



**HAL**  
open science

# Musical Training for Auditory Rehabilitation in Hearing Loss

Jacques Pesnot Lerousseau, Céline Hidalgo, Daniele Schön

► **To cite this version:**

Jacques Pesnot Lerousseau, Céline Hidalgo, Daniele Schön. Musical Training for Auditory Rehabilitation in Hearing Loss. *Journal of Clinical Medicine*, 2020, 9 (4), pp.1058. 10.3390/jcm9041058 . hal-03354598

**HAL Id: hal-03354598**

**<https://hal.science/hal-03354598>**

Submitted on 25 Sep 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Opinion

# Musical Training for Auditory Rehabilitation in Hearing Loss

Jacques Pesnot Lerousseau \*, Céline Hidalgo and Daniele Schön

Institut de Neurosciences des Systèmes, Aix Marseille University, Inserm, INS, 13005 Marseille, France; celine.hidalgo@univ-amu.fr (C.H.); daniele.schon@univ-amu.fr (D.S.)

\* Correspondence: jacques.pesnot-lerousseau@univ-amu.fr

Received: 16 February 2020; Accepted: 6 April 2020; Published: 8 April 2020



**Abstract:** Despite the overall success of cochlear implantation, language outcomes remain suboptimal and subject to large inter-individual variability. Early auditory rehabilitation techniques have mostly focused on low-level sensory abilities. However, a new body of literature suggests that cognitive operations are critical for auditory perception remediation. We argue in this paper that musical training is a particularly appealing candidate for such therapies, as it involves highly relevant cognitive abilities, such as temporal predictions, hierarchical processing, and auditory-motor interactions. We review recent studies demonstrating that music can enhance both language perception and production at multiple levels, from syllable processing to turn-taking in natural conversation.

**Keywords:** musical training; hearing loss; cochlear implants; speech development

## 1. Introduction

Before the advent of text messages and email, speech was undoubtedly the most common means of communication. Yet, even nowadays, speech or verbal communication seems to be the most efficient and easier way of conveying thoughts. Speech uses many combinations of phonetic units, vowels, and consonants to convey information. These units can be distinguished because each has a specific spectro-temporal signature. Nonetheless, one may wonder which acoustic features are fundamental for speech perception. Surprising as it may seem, the answer is that there probably are no necessary acoustic features for speech perception. This is because speech perception is an interplay between top-down linguistic predictions and bottom-up sensory signals. Sine-wave speech, for instance, although discarding most acoustic features of natural speech, except the dynamics of vocal resonances, can still be intelligible [1]. If speech perception indeed depended upon the specific spectro-temporal features of consonants and vowels, then a listener hearing sinusoidal signals should not perceive words. This is similar to the fact that *it deosn't mtttaer in waht oredr the ltteers in a wrod are, the olny iprmoetnt tihng is taht the frist and lsat ltteer be at the rghit pclae. The rset can be a toatl mses and you can sitll raed it wouthit porbelm. Tihs is bcuseae the huamn mnid atnicipaets the inoframtion folw.* A predictive coding framework minimizing the error between the sensory input and the predicted signal provides an elegant account of how prior knowledge can radically change what we hear.

This perspective is relevant in the study of the potential benefits of music-making to speech and language abilities. First, because both music and language share the distinctive feature of being dynamically organized in time, they require temporally resolved predictions. Second, in the predictive coding perspective, music and language share the same universal computations, despite surface differences. The similarity might, therefore, exist at the algorithmic level even though specific implementation may differ. Although a view of language and music has dominated the last century as two highly distinct human domains, each with a different neural implementation (modularity of functions), the current view is rather different. To perceive both music and language, one needs to be

able to discriminate sounds. Sounds can be characterized in terms of a limited number of spectral features, and these features are relevant to both musical and linguistic sounds. Both linguistic and musical sounds are categorized. This is important to make sense of the world by reducing its intrinsic variety to a finite and limited number of categories. However, we typically perceive sounds in a complex flow. This requires building a structure that evolves in time, considering the different elements of the temporal sequence. Our previous experience with these sounds will heavily influence the way sounds are perceived in a structure. Such experience generates internal models that allow us to make accurate predictions on upcoming events. These predictions can be made at different temporal scales and affect phoneme categorization, semantic, syntactic, and prosodic processing.

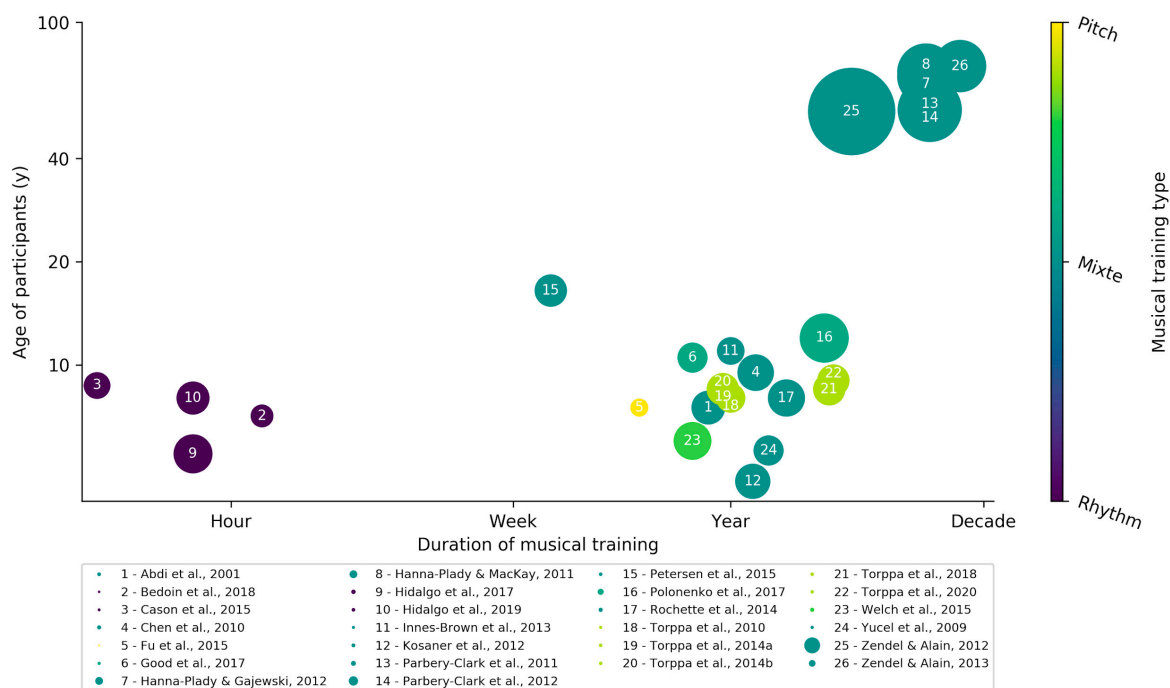
Once ascertained that music and language share several cognitive operations, one can wonder whether music training affects the way the brain processes language, and vice-versa. Indeed, if some of the operations required by music are also required by language, then one should be able to observe more efficient processing in musicians compared to nonmusicians whenever the appropriate language processing levels are investigated. Overall, there has been an increasing number of studies in the last decades, pointing to an improvement induced by music training at different levels of speech and language processing [2]. However, a debate on these effects remains open, especially at the high processing levels. The reason for this success seems to rely on, besides the sharing of sensory or cognitive processing mechanisms, the fact that music places higher demands on these mechanisms compared to speech [3]. This is particularly evident when considering pitch processing, temporal processing, and auditory scene analysis. Music-making requires a high level of both temporal and spectral prediction by adopting a prediction perspective. Indeed, music-making requires a highly precise level of temporal synchronization. It also requires segregating the sound of similar instruments, e.g.; the viola and the violin in a string quartet, which is only possible with an accurate prediction of the spectral content of the music. Importantly, these findings showing a benefit of musical training at different levels of speech and language processing have been underpinned by differences in terms of neural structures and dynamics between musicians and nonmusicians [4]. Altogether, the similarities between music and language in terms of cognitive operations and neural implementation, coupled to music-induced superior language skills, have provided a solid ground for the use of music in speech and language disorders. For reviews see [2,5,6].

