B*, 376(1835), 20200327.

Lerdahl, F., and Jackendoff, R. (1983). An overview of hierarchical structure in music. *Music Perception*, 1(2), 229–252.

Levey, S., and Raphael, L. J. (2002). Stress clash: Frequency and strategies of resolution. *Journal of the Acoustical Society of America*, 111(5), 2476–2476.

Lidji, P., Palmer, C., Peretz, I., and Morningstar, M. (2011). Listeners feel the beat: Entrainment to English and French speech rhythms. *Psychonomic Bulletin & Review*, 18, 1035–1041.

https://doi.org/10.1017/9781009295888.031 Published online by Cambridge University Press

London, J. (2004). *Hearing in Time: Psychological Aspects of Musical Meter*. Oxford University Press.
Marcus, S. M. (1981). Acoustic determinants of perceptual center (P-center) location. *Perception & Psychophysics*, 30, 247–256.
Matthews, T. E., Witek, M. A., Heggli, O. A., Penhune, V. B., and Vuust, P. (2019). The sensation of groove is affected by the interaction of rhythmic and harmonic complexity. *PLoS ONE*, 14(1), e0204539.
Mattys, S. L. (1997). The use of time during lexical processing and segmentation: A review. *Psychonomic Bulletin & Review*, 4(3), 310–329.
McAuley, J. D. (2010). Tempo and rhythm. In M. Riess Jones, R. Fay, and A. Popper (Eds.), *Music Perception*, pp. 165–199. Springer.
Merchant, H., and Honing, H. (2014). Are non-human primates capable of rhythmic entrainment? Evidence for the gradual audiomotor evolution hypothesis. *Frontiers in Neuroscience*, 7, 274. https://doi.org/10.3389/fnins.2013.00274
Morton, J., Marcus, S., and Frankish, C. (1976). Perceptual centers (P-centers). *Psychological Review*, 83(5), 405.
Nave-Blodgett, J. E., Snyder, J. S., and Hannon, E. E. (2021). Hierarchical beat perception develops throughout childhood and adolescence and is enhanced in those with musical training. *Journal of Experimental Psychology: General*, 150(2), 314.
Nave, K. M., Snyder, J. S., and Hannon, E. E. (2023). Sustained musical beat perception develops into late childhood and predicts phonological abilities. *Developmental Psychology*, 59(5), 829–844.
Nitin, R., Gustavson, D. E., Aaron, A. S., et al. (2023). Exploring individual differences in musical rhythm and grammar skills in school-aged children with typically developing language. *Scientific Reports*, 13(1), 2201. https://doi.org/10.1038/s41598-022-21902-0
Nolan, F., and Jeon, H. S. (2014). Speech rhythm: A metaphor? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1658), 20130396.
Obleser, J., and Kayser, C. (2019). Neural entrainment and attentional selection in the listening brain. *Trends in Cognitive Sciences*, 23(11), 913–926. https://doi.org/10.1016/j.tics.2019.08.004
Ozaki, Y., Tierney, A., and Pfordresher, P. Q. (2024). Globally, songs and instrumental melodies are slower and higher and use more stable pitches than speech: A registered report. *Science Advances*, 10(20), eadm9797. https://doi.org/10.1126/sciadv.adm9797
Ozernov-Palchik, O., and Patel, A. D. (2018). Musical rhythm and reading development: Does beat processing matter? *Annals of the New York Academy of Sciences*, 1423(1), 166–175.
Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, 11(4), 409–464.
Patel, A. D. (2010). *Music, Language, and the Brain*. Oxford University Press.
Peelle, J. E. (2012). The hemispheric lateralization of speech processing depends on what "speech" is: A hierarchical perspective. *Frontiers in Human Neuroscience*, 6, 309.
Peperkamp, S., Vendelin, I., and Dupoux, E. (2010). Perception of predictable stress: A cross-linguistic investigation. *Journal of Phonetics*, 38(3), 422–430. https://doi.org/10.1016/j.wocn.2010.04.001
Pitt, M. A., and Samuel, A. G. (1990). The use of rhythm in attending to speech. *Journal of Experimental Psychology: Human Perception and Performance*, 16(3), 564.
Plancher, G., Lévêque, Y., Fanuel, L., Piquandet, G., and Tillmann, B. (2018). Boosting maintenance in working memory with temporal regularities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(5), 812.
Pompino-Marschall, B. (1989). On the psychoacoustic nature of the P-center phenomenon. *Journal of Phonetics*, 17(3), 175–192.
Povel, D. J., and Essens, P. J. (1985). Perception of temporal patterns. *Music Perception*, 2(4), 411–440.
Prince, J. B., and Schmuckler, M. A. (2014). The tonal-metric hierarchy: A corpus analysis. *Music Perception*, 31, 254–270.

