

Experience-induced Malleability in Neural Encoding of *Pitch*, *Timbre*, and *Timing*

Implications for Language and Music

[Nina Kraus](#),^{a,b} [Erika Skoe](#),^a [Alexandra Parbery-Clark](#),^a and [Richard Ashley](#),^c

[Author information](#) [Copyright and License information](#) [Disclaimer](#)

The publisher's final edited version of this article is available at [Ann N Y Acad Sci](#)

[Go to:](#)

Abstract

Speech and music are highly complex signals that have many shared acoustic features. *Pitch*, *Timbre*, and *Timing* can be used as overarching perceptual categories for describing these shared properties. The acoustic cues contributing to these percepts also have distinct subcortical representations which can be selectively enhanced or degraded in different populations. Musically trained subjects are found to have enhanced subcortical representations of *pitch*, *timbre*, and *timing*. The effects of musical experience on subcortical auditory processing are pervasive and extend beyond music to the domains of

language and emotion. The sensory malleability of the neural encoding of *pitch*, *timbre*, and *timing* can be affected by lifelong experience and short-term training. This conceptual framework and supporting data can be applied to consider sensory learning of speech and music through a hearing aid or cochlear implant.

Keywords: brain stem, subcortical, musical training, cochlear implant

[Go to:](#)

Introduction

From the cochlea to the auditory cortex, sound is encoded at multiple locations along the ascending auditory pathway, eventually leading to conscious perception. While there is no doubt that the cortex plays a major role in the perception of speech, music, and other meaningful auditory signals, recent studies suggest that subcortical encoding of sound is not merely a series of passive, bottom-up processes successively transforming the acoustic signal into a more complex neural code. Rather, subcortical sensory processes dynamically interact with cortical processes, such as memory, attention, and multisensory integration, to

shape the perceptual system's response to speech and music.

In the last two decades there has been a surge in research devoted to how musical experience affects brain structure, cortical activity, and auditory perception. These three lines of research have uncovered several interesting byproducts of musical training. Musicians have brain structural differences not only in the motor cortices—the parts of the brain controlling hand/finger movement and coordination—but also in the auditory cortices.[1,2](#) In addition to structural differences, musicians show different patterns of neural activation. For example, musicians show stronger responses to simple, artificial tones and heightened responses to the sound of their own instrument compared to other instruments.[3–7](#) Interestingly, such cortical differences can be seen as early as 1 year after the onset of musical training[8](#) and extend to speech signals.[9,10](#) Recently, this line of research has moved to subcortical levels. This work, along with supporting data, will be presented here within the *pitch*, *timbre*, and *timing* conceptual

framework. In the final section of this review, we will switch the focus to cochlear implants and apply this conceptual framework to consider sensory learning of speech and music through an implant.

[Go to:](#)

Conceptual Framework for Studying Subcortical Responses: *Pitch*, *Timbre*, and *Timing*

Work from our laboratory^{[d](#)} points to *pitch*, *timbre*, and *timing* as having distinct subcortical representations which can be selectively enhanced or degraded in different populations.

Pitch, as defined by the Standard Acoustical Terminology of the Acoustical Society of America, is “that attribute of auditory sensation in terms of which sounds may be ordered on a scale extending from low to high” S12.01, P.34.^{[11](#)} For pure tones, the frequency, or cycles per second of the waveform, is the physical correlate of *pitch*; however when considering more complex sounds, *pitch* corresponds, in part, to the lowest resonant frequency, also known as the fundamental frequency (F₀).^{[e](#)} For speech, F₀ is dictated

by the rate of vocal fold vibration and for music it depends on the instrument. For example, the reed is the source of F_0 vibration for the oboe and clarinet, whereas the string is the source for the violin and guitar. For the purposes of this review, we use the word *pitch* as shorthand for referring to the information carried by the F_0 , and so in this context, *pitch* and F_0 are synonymous.

Timbre, also referred to as “sound color,” enables us to differentiate two sounds with the same *pitch*. *Timbre* is a multidimensional property resulting from the interaction of spectral and temporal changes associated with the harmonics of the fundamental along with the *timing* cues of the attack (onset) and decay (offset). Together this gives rise to the characteristic sound quality associated with a given instrument or voice. *Timbre* is also an important cue for distinguishing contrastive speech sounds (i.e., phonemes). As the vocal tract is shaped by the movement of the articulators during speech production, the resonance structure of the vocal tract changes and certain harmonics are attenuated while others are amplified. These amplified harmonics are known as

speech-formants and they are important for distinguishing phonemes. Our focus here is on the harmonic aspects of *timbre* and the corresponding subcortical representation.

Timing refers to the major acoustic landmarks in the temporal envelope of speech and music signals. For speech, *timing* arises from the alternating opening and closing of the articulators and from the interplay between laryngeal and supralaryngeal gestures. *Timing* also includes spectrotemporal features of speech, such as time-varying formants. As such, *timing* arises from the interplay between the actions of the source (glottal pulse train) and filter (articulators). For music, *timing* can be considered in conjunction with the temporal information contributing to *timbre* perception. Likewise, on a more global scale, it refers to the duration of sounds and their subsequent perceptual groupings into rhythm. For the purposes of this review, we will focus on the neural representation of transient temporal features, such as onsets and offsets occurring as fast as fractions of milliseconds.