When turning to hearing-impaired people, one should acknowledge that, despite early screenings, technology improvements, and intense speech therapy remediation, they still generally suffer from specific deficits in language comprehension and production. Without being exhaustive, children with cochlear implants (CI) and adults with hearing impairments have deficits in discriminating spectral features [7,8], phoneme categorization [9], and perception of speech in noise [8,10–13]. Beyond problems directly related to the implant's limitations in terms of spectral resolution, children with CI also suffer from a lack of auditory stimulation during the first months of life, which possibly leads to deficient structural connectivity [14] and deleterious cortical reorganization [15–17].

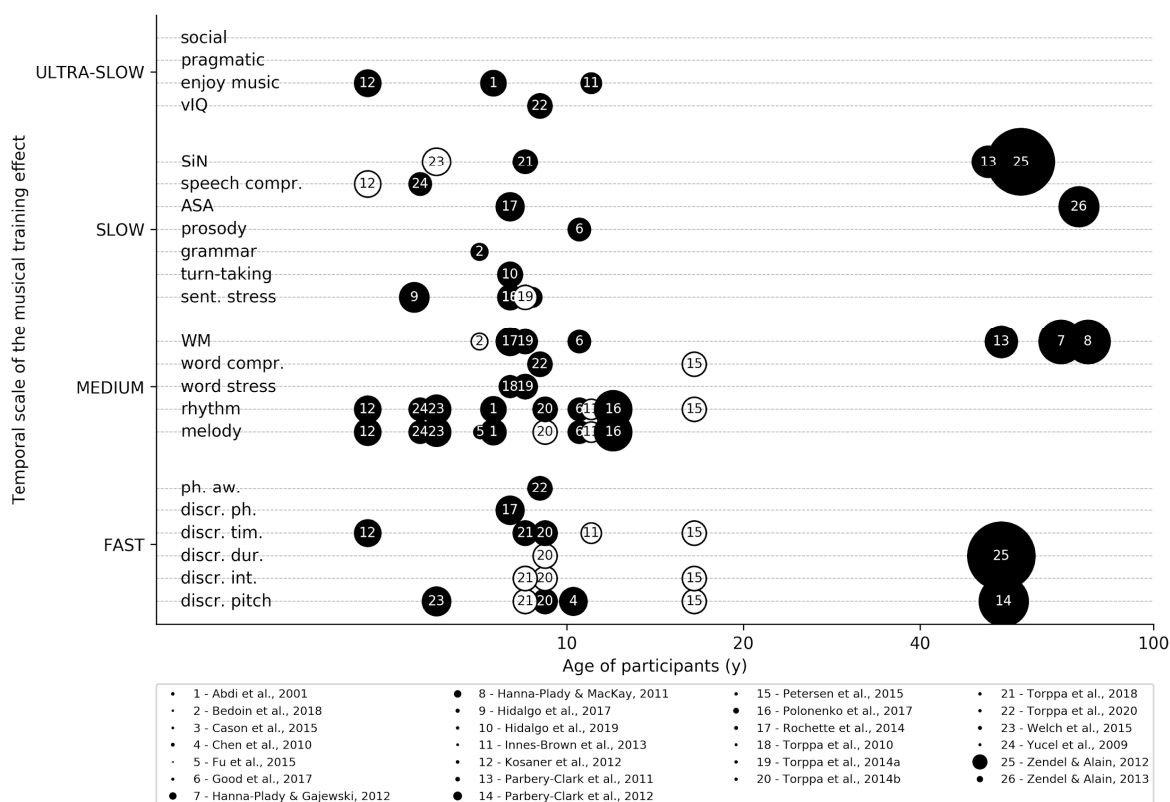
The aim of speech therapy (with children) is to reorient these pathological developmental trajectories. The beginning of a speech therapy usually consists of focusing on low-level features of speech, such as phoneme recognition or spectro-temporal pattern recognition. Later, more integrated features are progressively added, such as syntactic complexity, mental completion, and implicit statements comprehension. However, although benefiting from stimulation of language-related cognitive processes at different levels, hearing-impaired people rarely reach normal-level performance. For reviews, see [18] on the effects of auditory verbal therapy for hearing-impaired children; in [19] on computer-based auditory training for adults with hearing loss. Considering the vast literature on the benefits provided by intense musical training in normal hearing people, a new line of research has emerged to propose new therapies to hearing-impaired people. Although no studies have carefully contrasted standard speech therapy with musical training, the evidence is accumulating in favor of the latter. By providing complex auditory stimuli and tight interactions between perception and action,

musical training is thought to enhance top-down auditory processing and induce brain plasticity at multiple levels, even at shallow ones [20].

Indeed, music training has become increasingly common in hearing impairment rehabilitation, especially in children. Following this growing interest, new scientific literature has flourished and begun to examine the potential benefits that music training provides. We present here a review of most works assessing the potential benefit of musical training on hearing-impaired people (Figure 1). These results are organized concisely according to the level in the hierarchy of the auditory processes they address from spectro-temporal discrimination to the perception of speech in noise (Figure 2). For extensive reviews, we draw the attention of the reader toward two recent reviews on the topic, one for children with CI [21] and one for elderly adults [22]. Although we share the enthusiasm for this line of research. We would like to raise a note of caution on the fact that: (1) most studies are correlational, thus obviating definitive causal statements [23]. (2) Reported effects are generally moderate and (3) rigorous scientific methodology is sometimes hard to achieve in clinical settings with small samples and important inter-individual variability. Randomized controlled trials, with big samples of hearing-impaired participants randomly assigned to experimental and control groups, are still lacking and would be highly valuable to estimate musical training effects accurately. Despite these issues, encouraging results are provided in the literature, and musical training offers significant potential for future auditory remediation therapies.



**Figure 1.** Review of musical training protocols reported in scientific papers [7,24–48]. Each circle represents one study, plotted as a function of the average duration of the musical training and the average age of the participants. Color of the points indicates the content of the training, on a continuum from rhythmic only (drums only) to pitch only (songs, melodies) training. The size of the points indicates the sample size of the study (range: 6–163).



**Figure 2.** Review of musical training effects for hearing impaired people [7,24–48]. Each circle represents one study, plotted as a function of the average age of the participants and the precisely measured effect. Black circles: statistically significant effects; white: non-significant. The size of the points indicates the sample size of the study (range: 6–163).

## 2. Discriminate Sounds Based on Pitch, Duration, and Timbre

At a very rapid temporal scale, speech can be described by the fundamental characteristics of its constituting sounds: pitch, duration, and timbre. These features are critically related to the phonemic level of language, and as such, constitute a privileged target for hearing impairment remediation.

### 2.1. Pitch

Pitch, measured in the number of oscillatory cycles per second (Hz), is an important cue for both music and language. In speech, pitch relates to the notion of the fundamental frequency (F0), i.e.; the frequency of the vocal cords' vibration. F0, as well as sound harmonics F1, F2, etc.; are important cues to identify vowels, tones, sex of the locutor, and information conveyed by prosody. In tone languages, pitch variations also play a role in distinguishing words and grammatical categories. In music, pitch differentiates a high from a low note, and as such is the first constituent of melody. Importantly, small pitch variations, e.g.; 6% for a half-tone, are highly relevant in music, while this is not the case in speech.

It has been shown that musicians have better pitch discrimination abilities than non-musicians in controlled psychophysical studies using pure tones [49–52]. This advantage is correlated to a better ability to discriminate speech tones in adult musicians [53–55], as well as in 8-year-old children practicing music [56]. This advantage extends to vowel discrimination in natural speech, even after controlling for attention [57]. Neural correlates of this difference between musicians and non-musicians have been found in the brainstem response to sounds. Measured with an electroencephalograph (EEG), brainstem responses are acquired via multiple repetitions of a simple short stimulus, such as a syllable. As the brainstem response mimics the acoustic stimulus in both intensity, time, and frequency [58], a measure of acoustic fidelity can be established. It has been shown that musicians have a higher

acoustic fidelity in their brainstem response than nonmusicians [59]. This advantage can be naturally interpreted as a better encoding of the stimulus, leading to a better perception. However, it is also important to keep in mind that it may also reflect better prediction abilities, and stronger top-down connections between the cortex and the brainstem, notably via the cortico-fugal pathway [60].