https://doi.org/10.1017/9781009295888.031 Published online by Cambridge University Press

Prince, J. B., Tan, S. E. J., and Schmuckler, M. A. (2020). Listeners perceive complex pitch-temporal structure in melodies. *Memory & Cognition*, 48, 526–540.

Przybylski, L., Bedoin, N., Krifi-Papoz, S., et al. (2013). Rhythmic auditory stimulation influences syntactic processing in children with developmental language disorders. *Neuropsychology*, 27(1), 121–131.

Rasch, R. A. (1988). Timing and synchronization in ensemble performance. In J. Sloboda (Ed.), *Generative Processes in Music: The Psychology of Performance, Improvisation, and Composition*, pp. 70–90. Oxford University Press.

Rathcke, T., Lin, C. Y., Falk, S., and Dalla Bella, S. (2021). Tapping into linguistic rhythm. *Laboratory Phonology*, 12(1), 11. https://doi.org/10.5334/labphon.248

Ravignani, A., and Madison, G. (2017). The paradox of isochrony in the evolution of human rhythm. *Frontiers in Psychology*, 8, 1820.

Ravignani, A., and Norton, P. (2017). Measuring rhythmic complexity: A primer to quantify and compare temporal structure in speech, movement, and animal vocalizations. *Journal of Language Evolution*, 2(1), 4–19.

Richter, C. G., Thompson, W. H., Bosman, C. A., and Fries, P. (2015). A jackknife approach to quantifying single-trial correlation between covariance-based metrics undefined on a single-trial basis. *NeuroImage*, 114, 57–70.

Roeske, T. C., Tchernichovski, O., Poeppel, D., and Jacoby, N. (2020). Categorical rhythms are shared between songbirds and humans. *Current Biology*, 30(18), 3544–3555.

Savage, P. E., Loui, P., Tarr, B., et al. (2021). Music as a coevolved system for social bonding. *Behavioral and Brain Sciences*, 44, e59.

Seidl, A., French, B., Wang, Y., and Cristia, A. (2014). Toward establishing continuity in linguistic skills within early infancy. *Language Learning*, 64(s2), 165–183.

Selkirk, E. O. (1980). The role of prosodic categories in English word stress. *Linguistic Inquiry*, 11(3), 563–605.

Selkirk, E. O. (1995). Sentence prosody: Intonation, stress, and phrasing. *The Handbook of Phonological Theory*, 1, 550–569.

Selkirk, E.O. (1984). *Phonology and Syntax: The Relation between Sound and Structure*, pp. xv–476. MIT Press.

Senn, O., Kilchenmann, L., Von Georgi, R., and Bullerjahn, C. (2016). The effect of expert performance microtiming on listeners' experience of groove in swing or funk music. *Frontiers in Psychology*, 7, 1487.

Snyder, J., and Krumhansl, C. L. (2001). Tapping to ragtime: Cues to pulse finding. *Music Perception*, 18(4), 455–489.

Spiech, C., Endestad, T., Laeng, B., Danielsen, A., and Haghish, E. F. (2023). Beat alignment ability is associated with formal musical training not current music playing. *Frontiers in Psychology*, 14, 1034561.

Temperley, D. (2004). *The Cognition of Basic Musical Structures*. MIT Press.

Todd, N. P., and Lee, C. S. (2015). The sensory-motor theory of rhythm and beat induction 20 years on: A new synthesis and future perspectives. *Frontiers in Human Neuroscience*, 9, 444.

Trainor, L. J., and Cirelli, L. (2015). Rhythm and interpersonal synchrony in early social development. *Annals of the New York Academy of Sciences*, 1337(1), 45–52.