The Auditory brain stem Response

The auditory brain stem, an ensemble of nuclei belonging to the efferent and afferent auditory systems, receives and processes the output of the cochlea en route to higher centers of auditory processing. The auditory brain stem response (ABR), a highly replicable far-field potential recorded from surface electrodes placed on the scalp, reflects the acoustic properties of the sound stimulus with remarkable fidelity. In fact, when the electrical response is converted into an audio signal, the audio signal maintains a striking similarity to the eliciting stimulus.¹² Because of the transparency of this subcortical response, it is possible to compare the response *timing* and frequency composition to the corresponding features of the stimulus ([Fig. 1](#)). *Timing* features (including sound onsets, offsets, and format transitions) are represented in the brain stem response as large transient peaks, whereas *pitch* (F_0) and *timbre* (harmonics up to about 1000 Hz) information is represented as interspike intervals that match the periodicity of the signal, a phenomenon known as phase locking.^f By means of commonly employed digital signal processing tools, such as

autocorrelation^g and Fourier analysis,^h features relating to stimulus *pitch* and *timbre* can be extracted from the response. As a consequence of being such a highly replicable measure, incredibly subtle differences in the *timing* and phase locking of the ABR are indicative of sensory processing malleability and abnormality.



Figure 1

Schematic representation of *timing*, *pitch*, and *timbre* in the stimulus (black) and brain stem response (gray) waveforms. Top: The full view of the time-domain stimulus waveform “da.” The temporal features of the stimulus, including the sound offset and onset, are preserved in the response. The gray box demarcates six cycles of the fundamental frequency (F_0); a blowup of this section is plotted in the middle panel. Middle: Major waveform peaks occur at an interval of 10 ms (i.e., the periodicity of a 100-Hz signal). This stimulus periodicity, which elicits the perception of *pitch*, is faithfully represented in the response. Bottom: The left panel shows a closeup of an F_0 cycle. The harmonics of the stimulus are represented as small-amplitude fluctuations between the major F_0 peaks in the stimulus and response. In the right panel, the stimulus and response are plotted in the spectral domain. Frequencies important for the perception of *pitch* (100 Hz) and *timbre* (frequencies at multiples of 100 Hz) are maintained in the brain stem response.

Subcortical Representation of *Pitch*

Musicians have extensive experience manipulating *pitch* within the context of music. Work by the Kraus Laboratory^{9,13,14} shows that lifelong musical training is associated with heightened subcortical representations of both *musical* and *linguistic pitch*, suggesting transfer effects from music to speech processing.

Musacchia *et al.*¹⁴ employed an audiovisual (AV) paradigm to tap into the multisensory nature of music. Given that music performance involves the integration of auditory, visual, and tactile information, we hypothesized that lifelong musical practice would influence AV integration. Subcortical responses were compared in three conditions: AV, auditory alone (A), and visual alone (V). In the AV condition, subjects watched and listened to a movie of a person playing the cello or saying “da.” In the A condition, no movie was displayed, and in the V condition, no sounds were presented. For both musicians and nonmusicians, the *pitch* responses to both speech and music were larger in the multimodal condition (AV) compared to unimodal A condition. However, musicians showed

comparatively larger *pitch* response in both A and AV conditions (AV responses are plotted in [Fig. 2](#)), and more pronounced multimodal effects, that is, greater amplitude increase between A and AV conditions. In addition, *pitch* representation strongly correlated with years of musical practice, such that the longer a person had been playing, the larger the *pitch* response ([Fig. 3](#), top). When the cortical responses to the AV condition were examined, this *pitch* representation was positively correlated with the steepness of the P1–N1 slope, such that the sharper (i.e., more synchronous) the cortical response, the larger the *pitch* representation.⁹ Other aspects of these multisensory responses will be explored in the sections relating to subcortical representation of *timbre* and *timing*. Taken together these data indicate that multisensory training, such as is acquired with musical experience, has pervasive effects on subcortical and cortical sensory encoding mechanisms for both musical and speech stimuli and leads to training-induced malleability of sensory processing.

Figure 2

Grand average brain stem responses to the speech syllable “da” for both musician (red) and non-musician (black) groups in the audiovisual condition. Top: Amplitude differences between the groups are evident over the entire response waveform. These differences translate into enhanced *pitch* and *timbre* representation (see bottom panel). Auditory and visual components of the speech stimulus (man saying “da”) are plotted on top. Middle: Musicians exhibit faster (i.e., earlier) onset responses. The grand average brain stem responses in the top panel have been magnified here to highlight the onset response. The large response negativity (shaded region) occurs on average ~0.50 ms earlier for musicians compared to nonmusicians. Bottom. Fourier analysis shows musicians to have more robust amplitudes of the F₀ peak (100 Hz) and the peaks corresponding to the harmonics (200, 300, 400, 500 Hz) (left). To illustrate frequency tracking of *pitch* and *harmonics* over time, narrow-band spectrograms (right) were calculated to produce time–frequency plots (1-ms resolution) for the musician (right top) and non-musician groups (right bottom). Spectral amplitudes are plotted along a color continuum, with warmer colors corresponding to larger amplitudes and cooler colors representing smaller amplitudes. Musicians have more pronounced harmonic tracking over time. This is reflected in repeating parallel bands of color occurring at 100 Hz intervals. In contrast, the spectrogram for the nonmusician group is more diffuse, and the harmonics appear more faded (i.e., weaker) relative to the musician group. (Adapted from Musacchia *et al.*[9,14](#)) (In color in *Annals* online.)