Concerning concrete enhancement of pitch discrimination abilities in hearing-impaired people, musical training is effective for elderly adults, as well as for children with CI. For elderly adults, one correlational study [37] showed a musician's resilience to age-related delays in the neural timing of syllable subcortical representation. For children with CI, two studies reported differences in F0 discrimination between children having or not having a musical family environment [42]. Children living in a singing environment showed larger amplitude, and shorter latency of a P3a evoked potential, following changes in F0, indicating a change in the brain dynamics. Another study reported an enhancement of F0 discrimination following 2–36 months of musical schooling [61]. Finally, one valuable intervention study reported the same enhancement after two school terms of weekly musical training, although not related to speech-in-noise perception amelioration [45].

## 2.2. Duration

The ability to estimate and discriminate sound duration is crucial for recognizing consonants in speech [62]. It constitutes the main contrast between pairs of voiced consonants. For example, the only difference between a “p” and a “b” is the relative timing between the burst onset and the voice onset, a relation called voice onset time (VOT). In music, duration discrimination is relevant to distinguish different rhythmic patterns, as well as a different interpretation of the same melody (staccato vs. legato).

Multiple studies have shown that musicians have better timing abilities. They achieve lower gap detection thresholds [63] and better temporal intervals discrimination [64,65]. These perceptual benefits are linked to speech-related abilities, insofar as musicians are also better at discriminating syllable duration [66]. Similarly to the pitch dimension, these behavioral measures are linked to neural data, especially to the brainstem response consistency. Indeed, the musician's brainstems encode better rapid transients in speech sounds [67]).

Children with CI suffer from a deficit in duration discrimination abilities, resulting in troubles with phoneme perception [9] and phonological awareness [68,69]. Musical training could, therefore, be a valuable resource for rehabilitation. Indeed, a recent correlational study has investigated the phonemic discrimination abilities of children with CI following 1.5 to 4 years of musical experience [40]). Regression analysis indicated that these children had better scores after training and that musical lessons were partly driving the improvement.

## 2.3. Timbre

Timbre is a more subtle property of sound. It is related to high order properties, such as harmonics and temporal correlations. Two sounds can have the same pitch and duration, but different timbres. For example, the only difference between a note played on a clarinet and the same note played on a piano is their timbre. In both music and speech, timbre is essential for the notion of categorization. Despite differences in timbre, the musician must recognize notes or chords, even though they are played on different instruments. Similarly, a listener must recognize phonemes even if different talkers pronounce them with different voices.

Indeed, musicians have better timbre perception of both musical instruments and voice [70], and reduced susceptibility to the timbral influence of pitch perception [71]. Their brainstem responses to various instrument timbres are also more accurate [72]. Interestingly, this advantage generalizes to language stimulus, such as vowels [73].

Musical training has proven to be effective in children with CI to improve timbre perception. Although not correlated to changes in auditory speech perception, one study has shown improvement in the capacity of recognizing songs, tunes, and timbre in children with CI after an 18 months musical

training program [35]. Another study has revealed that the P3a evoked potential to change in timbre was earlier in children living in a singing environment, suggesting that their timbre perception was enhanced [42].

### 3. Exploit the Temporal Structure and Group Sounds Together

The ability to extract fundamental properties of individual sounds is essential for speech comprehension, especially at the phonemic level, and musical training has proven to be effective at enhancing this ability. As such, it is a promising tool for remediation in hearing-impaired people, both for children with CI and old adults. However, speech and music are not reducible to a sum of their constituting sounds. At a slower temporal scale, speech and music are both defined by a hierarchy of embedded rhythms and by multiple levels of chunks.

#### 3.1. Temporal Structure

Speech and music are both characterized by a hierarchy of rhythms. Short units, such as phonemes and notes, are embedded into longer units, such as words or melodic phrases. The distribution of the frequency of items at each level is not random. Still, it follows a systematic pattern, reproducible across languages: in speech, phonemes happen around 15–40 Hz (15 to 40 per second), syllables around 4–8 Hz, and words around 1–2 Hz [74]. As non-random processes, these rhythms allow the brain to compute expectations and predictions about when the next sound will come. It has been suggested that the synchronization between brain oscillations and speech rhythm, referred to as “entrainment,” is the primary mechanism allowing temporal predictions in speech, and in fine speech comprehension [75,76]. Indeed, the speech multiple frequencies coincide with natural rhythms of brain activity, especially gamma, theta, and delta rhythms [77]. A similar hierarchical structure is involved in music, i.e.; notes are inside rhythms and rhythms inside longer melodic phrases. Similar brain oscillatory mechanisms subtend the extraction of this temporal structure [78]. Furthermore, extracting the temporal structure is also involved in interpersonal synchronization and group playing, where fine-grained temporal synchronization, accurate temporal predictions, and low temporal jitters between players is needed. From the neural standpoint, music and speech both activate the dorsal pathway of the auditory network [79], especially the interaction between the motor and auditory systems [80].

Similar to low-level aspects of sounds, musicians are better at extracting the temporal structure of an auditory stream. They are generally better in rhythm perception [64], rhythm production [27,81], and audiomotor synchronization to both simple [82–84] and complex rhythms [85]. The enhanced ability to track the beat positively correlates with auditory neural synchrony and more precisely with subcortical response consistency [86]. Musicians are also better at performing hierarchical operations, such as grouping sounds in a metrical structure [87,88].

Concerning rehabilitation, there has been a lot of evidence in various clinical populations in favor of rhythmic musical training (for reviews see [5,89]). While music intervention for language disorders has initially been proposed for aphasia [90,91], in particular a music rhythm intervention [92]), in recent years musical interventions have become quite common for children with dyslexia. More precisely, since dyslexia has been characterized as involving a deficit in timing abilities [93], the rhythmic properties of music have been regarded as potentially beneficial to normalize the developmental trajectory [94]. One randomized controlled trial has shown the specific positive effect of 2 years of musical training on phonological awareness in children with a reading disorder [95]. One study showed that 3 years of musical schooling was associated with improvements in phonological awareness and P1/N1 response complex in neurotypical adolescents ([96]; however, see [97]).

For hearing-impaired people, it has been shown that short rhythmic training can enhance the ability to exploit the temporal structure. One study using a priming paradigm showed that rhythmic primes congruent with the metrical structure of target sentences could help children with CI to repeat better phonemes, words, and sentences [26]. A series of two studies have also demonstrated that

30 min of rhythmic training induces a better detection of word rhythmic regularity during verbal exchanges [32] and more precise turn-taking [33] in 5–10 years old children with CI.

### 3.2. Working Memory

Working memory, the ability to temporally maintain and manipulate information [98], is critical for language comprehension [99]. From a general point of view, working memory is important to interpret current items in the light of previous ones, to form chunks and to build expectations about future coming items. Indeed, keeping track of the beginning of a sequence to predict the rest is one fundamental computation for speech comprehension. It has been proven on large meta-analyses that working memory, as both storage and process, is a good predictor (Pearson's  $\rho = 0.30$  to  $0.52$ ) of global speech comprehension [100]. In music, working memory is also strongly involved in both perception and production, e.g.; when keeping track of a chord progression or when memorizing the score online to play on time.

While the ability to anticipate events and learn statistical dependencies between items is enhanced in musicians [101], children with CI seem to have a specific deficit ([102]; however see [103]). Furthermore, both verbal [104] and non-verbal [105] working memory is impaired in children with CI.

Once again, musical training has shown some benefits for working memory in hearing-impaired individuals. Children with CI enrolled in a musical training program for at least one year have a better auditory working memory [7,40]. One intervention study has also shown an enhancement of memory for melodies after six months of piano lessons compared to six months of visual training [29]. In older people, evidence suggests that musicians with at least 10 years of musical training have better verbal [30,36] and non-verbal ([31]; see [106] for a meta-analysis) working memory. However, the reported effects are weak and inconsistent.