Turk, A., and Shattuck-Hufnagel, S. (2013). What is speech rhythm? A commentary on Arvaniti and Rodriquez, Krivokapić, and Goswami and Leong. *Laboratory Phonology*, 4(1), 93–118.

Ullal-Gupta, S., Vanden Bosch der Nederlanden, C. M., Tichko, P., Lahav, A., and Hannon, E. E. (2013). Linking prenatal experience to the emerging musical mind. *Frontiers in Systems Neuroscience*, 7, 48.

Vanden Bosch der Nederlanden, C. M., and Vouloumanos, A. (2021). Infant biases for detecting speech in complex scenes. Developmental Psychology, 57(9), 1411.

https://doi.org/10.1017/9781009295888.031 Published online by Cambridge University Press

Vanden Bosch der Nederlanden, C. M., Joanisse, M. F., and Grahn, J. A. (2020). Music as a scaffold for listening to speech: Better neural phase-locking to song than speech. *NeuroImage*, 214, 116767.

Vanden Bosch der Nederlanden, C. M., Snyder, J. S., and Hannon, E. E. (2016). Children use object-level category knowledge to detect changes in complex auditory scenes. *Developmental Psychology*, 52(11), 1867.

Vanden Bosch der Nederlanden, C. M., Joanisse, M. F., Grahn, J. A., Snijders, T. M., and Schoffelen, J. M. (2022a). Familiarity modulates neural tracking of sung and spoken utterances. *NeuroImage*, 252, 119049.

Vanden Bosch der Nederlanden, C.M., Zaragoza, C., Rubio-Garcia, A., Clarkson, E., and Snyder, J.S. (2018). Change detection in complex auditory scenes is predicted by auditory memory, pitch perception, and years of musical training. *Psychological Research*, 84, 585–601. https://doi.org/10.1007/s00426-018-1072-x

Vanden Bosch der Nederlanden, C. M., Qi, X., Sequeira, S., et al. (2022b). Developmental changes in the categorization of speech and song. *Developmental Science*, 26(5), e13346. https://doi.org/10.1111/desc.13346

Villing, R. (2010). Hearing the moment: Measures and models of the perceptual centre. Doctoral dissertation, National University of Ireland Maynooth.

Weineck, K., Wen, O. X., and Henry, M. J. (2022). Neural synchronization is strongest to the spectral flux of slow music and depends on familiarity and beat salience. *ELife*, 11, e75515.

White, L., and Mattys, S. L. (2007). Calibrating rhythm: First language and second language studies. *Journal of Phonetics*, 35(4), 501–522.

Wiget, L., White, L., Schuppler, B., et al. (2010). How stable are acoustic metrics of contrastive speech rhythm? *Journal of the Acoustical Society of America*, 127(3), 1559–1569.

Woodruff Carr, K., White-Schwoch, T., Tierney, A. T., Strait, D. L., and Kraus, N. (2014). Beat synchronization predicts neural speech encoding and reading readiness in preschoolers. *Proceedings of the National Academy of Sciences*, 111(40), 14559–14564.

Yu, C. Y., Cabildo, A., Grahn, J. A., and Vanden Bosch der Nederlanden, C. M. (2023). Perceived rhythmic regularity is greater for song than speech: Examining acoustic correlates of rhythmic regularity in speech and song. *Frontiers in Psychology*, 14, 1167003.

Zatorre, R. J. (2015). Musical pleasure and reward: Mechanisms and dysfunction. *Annals of the New York Academy of Sciences*, 1337(1), 202–211.

Zion-Golumbic, E., and Schroeder, C. E. (2012). Attention modulates "speech-tracking" at a cocktail party. *Trends in Cognitive Sciences*, 16(7), 363–364.

Zhao, J., Al-Aidroos, N., and Turk-Browne, N. B. (2013). Attention is spontaneously biased toward regularities. *Psychological Science*, 24(5), 667–677.

Zoefel, B., Archer-Boyd, A., and Davis, M. H. (2018). Phase entrainment of brain oscillations causally modulates neural responses to intelligible speech. *Current Biology*, 28(3), 401–408.

https://doi.org/10.1017/9781009295888.031 Published online by Cambridge University Press