Figure 3

Neural enhancement varies according to the extent (top) and onset (bottom) of musical practice. Top: The number of years (over the last 10 years) of consistent practice is correlated with the strength of subcortical *pitch* encoding. Thus, the longer an individual has been practicing music, the larger the F0 amplitude. (Adapted from Musacchia *et al.*[14](#)) Bottom: The precision of brain stem *pitch* tracking is associated with the age that musical training began. Subjects who started earlier show a higher degree of *pitch* tracking. [N.B.: “Perfect” *pitch* tracking (i.e., no deviation between the stimulus *pitch* trajectory and response *pitch* trajectory) would be plotted as a 1 along the y-axis.] (Adapted from Wong *et al.*[13](#))

In music and language, *pitch* changes convey melodic and semantic or pragmatic information. Recently, a number of studies have looked at the representation of linguistic *pitch* contours (i.e., sounds which change in *pitch* over time) in the brain stem response. In Mandarin Chinese, unlike English, *pitch* changes signal lexical semantic changes. Compared to native English speakers, Mandarin Chinese speakers have stronger and more precise brain stem phase locking to Mandarin *pitch* contours, suggesting that the subcortical representation of *pitch* can be influenced by linguistic experience.[15,16](#) Using a similar paradigm, we explored the idea that musical *pitch* experience can lead to

enhanced linguistic *pitch* tracking.¹³ ABRs were recorded to three Mandarin tone contours: tone 1 (level contour), tone 2 (rising contour), and tone 3 (dipping contour). Musically trained native English speakers, with no knowledge of Mandarin or other tone languages, were found to have more accurate tracking of tone 3 ([Fig. 4](#)), a complex contour not occurring at the lexical (word) level in English.¹⁷ In addition, we found that the accuracy of *pitch* tracking was correlated with two factors: years of musical training and the age that musical training began ([Fig. 3](#), bottom). The differences between musicians and nonmusicians were less pronounced for tone 2 and not evident for tone 1. In contrast to tone 3, which only occurs at the phrase level in English, tones 1 and 2 are found at the word and syllable level. Taken together with the finding that musicians exhibit distinctive responses to emotionally salient *pitch* cues¹⁸ and enhanced *pitch* elements in musical chords²³ (reviewed below), we concluded that musical training alters subcortical sensory encoding of dynamic *pitch* contours, especially for complex and novel stimuli.



Figure 4

Pitch tracking plots from a musician (left) and nonmusician (right). The thin black line represents the *pitch* contour of the stimulus (Mandarin tone 3), and the thick gray line represents the extracted *pitch* trajectory of the brain stem response. The musician's brain response follows the *pitch* of the stimulus more precisely, a phenomenon known as pitch tracking. (Adapted from Wong *et al.*[13](#))

The studies reviewed above investigated the effects of lifelong auditory (linguistic and musical) experience on the subcortical representation of *pitch*. Recent work from Song *et al.*[19](#) suggests that lifelong experience may not be necessary for engendering changes in the subcortical representation of *pitch*. In fact, we found that as few as eight training sessions (30 mins each) can produce more accurate and more robust subcortical *pitch* tracking in native-English-speaking adults. Interestingly, improvement occurred only for the most complex and least familiar *pitch* contour (tone 3).

Unlike musicians who have heightened *pitch* perception,[20,21](#) some individuals with autism spectrum disorders (ASD) are known to have issues with *pitch* perception in the context of language.

For example, these individuals often cannot take advantage of the prosodic aspects of language and have difficulty distinguishing a question (rising *pitch*) from a statement (level or falling *pitch*). Russo *et al.*[22](#) explored whether this prosodic deficit was related to subcortical representation of *pitch*. We found that a subset of autistic children showed poor *pitch* tracking to syllables with linearly rising and falling *pitch* contours. Given that the subcortical representation of *pitch* can be enhanced with short-term linguistic *pitch* training and lifelong musical experience, this suggests that some children with ASD might benefit from an auditory training paradigm that integrates musical and linguistic training as a means of improving brain stem *pitch* tracking.

Subcortical Representation of *Timbre*

A growing body of research is showing that musicians represent the harmonics of the stimulus more robustly than their nonmusician counterparts.[9,18,23](#) This is evident for a whole host of stimuli including speech and emotionally affective sounds as well as musical sounds. Lee *et al.*[23](#) recorded brain stem responses to

harmonically rich musical intervals and found that musicians had heightened responses to the harmonics, as well as the combination tonesⁱ produced by the interaction of the two notes of the interval. In music, the melody is typically carried by the upper voice and the ability to parse out the melody from other voices is a fundamental musical skill. Consistent with previous behavioral and cortical studies,[24–27](#) we found that musicians demonstrated larger subcortical responses to the harmonics of the upper note relative to the lower note. In addition, an acoustic correlate of consonance perception (i.e., temporal envelope) was more precisely represented in the musician group. When two tones are played simultaneously, the two notes interact to create periodic amplitude modulations. These modulations generate the perception of “beats,” “smoothness,” and “roughness,” and contribute to the sensory consonance of the interval. Thus by actively attending to the upper note of a melody and the harmonic relation of concurrent tones, musicians may develop specialized sensory systems for processing behaviorally relevant aspects of musical signals. These specializations likely occur throughout the course of musical training—a

viewpoint supported by a correlation between the length of musical training (years) and the extent of subcortical enhancements.