## 4. Perceive Speech in Noise, Prosody, and Syntax

### 4.1. Speech in Noise

Speech perception in noise, e.g.; speaking on the phone or in social gatherings, is recognized as having a strong impact on the quality of life and is systematically impaired in hearing-impaired people [8,10–13]. Although several cognitive processes are involved in perceiving speech in noise, such as attention, memory, and listening skills, the ability to analyze auditory scenes stands first. Auditory scene analysis is defined as the capacity to group sound elements coming from one source (one speaker) and segregate elements arising from other sources (other speakers) [22]. In music, auditory scene analysis is both prevalent and challenging in ensemble playing, for instance, when listening to a string quartet wherein the four voices are intermingled strongly.

There is evidence that musical training provides an advantage in the ability to group and segregate auditory sources [47]. This includes pure tones, harmonic complexes, or even spectrally limited “unresolved” harmonic complexes [107]. Musicians are better at perceiving and encoding speech in difficult listening situations. They better discriminate and encode the F0 and F1 of a vowel in the presence of severe reverberation [108]. They encode speech in noise more precisely in the brainstem response [109–111]. This encoding advantage extends to the perception of speech in noise [112]. It has been evaluated to  $\sim 0.7$  dB of signal-to-noise ratio, which translates to  $\sim 5$ – $10\%$  improvement in speech recognition performance [22,113]. However, it should be noted that this point is debated heavily and other results fail to replicate [114–118]. One intervention study has shown that the perception of speech in noise was better in normal-hearing children after two years of musical training compared to 1 year [119]. However, results should be mitigated by the fact that the advantage appeared only after two years, and not after one year.

In hearing-impaired people, although no intervention study has been conducted, several quasi-experimental or cross-sectional studies have provided encouraging results. Children with CI who choose to sing at home have enhanced speech in noise [43] and children with hearing



impairment who receive musical training have better auditory scene analysis [40]. These results are to be interpreted with caution, as sample sizes are small ( $n \approx 15$  per group), confounding factors are not thoroughly considered (e.g.; no randomization), and thus, causal interpretation of these results should be avoided. In older people, studies on life-long musicians have shown speech in noise perception [36,47] and auditory scene analysis enhancement [48].

#### 4.2. Prosody

Prosody refers to the variation of intensity, duration, and spectral features across time in speech. It is an integrated property, as it concerns the dynamics of the entire phrase, for example, the rise of F0 during a question sentence. It conveys emotions, contextual information, and meaning. It is very close to the melodic aspect of music. Indeed, musicians have better discrimination and identification of emotional prosody [120], and a higher and earlier evoked response to changes in prosody [121]. In children with CI, one study showed that six months of piano lessons was associated with enhanced emotional prosody perception, although not significantly different from a control group who did not receive musical training [29].

#### 4.3. Syntax

Finally, at the highest level of analysis, language is composed of embedded syntactic structures [122,123]. Similar to language, music can also be described as having a syntactic structure [124], with its own grammar. For example, in a given context, certain words are grammatically incorrect (\*I have *hungry*). Similarly, in a musical sequence, certain chords or notes are highly unexpected (\* Bm, G, C, A#). For language, these grammatical errors are associated with specific evoked potentials, such as the early left anterior negativity [125] or the P600 [126]. Similar components are evoked by musical “grammatical” errors (ERAN [127]; P600 [128]; for reviews see [129,130]).

Musical training has been effective in restoring evoked potentials associated classically with syntactic processing in various clinical populations. For example, the P600 component is present after three minutes of rhythmic priming in patients with basal ganglia lesions [131]). In contrast, it was proven absent in this type of patient [132]. Similarly, rhythmic primes have been beneficial for grammaticality judgment in normal children, as well as in children with specific language impairment [133,134]. In children with hearing impairment, only one study has been conducted. It suggests a positive, although moderate, effect of rhythmic primes on grammatical judgement, possibly mediated by an effect on speech perception [25]. No effect was revealed for syntactic comprehension. In these studies, children were tested immediately after the rhythmic primes. Further work is needed to establish the presence and the duration of any carryover effect. Indeed, to be useful for therapy, these positive effects must last over time, at least in the order of the day.

### 5. Toward a Remediation of Dialogue by Musical Training

Communication and verbal coordination often go along with a largely automatic process of “interactive alignment” [135]. Two individuals talking with each other simultaneously align their neural dynamics at different linguistic levels by imitating each other’s choices of voice quality, speech rate, prosodic contour, grammatical forms, and meanings [136]. Multi-level alignment improves communication by optimizing turn-taking and coordination behaviors. Importantly, communication is a central feature of social behavior. Furthermore, communicating requires being able to predict others’ actions and integrate these predictions in a joint-action model. Joint-action and prediction can be found in several types of human (and animal) behavior. For all these examples, to succeed, it is not enough to recognize the partners’ actions; rather, it is essential to be able to predict them in advance. Failure in doing so will most often imply the inability to coordinate the timing of each other’s actions [137].

Clearly, music-making is highly demanding in terms of temporal precision and flexibility in interpersonal coordination at multiple timescales and across different sensory modalities. Such coordination is supported by cognitive-motor skills that enable individuals to represent joint

action goals and to anticipate, attend, and adapt to other's actions in real-time. Several studies show that music training can have a facilitatory effect at several levels of speech and language processing [3,6,60]. Meanwhile, synchronized activities lead to increased cooperative and altruistic behaviors in adults [138], children [139], and infants [140]. While little is known of the effect of music training on interpersonal verbal coordination and social skills in individuals with HL, the few existing studies [32,33] show that this may be a promising avenue of research.

Overall, hearing devices are developed for rehabilitating language comprehension and production. *In fine*, the long term goal is to improve communication in hearing-impaired people. Music, especially ensemble playing, offers a particularly stimulating social situation. As such, it is a promising tool for developing social aspects of language, such as turn-taking, language level flexibility, role-playing or even joking. These are aspects that have a tremendous impact on the quality of life of hearing-impaired people.

**Author Contributions:** Writing—original draft preparation, J.P.L.; writing—review and editing, J.P.L.; C.H. and D.S.; supervision, D.S.; project administration, D.S.; funding acquisition, D.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by AGIR POUR L'AUDITION foundation, grant number RD-2016-9, AGENCE NATIONALE DE LA RECHERCHE, grant number ANR-11-LABX-0036 (BLRI) and ANR-16-CONV-0002 (ILCB), and the EXCELLENCE INITIATIVE of Aix-Marseille University (A\*MIDEX).

**Acknowledgments:** We thank E. Truy for his kind invitation, as well as two anonymous reviewers for their constructive criticisms.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Shannon, R.V.; Zeng, F.G.; Kamath, V.; Wygonski, J.; Ekelid, M. Speech recognition with primarily temporal cues. *Science* **1995**, *270*, 303–304. [[CrossRef](#)]
- Schon, D.; Morillon, B. Music and Language. In *The Oxford Handbook of Music and the Brain*; Oxford Press: Oxford, UK, 2018.
- Patel, A.D. Why would Musical Training Benefit the Neural Encoding of Speech? The OPERA Hypothesis. *Front. Psychol.* **2011**, *2*, 142. [[CrossRef](#)]
- Herholz, S.C.; Zatorre, R.J. Musical training as a framework for brain plasticity: Behavior, function, and structure. *Neuron* **2012**, *76*, 486–502. [[CrossRef](#)] [[PubMed](#)]
- Fujii, S.; Wan, C.Y. The role of rhythm in speech and language rehabilitation: The SEP hypothesis. *Front. Hum. Neurosci.* **2014**, *8*, 777. [[CrossRef](#)] [[PubMed](#)]
- Schön, D.; Tillmann, B. Short- and long-term rhythmic interventions: Perspectives for language rehabilitation. *Ann. N. Y. Acad. Sci.* **2015**, *1337*, 32–39. [[CrossRef](#)] [[PubMed](#)]
- Torppa, R.; Faulkner, A.; Huottilainen, M.; Järvikivi, J.; Lipsanen, J.; Laasonen, M.; Vainio, M. The perception of prosody and associated auditory cues in early-implanted children: The role of auditory working memory and musical activities. *Int. J. Audiol.* **2014**, *53*, 182–191. [[CrossRef](#)]
- Moore, B.C.J. Speech processing for the hearing-impaired: Successes, failures, and implications for speech mechanisms. *Speech Commun.* **2003**, *41*, 81–91. [[CrossRef](#)]
- Eisenberg, L.S. Current state of knowledge: Speech recognition and production in children with hearing impairment. *Ear Hear.* **2007**, *28*, 766–772. [[CrossRef](#)]
- Asp, F.; Mäki-Torkko, E.; Karltorp, E.; Harder, H.; Hergils, L.; Eskilsson, G.; Stenfelt, S. Bilateral versus unilateral cochlear implants in children: Speech recognition, sound localization, and parental reports. *Int. J. Audiol.* **2012**, *51*, 817–832. [[CrossRef](#)]
- Caldwell, A.; Nittrouer, S. Speech perception in noise by children with cochlear implants. *J. Speech Lang. Hear. Res.* **2013**, *56*, 13–30. [[CrossRef](#)]
- Crandell, C.C. Individual differences in speech recognition ability: Implications for hearing aid selection. *Ear Hear.* **1991**, *12*, 100S–108S. [[CrossRef](#)] [[PubMed](#)]
- Humes, L.E.; Halling, D.; Coughlin, M. Reliability and Stability of Various Hearing-Aid Outcome Measures in a Group of Elderly Hearing-Aid Wearers. *J. Speech Lang. Hear. Res.* **1996**, *39*, 923–935. [[CrossRef](#)] [[PubMed](#)]