The link between behavior and subcortical enhancements is also directly supported by Musacchia *et al.*,⁹ who found that better performance on a *timbre* discrimination task was associated with larger subcortical representations of *timbre*. *Timbre* was also an important distinguishing factor for separating out musicians from nonmusicians. As a group, the musically trained subjects had heightened representation of the harmonics ([Fig. 2](#), bottom).

Furthermore, when the subjects were analyzed along a continuum according to the age musical training began, subjects who started at a younger age were found to have larger *timbre* representations compared to those who began later in life. In addition, a correlation was found between cortical response *timing* and subcortical *timbre* encoding, which may be indicative of cortical structures being active in the processing of more subtle stimulus features.

Subcortical Representation of *Timing*

Timing measures provide insight into the accuracy with which the brain stem nuclei synchronously respond to acoustic stimuli. The hallmark of normal perception is an accurate representation of the temporal features of sound. In fact, disruptions on the order of fractions of milliseconds are clinically significant for the diagnosis of hearing loss, brain stem pathology, and certain learning disorders. Compared to normally hearing nonmusicians, musicians have more precise subcortical representation of *timing*, resulting in earlier (i.e., faster) and larger onset peaks^{14,18} ([Fig. 2](#), middle).

Furthermore, the results of these studies suggest an intricate relationship between years of musical practice and neural representation of *timing*. Taken together, the outcomes of our correlational analyses show that subcortical sensory malleability is dynamic and continues beyond the first few years of musical training.

[Go to:](#)

Summary: Music Experience and Neural Plasticity

Transfer Effects

By binding together multimodal information and actively engaging cognitive and attentional mechanisms, music is an effective vehicle for auditory training.[29,30](#) By showing that the effects of musical experience on the nervous system's response to sound are pervasive and extend beyond music,[9,13,14,18,31](#) work from our laboratory fits within the larger scientific body of evidence. We find transfer effects between the musical domain and the speech domain resulting in enhanced subcortical representation of linguistic stimuli.[9,13,14](#) However, these enhancements are not only specific to musical and linguistic stimuli, but also occur with non-linguistic emotionally rich stimuli as well. Strait *et al.*[18](#) (also appearing in this volume[31a](#)) recorded ABRs to the sound of a baby's cry, an emotionally laden sound. Compared to the nonmusician cohort, musicians showed enhanced *pitch* and *timbre* amplitudes to the most spectrally complex section of the sound, and attenuated responses to the more periodic, less complex section. These results provide the first biological evidence for enhanced perception of emotion in musicians[32,33](#) and

indicate the involvement of subcortical mechanisms in processing of vocally expressed emotion. Another compelling finding is that extensive auditory training can lead to both enhancement and efficiency (i.e., smaller amplitudes are indicative of allocation of fewer neural resources) of subcortical processing, with both enhancement and economy being evident in the subcortical response to a single acoustic stimulus. This finding reinforces the idea that subcortical responses to behaviorally relevant signals are not hardwired, but are malleable with auditory training.

The multisensory nature of music may also have an impact on vocal production by engaging auditory/vocal-motor mechanisms. Stegemöller and colleagues³¹ recorded speech and song samples from musicians and non-musicians. Vocal productions were analyzed using a statistical analysis of frequency ratios.³⁴ The vocal productions (speech and music) of both groups showed energy concentrations at ratios corresponding to the 12-tone musical scale. However, musicians' samples were smoother and had fewer deviant (i.e., non 12-tone ratio) peaks ([Fig. 5](#)), showing

that musicians had less harmonic jitter in their voices. This pattern was apparent even in the speech condition, where nonmusicians were found to differ from the vocally trained subjects in the musician group. This suggests that musical vocal training has an impact on vocal tract resonance during speech production. Also notable is that the musicians who did not undergo vocal training (instrumentalists) had smoother spectra for the song samples. Therefore, exposure to the 12-tone scale through instrumental training can be seen to influence vocal production, indicating a transfer from the auditory to the motor modalities.



Figure 5

Normalized spectra of speech (top two traces) and song (bottom two traces) tokens for non-musicians and vocalists. Prominent peaks in the spectra correspond to the intervals of the 12-tone scale. Unison, Perfect 4th, Perfect 5th, Major 6th, and Octave are labeled and represent the most well-defined spectral peaks in the speech and song tokens. Compared to nonmusicians, vocalists and professional musicians (not plotted) have smoother normalized spectra which include fewer unexpected (non-12-tone interval) peaks. The encircled portion of (A) is magnified in (B) to show the decrease in the number of unexpected peaks from speech to song, and from no musical experience to trained vocal experience. (Adapted from Stegemöller *et*

al.³¹)

Subcortical Enhancements and the Interaction of Top-down Processes

At first blush, it would appear that musical training is akin to a volume knob, leading to musicians' processing sounds as if they were presented at a louder decibel level. While it is clear that musicians show subcortical enhancements for *pitch*, *timbre*, and *timing*, a simple stimulus-independent gain effect cannot explain all of the results reviewed above. A better analogy is that musical training helps to focus auditory processing, much in the same way that glasses help to focus vision, and that this leads to clearer and more fine-grained subcortical representations. If only a gain effect was operative, we might expect all stimuli and all stimulus features to show more or less equivalent enhancements. However, available data do not support this stimulus-independent view. What we find instead is that only certain stimuli¹³ or certain aspects of the stimuli are enhanced in musicians.^{14,18,23} So while musical training might help focus auditory processing at a subcortical level, it does not do so blindly. Instead the behavioral relevance and complexity of the stimulus likely

influences how the sensory system responds. This suggests that higher-level cognitive factors are at play. In order to obtain auditory acuity, musicians actively engage top-down mechanisms, such as attention, memory, and context, and it is this binding of sensory acuity and cognitive demands that may in fact drive the subcortical enhancements we observe in musicians. Our findings suggest that higher-order processing levels (i.e., cortical) have efficient feedback pathways to lower-order (i.e., brain stem) processing levels. This top-down feedback is likely mediated by the corticofugal pathway, a vast track of efferent fibers that link together the cortex and lower structures.[35–38](#) While the corticofugal system has been extensively studied in animal models, the direct involvement of this efferent system in human auditory processing has also been demonstrated by Perrot and colleagues.[39](#) In the animal model, the corticofugal system works to fine-tune subcortical auditory processing of behaviorally relevant sounds by linking learned representations and the neural encoding of the physical acoustic features. This can lead to short-term plasticity and eventually long-term reorganization of subcortical sound encoding (for a

review see Suga *et al.*[35](#)). Importantly, corticofugal modulation of specific auditory information is evident in the earliest stages of auditory processing.[6](#) It is therefore our view that corticofugal mechanisms apply to human sensory processing, and can account, at least in part, for the pattern of results observed in musicians. Consistent with this corticofugal hypothesis and observations of experience-dependent sharpening of primary auditory cortex receptive fields,[7,40](#) we maintain that subcortical enhancements do not result simply from passive, repeated exposure to musical signals or pure genetic determinants. Instead, the refinement of auditory sensory encoding is driven by a *combination* of these factors and behaviorally relevant experiences, such as lifelong music making. This idea is reinforced by correlational analyses showing that subcortical enhancements vary as a function of musical experience[9,13,14,18,23](#) ([Fig. 3](#)).

Go to:

When Auditory Processing Goes Awry

Impaired auditory processing is the hallmark of several clinical conditions, such as auditory-processing disorder

(APD), a condition characterized by difficulty perceiving speech in noisy environments. Work from our laboratory has shown that a significant subset of children with language-based learning problems, such as dyslexia, where APD is common, show irregular subcortical representations of *timing* and *timbre* (harmonics), but not *pitch*.^{28,41} This pattern is consistent with the phonological processing problems inherent in reading disorders. Our research into the subcortical representation of speech in the learning-impaired population has been translated into a clinical tool, BioMARK (Biological Marker of Auditory Processing; see Clinical Technologies at <http://www.brainvolts.northwestern.edu/>). This test provides a standardized metric of auditory encoding and can be used to disentangle roles of *pitch*, *timbre*, and *timing* in normal and disordered auditory processing.

For a significant number of children with reading disabilities, sound is atypically encoded at multiple levels of the auditory system—the auditory brain stem,^{28,41–44} the auditory cortex^{45–47} or both^{48–50}—

suggesting a complex interaction between subcortical and cortical levels. Thus, the deficits we find in language impairment, such as developmental dyslexia^{28,48} ([Fig. 6](#)) and ASD,²² might be the consequence of faulty or suboptimal corticofugal engagement of auditory activity.



[Figure 6](#)

Brain stem responses from a child with reading difficulties (top), a young adult with typical hearing (middle) and a professional musician (bottom). Note the differences in waveform morphology, with the musician having larger and more defined (sharper) peaks.

Further evidence for the dynamic nature of subcortical auditory processing can be found by studying the effects of short-term training in children. After undergoing an 8-week commercially available auditory training program, children with language-based learning impairments showed improved subcortical response *timing* for speech signals presented in background noise.⁵¹ Because the auditory training was not specific to speech perception in noise, it raises the possibility that training-induced brain stem plasticity was mediated by top-down, cortically driven processes, a conclusion also supported by work from de Boer and

Go to:

Cochlear Implants and Music Perception

Cochlear implants (CIs) have proven to be enormously successful in engendering speech perception, especially in quiet settings, yet music perception is still below par. This is perhaps not surprising given that CI processing strategies are primarily designed to promote speech perception and thereby provide only a rough estimation of spectral shape, despite comparably fine-grained temporal resolution. While both speech and music have spectral and temporal elements, the weighting of these elements is not the same: speech perception requires more temporal precision whereas music perception requires more spectral precision.[53](#) The CI user's poor performance on musical tasks can be explained in large part by this underlying CI processing scheme and the acoustic differences between speech and music.

Real-world music listening requires the integration of multiple cues including *pitch*, *timing* (e.g., tempo and rhythm), and *timbre* (e.g., instrument identification).

For research purposes, music can be analytically decomposed into perceptual tasks that tap into each individual element. The *pitch*, *timbre*, and *timing* model that we employ in our laboratory for studying brain stem responses is also a useful trichotomy for assessing CI performance on musical tasks. With respect to *timing* tasks, the general consensus in the CI literature is that CI users and normal-hearing listeners have nearly comparable performances, yet the CI users perform far below average on *timbre* and *pitch* tasks.[54–60](#) On *timbre* tasks, CI wearers often have a difficult time telling two instruments apart.[54–56,58,60](#) However, despite this well-documented performance, Koelsch and colleagues[61](#) have demonstrated that *timbral* differences can elicit subliminal cortical responses. This suggests that even though many CI users cannot formally acknowledge differences in sound quality, these differences may in fact be registered in the brain.