14. Feng, G.; Ingvalson, E.M.; Grieco-Calub, T.M.; Roberts, M.Y.; Ryan, M.E.; Birmingham, P.; Burrowes, D.; Young, N.M.; Wong, P.C.M. Neural preservation underlies speech improvement from auditory deprivation in young cochlear implant recipients. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E1022–E1031. [[CrossRef](#)] [[PubMed](#)]
15. Lee, D.S.; Lee, J.S.; Oh, S.H.; Kim, S.K.; Kim, J.W.; Chung, J.K.; Lee, M.C.; Kim, C.S. Cross-modal plasticity and cochlear implants. *Nature* **2001**, *409*, 149–150. [[CrossRef](#)] [[PubMed](#)]
16. Lee, H.J.; Kang, E.; Oh, S.-H.; Kang, H.; Lee, D.S.; Lee, M.C.; Kim, C.-S. Preoperative differences of cerebral metabolism relate to the outcome of cochlear implants in congenitally deaf children. *Hear. Res.* **2005**, *203*, 2–9. [[CrossRef](#)] [[PubMed](#)]
17. Lee, H.-J.; Giraud, A.-L.; Kang, E.; Oh, S.-H.; Kang, H.; Kim, C.-S.; Lee, D.S. Cortical activity at rest predicts cochlear implantation outcome. *Cereb. Cortex* **2007**, *17*, 909–917. [[CrossRef](#)]
18. Kaipa, R.; Danser, M.L. Efficacy of auditory-verbal therapy in children with hearing impairment: A systematic review from 1993 to 2015. *Int. J. Pediatr. Otorhinolaryngol.* **2016**, *86*, 124–134. [[CrossRef](#)]
19. Henshaw, H.; Ferguson, M.A. Efficacy of individual computer-based auditory training for people with hearing loss: A systematic review of the evidence. *PLoS ONE* **2013**, *8*, e62836. [[CrossRef](#)]
20. Chandrasekaran, B.; Kraus, N. The scalp-recorded brainstem response to speech: Neural origins and plasticity. *Psychophysiology* **2010**, *47*, 236–246. [[CrossRef](#)]
21. Torppa, R.; Huutilainen, M. Why and how music can be used to rehabilitate and develop speech and language skills in hearing-impaired children. *Hear. Res.* **2019**, *380*, 108–122. [[CrossRef](#)]
22. Alain, C.; Zendel, B.R.; Hutka, S.; Bidelman, G.M. Turning down the noise: The benefit of musical training on the aging auditory brain. *Hear. Res.* **2014**, *308*, 162–173. [[CrossRef](#)] [[PubMed](#)]
23. Gfeller, K. Music as Communication and Training for Children with Cochlear Implants. In *Pediatric Cochlear Implantation*; Young, N.M., Iler Kirk, K., Eds.; Springer: New York, NY, USA, 2016; ISBN 978-1-4939-2787-6.
24. Abdi, S.; Khalessi, M.H.; Khorsandi, M.; Gholami, B. Introducing music as a means of habilitation for children with cochlear implants. *Int. J. Pediatr. Otorhinolaryngol.* **2001**, *59*, 105–113. [[CrossRef](#)]
25. Bedoin, N.; Besombes, A.M.; Escande, E.; Dumont, A.; Lalitte, P.; Tillmann, B. Boosting syntax training with temporally regular musical primes in children with cochlear implants. *Ann. Phys. Rehabil. Med.* **2018**, *61*, 365–371. [[CrossRef](#)] [[PubMed](#)]
26. Cason, N.; Hidalgo, C.; Isoard, F.; Roman, S.; Schön, D. Rhythmic priming enhances speech production abilities: Evidence from prelingually deaf children. *Neuropsychology* **2015**, *29*, 102–107. [[CrossRef](#)]
27. Chen, J.L.; Penhune, V.B.; Zatorre, R.J. Moving on time: Brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training. *J. Cogn. Neurosci.* **2008**, *20*, 226–239. [[CrossRef](#)]
28. Fu, Q.-J.; Galvin, J.J.; Wang, X.; Wu, J.-L. Benefits of music training in mandarin-speaking pediatric cochlear implant users. *J. Speech Lang. Hear. Res.* **2015**, *58*, 163–169. [[CrossRef](#)]
29. Good, A.; Gordon, K.A.; Papsin, B.C.; Nespoli, G.; Hopyan, T.; Peretz, I.; Russo, F.A. Benefits of music training for perception of emotional speech prosody in deaf children with cochlear implants. *Ear Hear.* **2017**, *38*, 455–464. [[CrossRef](#)]
30. Hanna-Pladdy, B.; Gajewski, B. Recent and past musical activity predicts cognitive aging variability: Direct comparison with general lifestyle activities. *Front. Hum. Neurosci.* **2012**, *6*, 198. [[CrossRef](#)]
31. Hanna-Pladdy, B.; MacKay, A. The relation between instrumental musical activity and cognitive aging. *Neuropsychology* **2011**, *25*, 378–386. [[CrossRef](#)]
32. Hidalgo, C.; Falk, S.; Schön, D. Speak on time! Effects of a musical rhythmic training on children with hearing loss. *Hear. Res.* **2017**, *351*, 11–18. [[CrossRef](#)]
33. Hidalgo, C.; Pesnot-Lerousseau, J.; Marquis, P.; Roman, S.; Schön, D. Rhythmic training improves temporal anticipation and adaptation abilities in children with hearing loss during verbal interaction. *J. Speech Lang. Hear. Res.* **2019**, *62*, 3234–3247. [[CrossRef](#)] [[PubMed](#)]
34. Innes-Brown, H.; Marozeau, J.P.; Storey, C.M.; Blamey, P.J. Tone, rhythm, and timbre perception in school-age children using cochlear implants and hearing aids. *J. Am. Acad. Audiol.* **2013**, *24*, 789–806. [[CrossRef](#)] [[PubMed](#)]
35. Koşaner, J.; Kilinc, A.; Deniz, M. Developing a music programme for preschool children with cochlear implants. *Cochlear Implants Int.* **2012**, *13*, 237–247. [[CrossRef](#)] [[PubMed](#)]
36. Parbery-Clark, A.; Strait, D.L.; Anderson, S.; Hittner, E.; Kraus, N. Musical experience and the aging auditory system: Implications for cognitive abilities and hearing speech in noise. *PLoS ONE* **2011**, *6*, e18082. [[CrossRef](#)]