When it comes to *pitch* perception, CI users could be described as having an extreme form of amusia (tone deafness). For example, whereas normally hearing adults can easily tell the difference between two

adjacent keys on a piano (i.e., 1 semitone difference), for the average postlingually implanted CI wearer, the notes must be at least 7 keys apart.[54](#) However, even if implantation occurs later in life, recent work by Guiraud and colleagues,[62](#) indicates that CIs can help reverse the effects of sensory deprivation by reorganizing how spectral information is mapped in the cortex.

For CI users, rehabilitative therapy has traditionally focused on improving speech perception and production. Despite numerous anecdotal and case reports showing that music therapy is being integrated into the rehabilitative process, the effects of musical training after CI implantation have garnered little scientific attention. Nevertheless, two known published reports reinforce the idea that focused short-term training can improve *timbre* and *pitch* perception.[54,63](#)

While vocoded sounds—sounds that have been manipulated to simulate the input that CI users receive—cannot fully mimic the CI acoustic experience, they serve as a useful surrogate for studying how the nervous system deals with degraded sensory input before and

after training. Studies are currently under way in our laboratory to explore how the normal hearing system encodes *pitch*, *timbre*, and *timing* features of speech and musical stimuli, and their vocoded counterparts. Special attention will be paid to the relationship between musical experience and how vocoded and more natural conditions are differentially represented at subcortical and cortical levels.

Because of magnetic and electromagnetic interference from the CI transmitter, magnetoencephalography and magnetic resonance imaging cannot be performed while a person is wearing a CI. Although an electrical artifact can plague electrophysiological recordings from CI wearers, techniques have been developed to minimize these effects in cortical potentials.[64,65](#) ABRs to speech and music have the capacity to be a highly objective and revealing measure of auditory processing in normal subjects listening to vocoded sounds, and with technological advances speech- and music-evoked ABRs may eventually be recorded in CI users. This work would complement the existing literature that has documented the integrity and plasticity of the CI user's

subcortical auditory pathways using simple click stimuli.[66,67](#)

Furthermore, in order to promote large scale and cross-laboratory/cross-clinic comparisons there is a need for standardized measurements of electrophysiology (equivalent to BioMARK) and music perception in this population (for three examples of music tests, see Nimmons *et al.*,[68](#) Cooper *et al.*,[69](#) and Spitzer *et al.*[70](#)). The benchmark of an effective test is one that can track changes before and after training, and is also sensitive enough to keep up with advancing CI technologies.

Speech and music perception are without question constrained by the current state of CI technology. However, technology alone cannot explain the highly variable performance across implantees, including the exceptional cases of children and adults who demonstrate near-normal *pitch* perception and production.[71,72](#) These “super-listeners” serve as beacons for where commonplace CI performance can aspire in the near future.

While most CI wearers have limited musical experience

before implantation,[73](#) a growing number of trained musicians are receiving implants. These individuals seem to have an advantage when it comes to music perception through a CI, especially for *pitch* perception. This underscores the important role that music experience plays in shaping sensory skills and lends further support for experience-dependent corticofugal (top-down) modulation of cortical and subcortical auditory pathway.[13,35,39,74](#) Through the use of electrophysiology and standardized music tests, we will gain better insight into the biological processes underlying super-listeners and ordinary listeners, which will ultimately lead to more refined CI technology and improved music enjoyment among CI users.

[Go to:](#)

Conclusion and Future Outlook

Subcortical auditory processes are dynamic and not hardwired. As discussed here, auditory sensory processing interacts with other modalities (e.g., visual and motor influences) and is influenced by language and music experience. The role of subcortical auditory processes in perception and cognition is far from

understood, but available data suggest a rich interplay between the sensory and cognitive processes involved in language and music, and a common subcortical pathway for these functions. It appears that in the normal system, music and language experience fundamentally shape auditory processing that occurs early in the sensory processing stream.[13–16,18,19,23](#) This top-down influence is likely mediated by the extensive corticofugal circuitry of descending efferent fibers that course from the cortex to the cochlea.[75](#) In order to facilitate sensory learning, the impaired system can capitalize on the shared biological resources underlying the neural processing of language and music, the impact music has on auditory processing and multisensory integration, and the apparent cognitive-sensory reciprocity.

Context-dependent Plasticity and Strength of Subcortical Encoding of Musical Sounds Independently Underlie Pitch Discrimination for Music Melodies

Author links open overlay panel

Xiaochen Zhang ^{a c}, Qin Gong ^{a b}

[Show more](#)

Add to Mendeley

Share

Cite

<https://doi.org/10.1016/j.neuroscience.2021.07.032>

[Get rights and content](#)

Highlights

- Subcortical neural correlates to music pitch perception were investigated with the frequency-following response (FFR).
- Weaker FFR inter-trial phase-locking (ITPL), greater context-dependent plasticity, and better music pitch perception were related.
- Larger FFR spectral strength was associated with better music pitch perception, only when the FFR ITPL was partialled out.
- Context-dependent plasticity, FFR ITPL, and spectral strength were more associated with pitch than with rhythm perception.

Abstract

Subcortical auditory nuclei contribute to pitch perception, but how subcortical sound encoding is related to pitch processing for music perception remains unclear. Conventionally, enhanced subcortical sound encoding is considered underlying superior pitch discrimination. However, associations between superior auditory perception and the context-dependent plasticity of subcortical sound encoding are also documented. Here, we explored the subcortical neural correlates to music pitch perception by analyzing frequency-following responses (FFRs) to musical sounds presented in a predictable context and a random context. We found that the FFR inter-trial phase-locking (ITPL) was negatively correlated with behavioral performances of discrimination of pitches in music melodies. It was also negatively correlated with the plasticity indices measuring the variability of FFRs to physically identical sounds between the two contexts. The plasticity indices were consistently positively correlated with pitch discrimination performances, suggesting the subcortical context-dependent plasticity underlying music pitch perception. Moreover, the raw FFR spectral strength was not significantly correlated with pitch discrimination performances. However, it was positively correlated with behavioral performances when the FFR ITPL was controlled by partial correlations, suggesting that the strength of subcortical sound encoding underlies music pitch perception. When the spectral strength was controlled by partial correlations, the negative ITPL-behavioral correlations were maintained. Furthermore, the FFR ITPL, the plasticity indices, and the FFR spectral strength were more correlated with pitch than with rhythm discrimination performances. These findings suggest that the context-dependent plasticity and the strength of subcortical encoding of musical sounds are independently and perhaps specifically associated with pitch perception for music melodies.

Tonality decides how much we can appreciate the music. ¹¹ Neural mechanisms of musical syntax and tonality, and the effect of musicianship Lei Jiang^{1,2†}, Ruiqing Zhang^{1†}, Lily Tao¹, Yuxin Zhang³, Yongdi Zhou^{4,5*}, Qing Cai^{1,6,7*} ^{3 4 1}Key Laboratory of Brain Functional Genomics (MOE & STCSM), Affiliated Mental Health⁵ Center, School of Psychology and Cognitive Science, East China Normal University, Shanghai 6 200062, China ^{2 7} School of Music, East China Normal University, Shanghai 200241, China ^{3 8} Shanghai High School International Division, Shanghai 200231, China ^{4 9} School of Psychology, Shenzhen University, Shenzhen 518061, China ^{10 5} Krieger Mind/Brain Institute, Johns Hopkins University, Baltimore, MD 21218, United States ^{6 11} Shanghai Key Laboratory of Magnetic Resonance, East China Normal University, Shanghai 12 200062, China ^{13 7} NYU-ECNU Institute of Brain and Cognitive Science, New York University Shanghai, 14 Shanghai 200062, China ^{† 15} These authors contributed equally. ¹⁶

*Correspondence: ¹⁷ Qing Cai, E-mail: miao.cai@gmail.com; ¹⁸ Yongdi Zhou, E-mail:

ydzhou.icn@gmail.com; 19 Keywords: music, tonality, syntax, hierarchical structure, fMRI, informational 20 connectivity available under aCC-BY-NC-ND 4.0 International license. (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made bioRxiv preprint doi: <https://doi.org/10.1101/2022.10.14.512259>; this version posted October 17, 2022. The copyright holder for this preprint Tonality decides how much we can appreciate the music. 221 Abstract 22 The neural basis for the processing of musical syntax has previously been examined almost 23 exclusively in classical tonal music, which is characterized by strictly organized hierarchical 24 structure. The present study investigated the neural mechanisms for processing musical syntax25 across genres varying in tonality - classical, impressionist, and atonal music - and, in addition, 26 examined how musicianship modulates such processing. Results showed that, first, the dorsal 27 stream, including bilateral inferior frontal gyrus and superior temporal gyrus, plays a key rolein28 the perception of tonality. Second, right fronto-temporal regions were crucial in allowing29 musicians to outperform non-musicians in musical syntactic processing; musicians also benefit 30 from a cortical-subcortical network including pallidum and cerebellum, suggesting more 31 auditory–motor interaction in musicians than in non-musicians. Third, left pars triangularis 32 carries out on-line computations independently of tonality and musicianship, whereas right pars33 triangularis is sensitive to tonality and partly dependent on musicianship. Finally, unlike tonal 34 music, processing of atonal music could not be differentiated from that of scrambled notes, both35 behaviorally and neurally, even among musicians. The present study highlightsthe importanceof36 studying varying music genres and experience levels, and provides a better understandingof 37 musical syntax and tonality processing and how such processing is modulated by music 38 experience. 39 1 Introduction 40 Throughout the history of humanity, music has been a key component in social and cultural 41 interactions. How people communicate with music, namely how listeners perceive music syntax42 has been the subject of investigation in neuroscience. Some have suggested parallels between43 music processing and language processing. Currently, however, the neural mechanisms of tonal 44 music perception are still uncertain. Some evidence has been provided by studies on Western45 classical music. The organization of pitches or chords in classical harmonic musical sequence46 tends to begin with the main tone or chord, and usually returns to the main tone or chord at the47 end. Other genres of music involve different structures, and may, thus, entail different processing48 mechanisms to classical music. 49 Animal studies have shown that, in marmosets, harmonic template neurons sensitive to spectral 50 regularity of harmonic complex sounds are distributed across the primary auditory cortexandthe51 neighboring primary-like rostral area (Feng & Wang, 2017). In humans, widely distributed52 frontal and temporal regions have been involved in the precessing of classical music. Among53 these regions, the left inferior frontal gyrus (IFG) has been suggested to be the most important 54 site offering computational resources for both linguistic and musical syntax (Patel, 2003, Patel et 55 al., 2008; Kunert et al., 2015). Electrophysiological studies have suggested that patients with56 lesions in left IFG show abnormal musical syntax processing and impaired behavioral 57 performance in the processing of irregular chord sequences, and that left IFG is the key region58 for the processing of syntax in a domain-general way (Sammler, Koelsch, & Friederici, 2011; 59 Patel et al., 2008). Furthermore, music processing, like language processing, may also involve60 shared dorsal and ventral neural networks, underlying structure and meaning processing 61 respectively (Koelsch & Siebel, 2005; Musso et al., 2015). The dorsal stream–includingIFG, available under aCC-BY-NC-ND 4.0 International license. (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made bioRxiv preprint doi: <https://doi.org/10.1101/2022.10.14.512259>; this version posted October 17, 2022. The copyright holder for