37. Parbery-Clark, A.; Anderson, S.; Hittner, E.; Kraus, N. Musical experience offsets age-related delays in neural timing. *Neurobiol. Aging* **2012**, *33*, 1483.e1-4. [[CrossRef](#)]
38. Petersen, B.; Weed, E.; Sandmann, P.; Brattico, E.; Hansen, M.; Sørensen, S.D.; Vuust, P. Brain responses to musical feature changes in adolescent cochlear implant users. *Front. Hum. Neurosci.* **2015**, *9*, 7. [[CrossRef](#)]
39. Polonenko, M.J.; Giannantonio, S.; Papsin, B.C.; Marsella, P.; Gordon, K.A. Music perception improves in children with bilateral cochlear implants or bimodal devices. *J. Acoust. Soc. Am.* **2017**, *141*, 4494. [[CrossRef](#)] [[PubMed](#)]
40. Rochette, F.; Moussard, A.; Bigand, E. Music lessons improve auditory perceptual and cognitive performance in deaf children. *Front. Hum. Neurosci.* **2014**, *8*, 488. [[CrossRef](#)]
41. Torppa, R.; Faulkner, A.; Järvikivi, J. Acquisition of focus by normal hearing and cochlear implanted children: The role of musical experience. *Speech Prosody* **2010**.
42. Torppa, R.; Huutilainen, M.; Leminen, M.; Lipsanen, J.; Tervaniemi, M. Interplay between singing and cortical processing of music: A longitudinal study in children with cochlear implants. *Front. Psychol.* **2014**, *5*, 1389. [[CrossRef](#)]
43. Torppa, R.; Faulkner, A.; Kujala, T.; Huutilainen, M.; Lipsanen, J. Developmental Links Between Speech Perception in Noise, Singing, and Cortical Processing of Music in Children with Cochlear Implants. *Music Percept.* **2018**, *36*, 156–174. [[CrossRef](#)]
44. Torppa, R.; Faulkner, A.; Laasonen, M.; Lipsanen, J.; Sammler, D. Links of prosodic stress perception and musical activities to language skills of children with cochlear implants and normal hearing. *Ear Hear.* **2020**, *41*, 395–410. [[CrossRef](#)] [[PubMed](#)]
45. Welch, G.F.; Saunders, J.; Edwards, S.; Palmer, Z.; Himonides, E.; Knight, J.; Mahon, M.; Griffin, S.; Vickers, D.A. Using singing to nurture children's hearing? A pilot study. *Cochlear Implants Int.* **2015**, *16* (Suppl. 3), S63–S70. [[CrossRef](#)]
46. Yucel, E.; Sennaroglu, G.; Belgin, E. The family oriented musical training for children with cochlear implants: Speech and musical perception results of two year follow-up. *Int. J. Pediatr. Otorhinolaryngol.* **2009**, *73*, 1043–1052. [[CrossRef](#)] [[PubMed](#)]
47. Zendel, B.R.; Alain, C. Musicians experience less age-related decline in central auditory processing. *Psychol. Aging* **2012**, *27*, 410–417. [[CrossRef](#)] [[PubMed](#)]
48. Zendel, B.R.; Alain, C. The influence of lifelong musicianship on neurophysiological measures of concurrent sound segregation. *J. Cogn. Neurosci.* **2013**, *25*, 503–516. [[CrossRef](#)]
49. Bidelman, G.M.; Gandour, J.T.; Krishnan, A. Musicians and tone-language speakers share enhanced brainstem encoding but not perceptual benefits for musical pitch. *Brain Cogn.* **2011**, *77*, 1–10. [[CrossRef](#)]
50. Kishon-Rabin, L.; Amir, O.; Vexler, Y.; Zaltz, Y. Pitch discrimination: Are professional musicians better than non-musicians? *J. Basic Clin. Physiol. Pharmacol.* **2001**, *12*, 125–143. [[CrossRef](#)] [[PubMed](#)]
51. Micheyl, C.; Delhommeau, K.; Perrot, X.; Oxenham, A.J. Influence of musical and psychoacoustical training on pitch discrimination. *Hear. Res.* **2006**, *219*, 36–47. [[CrossRef](#)]
52. Spiegel, M.F.; Watson, C.S. Performance on frequency-discrimination tasks by musicians and nonmusicians. *J. Acoust. Soc. Am.* **1984**, *76*, 1690–1695. [[CrossRef](#)]
53. Gottfried, T.L.; Staby, A.M.; Ziemer, C.J. Musical experience and Mandarin tone discrimination and imitation. *J. Acoust. Soc. Am.* **2004**, *115*, 2545. [[CrossRef](#)]
54. Gottfried, T.L.; Riester, D. Relation of pitch glide perception and Mandarin tone identification. *J. Acoust. Soc. Am.* **2000**, *108*, 2604. [[CrossRef](#)]
55. Marie, C.; Delogu, F.; Lampis, G.; Belardinelli, M.O.; Besson, M. Influence of musical expertise on segmental and tonal processing in Mandarin Chinese. *J. Cogn. Neurosci.* **2011**, *23*, 2701–2715. [[CrossRef](#)] [[PubMed](#)]
56. Magne, C.; Schön, D.; Besson, M. Musician children detect pitch violations in both music and language better than nonmusician children: Behavioral and electrophysiological approaches. *J. Cogn. Neurosci.* **2006**, *18*, 199–211. [[CrossRef](#)] [[PubMed](#)]
57. Sares, A.G.; Foster, N.E.V.; Allen, K.; Hyde, K.L. Pitch and time processing in speech and tones: The effects of musical training and attention. *J. Speech Lang. Hear. Res.* **2018**, *61*, 496–509. [[CrossRef](#)]
58. Krizman, J.; Kraus, N. Analyzing the FFR: A tutorial for decoding the richness of auditory function. *Hear. Res.* **2019**, *382*, 107779. [[CrossRef](#)]
59. Musacchia, G.; Sams, M.; Skoe, E.; Kraus, N. Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 15894–15898. [[CrossRef](#)]

60. Kraus, N.; Chandrasekaran, B. Music training for the development of auditory skills. *Nat. Rev. Neurosci.* **2010**, *11*, 599–605. [[CrossRef](#)]
61. Chen, J.K.-C.; Chuang, A.Y.C.; McMahon, C.; Hsieh, J.-C.; Tung, T.-H.; Li, L.P.-H. Music training improves pitch perception in prelingually deafened children with cochlear implants. *Pediatrics* **2010**, *125*, e793–e800. [[CrossRef](#)]
62. Rosen, S. Temporal information in speech: Acoustic, auditory and linguistic aspects. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **1992**, *336*, 367–373. [[CrossRef](#)]
63. Kuman, P.V.; Rana, B.; Krishna, R. Temporal processing in musicians and non-musicians. *J. Hear Sci.* **2014**.
64. Rammsayer, T.; Altenmüller, E. Temporal information processing in musicians and nonmusicians. *Music Percept.* **2006**, *24*, 37–48. [[CrossRef](#)]
65. Wang, X.; Ossher, L.; Reuter-Lorenz, P.A. Examining the relationship between skilled music training and attention. *Conscious. Cogn.* **2015**, *36*, 169–179. [[CrossRef](#)] [[PubMed](#)]
66. Marie, C.; Magne, C.; Besson, M. Musicians and the metric structure of words. *J. Cogn. Neurosci.* **2011**, *23*, 294–305. [[CrossRef](#)] [[PubMed](#)]
67. Parbery-Clark, A.; Tierney, A.; Strait, D.L.; Kraus, N. Musicians have fine-tuned neural distinction of speech syllables. *Neuroscience* **2012**, *219*, 111–119. [[CrossRef](#)]
68. Ambrose, S.E.; Fey, M.E.; Eisenberg, L.S. Phonological awareness and print knowledge of preschool children with cochlear implants. *J. Speech Lang. Hear. Res.* **2012**, *55*, 811–823. [[CrossRef](#)]
69. Soleymani, Z.; Mahmoodabadi, N.; Nouri, M.M. Language skills and phonological awareness in children with cochlear implants and normal hearing. *Int. J. Pediatr. Otorhinolaryngol.* **2016**, *83*, 16–21. [[CrossRef](#)] [[PubMed](#)]
70. Chartrand, J.-P.; Belin, P. Superior voice timbre processing in musicians. *Neurosci. Lett.* **2006**, *405*, 164–167. [[CrossRef](#)] [[PubMed](#)]
71. Pitt, M.A. Perception of pitch and timbre by musically trained and untrained listeners. *J. Exp. Psychol. Hum. Percept. Perform.* **1994**, *20*, 976–986. [[CrossRef](#)]
72. Musacchia, G.; Strait, D.; Kraus, N. Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians. *Hear. Res.* **2008**, *241*, 34–42. [[CrossRef](#)]
73. Intartaglia, B.; White-Schwoch, T.; Kraus, N.; Schön, D. Music training enhances the automatic neural processing of foreign speech sounds. *Sci. Rep.* **2017**, *7*, 12631. [[CrossRef](#)] [[PubMed](#)]
74. Ding, N.; Patel, A.D.; Chen, L.; Butler, H.; Luo, C.; Poeppel, D. Temporal modulations in speech and music. *Neurosci. Biobehav. Rev.* **2017**, *81*, 181–187. [[CrossRef](#)] [[PubMed](#)]
75. Giraud, A.-L.; Poeppel, D. Cortical oscillations and speech processing: Emerging computational principles and operations. *Nat. Neurosci.* **2012**, *15*, 511–517. [[CrossRef](#)] [[PubMed](#)]
76. Ding, N.; Melloni, L.; Zhang, H.; Tian, X.; Poeppel, D. Cortical tracking of hierarchical linguistic structures in connected speech. *Nat. Neurosci.* **2016**, *19*, 158–164. [[CrossRef](#)] [[PubMed](#)]
77. Buzsáki, G.; Draguhn, A. Neuronal oscillations in cortical networks. *Science* **2004**, *304*, 1926–1929. [[CrossRef](#)] [[PubMed](#)]
78. Nozaradan, S.; Peretz, I.; Missal, M.; Mouraux, A. Tagging the neuronal entrainment to beat and meter. *J. Neurosci.* **2011**, *31*, 10234–10240. [[CrossRef](#)]
79. Rauschecker, J.P.; Scott, S.K. Maps and streams in the auditory cortex: Nonhuman primates illuminate human speech processing. *Nat. Neurosci.* **2009**, *12*, 718–724. [[CrossRef](#)]
80. Zatorre, R.J.; Chen, J.L.; Penhune, V.B. When the brain plays music: Auditory-motor interactions in music perception and production. *Nat. Rev. Neurosci.* **2007**, *8*, 547–558. [[CrossRef](#)]
81. Drake, C. Reproduction of musical rhythms by children, adult musicians, and adult nonmusicians. *Percept. Psychophys.* **1993**, *53*, 25–33. [[CrossRef](#)]
82. Krause, V.; Pollok, B.; Schnitzler, A. Perception in action: The impact of sensory information on sensorimotor synchronization in musicians and non-musicians. *Acta Psychol. (Amst.)* **2010**, *133*, 28–37. [[CrossRef](#)]
83. Repp, B.H. Sensorimotor synchronization and perception of timing: Effects of music training and task experience. *Hum. Mov. Sci.* **2010**, *29*, 200–213. [[CrossRef](#)] [[PubMed](#)]
84. Repp, B.H.; Doggett, R. Tapping to a very slow beat: A comparison of musicians and nonmusicians. *Music Percept. Interdiscip. J.* **2007**, *24*, 367–376. [[CrossRef](#)]
85. Farrugia, N.; Benoit, C.; Harding, E. BAASTA: Battery for the Assessment of Auditory Sensorimotor and Timing Abilities. *Behav. Res. Methods* **2017**, *49*, 1128–1145. [[CrossRef](#)]

86. Tierney, A.; Kraus, N. Neural responses to sounds presented on and off the beat of ecologically valid music. *Front. Syst. Neurosci.* **2013**, *7*, 14. [[CrossRef](#)]
87. Bouwer, F.L.; Burgoyne, J.A.; Odijk, D.; Honing, H.; Grahn, J.A. What makes a rhythm complex? The influence of musical training and accent type on beat perception. *PLoS ONE* **2018**, *13*, e0190322. [[CrossRef](#)]
88. Kung, S.-J.; Tzeng, O.J.L.; Hung, D.L.; Wu, D.H. Dynamic allocation of attention to metrical and grouping accents in rhythmic sequences. *Exp. Brain Res.* **2011**, *210*, 269–282. [[CrossRef](#)] [[PubMed](#)]
89. Tierney, A.; Kraus, N. Music training for the development of reading skills. *Prog. Brain Res.* **2013**, *207*, 209–241. [[CrossRef](#)] [[PubMed](#)]
90. Albert, M.L.; Sparks, R.W.; Helm, N.A. Melodic intonation therapy for aphasia. *Arch. Neurol.* **1973**, *29*, 130–131. [[CrossRef](#)]
91. Sparks, R.; Helm, N.; Albert, M. Aphasia rehabilitation resulting from melodic intonation therapy. *Cortex* **1974**, *10*, 303–316. [[CrossRef](#)]
92. Stahl, B.; Kotz, S.A.; Henseler, I.; Turner, R.; Geyer, S. Rhythm in disguise: Why singing may not hold the key to recovery from aphasia. *Brain* **2011**, *134*, 3083–3093. [[CrossRef](#)]
93. Goswami, U.; Thomson, J.; Richardson, U.; Stainthorpe, R.; Hughes, D.; Rosen, S.; Scott, S.K. Amplitude envelope onsets and developmental dyslexia: A new hypothesis. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 10911–10916. [[CrossRef](#)]
94. Overy, K. Dyslexia and music. From timing deficits to musical intervention. *Ann. N. Y. Acad. Sci.* **2003**, *999*, 497–505. [[CrossRef](#)] [[PubMed](#)]
95. Flaugnacco, E.; Lopez, L.; Terribili, C.; Montico, M.; Zoia, S.; Schön, D. Music training increases phonological awareness and reading skills in developmental dyslexia: A randomized control trial. *PLoS ONE* **2015**, *10*, e0138715. [[CrossRef](#)] [[PubMed](#)]
96. Tierney, A.T.; Krizman, J.; Kraus, N. Music training alters the course of adolescent auditory development. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 10062–10067. [[CrossRef](#)]
97. Moreno, S.; Friesen, D.; Bialystok, E. Effect of music training on promoting preliterate skills: Preliminary causal evidence. *Music Percept. Interdisci. J.* **2011**, *29*, 165–172. [[CrossRef](#)]
98. Baddeley, A. Working memory. *Science* **1992**, *255*, 556–559. [[CrossRef](#)]
99. Baddeley, A. Working memory and language: An overview. *J. Commun. Disord.* **2003**, *36*, 189–208. [[CrossRef](#)]
100. Daneman, M.; Merikle, P.M. Working memory and language comprehension: A meta-analysis. *Psychon. Bull. Rev.* **1996**, *3*, 422–433. [[CrossRef](#)]
101. Francois, C.; Schön, D. Musical expertise boosts implicit learning of both musical and linguistic structures. *Cereb. Cortex* **2011**, *21*, 2357–2365. [[CrossRef](#)]
102. Conway, C.M.; Pisoni, D.B.; Anaya, E.M.; Karpicke, J.; Henning, S.C. Implicit sequence learning in deaf children with cochlear implants. *Dev. Sci.* **2011**, *14*, 69–82. [[CrossRef](#)]
103. Torkildsen, J. Von K.; Arciuli, J.; Haukedal, C.L.; Wie, O.B. Does a lack of auditory experience affect sequential learning? *Cognition* **2018**, *170*, 123–129. [[CrossRef](#)]
104. Nitttrouer, S.; Caldwell-Tarr, A.; Lowenstein, J.H. Working memory in children with cochlear implants: Problems are in storage, not processing. *Int. J. Pediatr. Otorhinolaryngol.* **2013**, *77*, 1886–1898. [[CrossRef](#)]
105. AuBuchon, A.M.; Pisoni, D.B.; Kronenberger, W.G. Short-Term and Working Memory Impairments in Early-Implanted, Long-Term Cochlear Implant Users Are Independent of Audibility and Speech Production. *Ear Hear.* **2015**, *36*, 733–737. [[CrossRef](#)]
106. Talamini, F.; Altoè, G.; Carretti, B.; Grassi, M. Musicians have better memory than nonmusicians: A meta-analysis. *PLoS ONE* **2017**, *12*, e0186773. [[CrossRef](#)] [[PubMed](#)]
107. Vliegen, J.; Oxenham, A.J. Sequential stream segregation in the absence of spectral cues. *J. Acoust. Soc. Am.* **1999**, *105*, 339–346. [[CrossRef](#)]
108. Bidelman, G.M.; Krishnan, A. Effects of reverberation on brainstem representation of speech in musicians and non-musicians. *Brain Res.* **2010**, *1355*, 112–125. [[CrossRef](#)] [[PubMed](#)]
109. Parbery-Clark, A.; Skoe, E.; Lam, C.; Kraus, N. Musician enhancement for speech-in-noise. *Ear Hear.* **2009**, *30*, 653–661. [[CrossRef](#)] [[PubMed](#)]
110. Slater, J.; Kraus, N.; Woodruff Carr, K.; Tierney, A.; Azem, A.; Ashley, R. Speech-in-noise perception is linked to rhythm production skills in adult percussionists and non-musicians. *Lang. Cogn. Neurosci.* **2017**, *33*, 1–8. [[CrossRef](#)] [[PubMed](#)]

111. Strait, D.L.; Parbery-Clark, A.; Hittner, E.; Kraus, N. Musical training during early childhood enhances the neural encoding of speech in noise. *Brain Lang.* **2012**, *123*, 191–201. [[CrossRef](#)] [[PubMed](#)]
112. Coffey, E.B.J.; Mogilever, N.B.; Zatorre, R.J. Speech-in-noise perception in musicians: A review. *Hear. Res.* **2017**, *352*, 49–69. [[CrossRef](#)]
113. Middelweerd, M.J.; Festen, J.M.; Plomp, R. Difficulties with Speech Intelligibility in Noise in Spite of a Normal Pure-Tone Audiogram. *Int. J. Audiol.* **1990**, *29*, 1–7. [[CrossRef](#)]
114. Boebinger, D.; Evans, S.; Rosen, S.; Lima, C.F.; Manly, T.; Scott, S.K. Musicians and non-musicians are equally adept at perceiving masked speech. *J. Acoust. Soc. Am.* **2015**, *137*, 378–387. [[CrossRef](#)] [[PubMed](#)]
115. Escobar, J.; Mussoi, B.S.; Silberer, A.B. The effect of musical training and working memory in adverse listening situations. *Ear Hear.* **2020**, *41*, 278–288. [[CrossRef](#)] [[PubMed](#)]
116. Madsen, S.M.K.; Marschall, M.; Dau, T.; Oxenham, A.J. Speech perception is similar for musicians and non-musicians across a wide range of conditions. *Sci. Rep.* **2019**, *9*, 10404. [[CrossRef](#)] [[PubMed](#)]
117. Madsen, S.M.K.; Whiteford, K.L.; Oxenham, A.J. Musicians do not benefit from differences in fundamental frequency when listening to speech in competing speech backgrounds. *Sci. Rep.* **2017**, *7*, 12624. [[CrossRef](#)] [[PubMed](#)]
118. Ruggles, D.R.; Freyman, R.L.; Oxenham, A.J. Influence of musical training on understanding voiced and whispered speech in noise. *PLoS ONE* **2014**, *9*, e86980. [[CrossRef](#)] [[PubMed](#)]
119. Slater, J.; Skoe, E.; Strait, D.L.; O’Connell, S.; Thompson, E.; Kraus, N. Music training improves speech-in-noise perception: Longitudinal evidence from a community-based music program. *Behav. Brain Res.* **2015**, *291*, 244–252. [[CrossRef](#)]
120. Thompson, W.F.; Schellenberg, E.G.; Husain, G. Decoding speech prosody: Do music lessons help? *Emotion* **2004**, *4*, 46–64. [[CrossRef](#)]
121. Schön, D.; Magne, C.; Besson, M. The music of speech: Music training facilitates pitch processing in both music and language. *Psychophysiology* **2004**, *41*, 341–349. [[CrossRef](#)]
122. Chomsky, N. Logical structures in language. *Am. Doc.* **1957**, *8*, 284–291. [[CrossRef](#)]
123. Hauser, M.D.; Chomsky, N.; Fitch, W.T. The faculty of language: What is it, who has it, and how did it evolve? *Science* **2002**, *298*, 1569–1579. [[CrossRef](#)]
124. Lerdahl, F.; Jackendoff, R.S. *A Generative Theory of Tonal Music*; MIT Press: Cambridge, MA, USA, 1996.
125. Friederici, A.D. Towards a neural basis of auditory sentence processing. *Trends Cogn. Sci. (Regul. Ed.)* **2002**, *6*, 78–84. [[CrossRef](#)]
126. Frisch, S.; Schlesewsky, M.; Saddy, D.; Alpermann, A. The P600 as an indicator of syntactic ambiguity. *Cognition* **2002**, *85*, B83–B92. [[CrossRef](#)]
127. Maess, B.; Koelsch, S.; Gunter, T.C.; Friederici, A.D. Musical syntax is processed in Broca’s area: An MEG study. *Nat. Neurosci.* **2001**, *4*, 540–545. [[CrossRef](#)] [[PubMed](#)]
128. Patel, A.D.; Gibson, E.; Ratner, J.; Besson, M.; Holcomb, P.J. Processing syntactic relations in language and music: An event-related potential study. *J. Cogn. Neurosci.* **1998**, *10*, 717–733. [[CrossRef](#)]
129. Koelsch, S. Neural substrates of processing syntax and semantics in music. In *Music that Works*; Haas, R., Brandes, V., Eds.; Springer: Vienna, Austria, 2009; pp. 143–153. ISBN 978-3-211-75120-6.
130. Patel, A.D. Language, music, syntax and the brain. *Nat. Neurosci.* **2003**, *6*, 674–681. [[CrossRef](#)] [[PubMed](#)]
131. Kotz, S.A.; Gunter, T.C.; Wonneberger, S. The basal ganglia are receptive to rhythmic compensation during auditory syntactic processing: ERP patient data. *Brain Lang.* **2005**, *95*, 70–71. [[CrossRef](#)]
132. Kotz, S.A.; Frisch, S.; von Cramon, D.Y.; Friederici, A.D. Syntactic language processing: ERP lesion data on the role of the basal ganglia. *J. Int. Neuropsychol. Soc.* **2003**, *9*, 1053–1060. [[CrossRef](#)]
133. Bedoin, N.; Bricseau, L.; Molinier, P.; Roch, D.; Tillmann, B. Temporally Regular Musical Primes Facilitate Subsequent Syntax Processing in Children with Specific Language Impairment. *Front. Neurosci.* **2016**, *10*, 245. [[CrossRef](#)]
134. Przybylski, L.; Bedoin, N.; Krifi-Papoz, S.; Herbillon, V.; Roch, D.; Léculier, L.; Kotz, S.A.; Tillmann, B. Rhythmic auditory stimulation influences syntactic processing in children with developmental language disorders. *Neuropsychology* **2013**, *27*, 121–131. [[CrossRef](#)]
135. Pickering, M.J.; Garrod, S. Toward a mechanistic psychology of dialogue. *Behav. Brain Sci.* **2004**, *27*, 169–190, discussion 190. [[CrossRef](#)]
136. Garrod, S.; Pickering, M.J. Joint action, interactive alignment, and dialog. *Top. Cogn. Sci.* **2009**, *1*, 292–304. [[CrossRef](#)] [[PubMed](#)]

137. Tuller, B.; Lancia, L. Speech dynamics: Converging evidence from syllabification and categorization. *J. Phon.* **2017**, *64*, 21–33. [[CrossRef](#)]
138. Wiltermuth, S.S.; Heath, C. Synchrony and cooperation. *Psychol. Sci.* **2009**, *20*, 1–5. [[CrossRef](#)] [[PubMed](#)]
139. Kirschner, S.; Tomasello, M. Joint music making promotes prosocial behavior in 4-year-old children. *Evol. Hum. Behav.* **2010**, *31*, 354–364. [[CrossRef](#)]
140. Cirelli, L.K.; Einarson, K.M.; Trainor, L.J. Interpersonal synchrony increases prosocial behavior in infants. *Dev. Sci.* **2014**, *17*, 1003–1011. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).