this preprint Tonicity decides how much we can appreciate the music. 362 anterior superior temporal gyrus (STG) and ventrolateral premotor cortex (PMC) – processes 63 harmonic relations and structural irregularities, predicts short-term upcoming harmonic 64 sequences (Koelsch & Siebel, 2005), and is involved independently of the type of musical 65 stimuli (Tillmann et al., 2006). The left IFG further connects to inferior parietal cortex and 66 middle temporal lobe through dorsal and ventral long association tracts (Musso et al., 2015). 67 Although previous studies have provided a good basis for the understanding of music processing, 68 so far almost all neuroscientific studies on music exclusively used Western classical music. 69 Classical Western music is characterized by strictly organized hierarchical structure, which may 70 not be the case across other music genres. It is important, therefore, to examine a variety of 71 music genres to provide a complete and unbiased picture (see also Brattico et al., 2013). Let us 72 take a closer look at two other music genres: impressionist music and atonal music. 73 Representative compositions of impressionism are partial to the diatonic scale. Impressionist 74 musicians such as Debussy divides an octave into six major second intervals of three kinds – 75 major second, major third, and tritone (Day-O’Connell, 2009). Atonal music exploits a 76 composition technique without the tonic center and functional relationship among notes or 77 chords. For example, in “A Survivor of Warsaw”, a representative atonal piece written by 78 Schoenberg, the twelve semitones are functionally equal, making it distinct from the major 79 minor system. Moreover, the size distribution of intervals in the scale of tonal music is generally 80 between one and three half-tones, and the grading progress is the main composition of the 81 melody lines. 82 In short, the diatonic scale in impressionist music and the combination of 12 equal half-tones in 83 atonal music both break the structural rules of classical music, either partially or completely. The 84 asymmetry of the scale, the limitation of sound levels, and the size distribution of intervals 85 within the scale are some important factors that differentiate tonal, impressionist, and atonal 86 music in music theory. According to the literature on music processing, if the interval 87 relationship to the tonal center (i.e. pitch-center relationship) disappears, the musical grammar 88 would be disrupted and listeners could feel weary (Lerdahl & Jackendoff, 1983). If this is the 89 case for atonal music, we should expect that the neural networks underlying the processing of the 90 regularities of pitch relationships and structure-based prediction to also work differently. 91 A further question is whether such neural activation is exclusively decided by the physical 92 features of musical stimuli, which is identical for all listeners; or if it rather reflects how the 93 music is perceived by individuals and, therefore, interacts with listeners’ music experience and 94 preference. For example, for a non-trained listener, music may simply be a series of notes and 95 beats, sometimes even a nuisance to the ear. For the romantic musician, in contrast, music can 96 communicate just as well, or even better than language. In other words, training and experience 97 matters. Previous findings have shown that the early right anterior negativity (ERAN) ERP 98 component is sensitive to music training (Koelsch et al., 2002b). A recent study further showed 99 that, in musicians, right IFG, as well as right posterior STG, superior temporal sulcus (STS), and 100 cerebellum are involved in the processing of musical structures, with resting state activity in 101 right IFG positively correlated with that in posterior STG and left Heschl’s gyrus (Bianco et al., 102 2016). However, only musicians were tested in that study, so it remains unclear how music 103 experience modulates music processing and whether this process interacts with tonality. 104 The present study aimed to investigate the neural mechanisms underlying the processing of available under aCC-BY-NC-ND 4.0 International license. (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made bioRxiv preprint doi: <https://doi.org/10.1101/2022.10.14.512259>; this version posted October 17, 2022. The copyright holder for this preprint Tonicity decides how much we can appreciate the music. 4105 musical syntax, as well as the impact of tonality and

expertise on such processing. To achieve this purpose, we included music genres that varied in tonality. Specifically, extending from previous studies on classical tonal music, we also examined impressionist music (relatively decreased tonality) and atonal music (no tonality). A second aim of the present study was to investigate how musicianship modulates musical structure processing, and how it interacts with different music genres – that is, whether music experience affects brain networks underlying music tonal syntactic processing.

2 Materials and Methods

2.1 Participants

Thirty-six healthy native Chinese speakers with normal hearing, recruited from East China Normal University or Shanghai Conservatory of Music, took part in this study. All participants were right-handed, confirmed using Edinburgh Handedness Inventory (Oldfield, 1971). Written informed consent was obtained from each participant, and the protocol of the present study was approved by the Committee on Human Research Protection at East China Normal University. All participants were paid for their participation.

Musicianship was determined using Music Experience Questionnaire. Half of the participants ($n=18$) were musicians (22.4 ± 2.1 years, 16 females) who majored in instrumental (17) or vocal (1) performance, and were immersed in a classical music environment for on average $3.3(2.4)$ hours per day; five of them reported having absolute pitch. They had on average 13.0 years of formal music training (± 3.2 , range 8 to 17 years), with an average age of onset of 5.5 years (± 1.3 , range 3 to 8 years). The other half of the participants ($n=18$) were non-musicians (21.3 ± 3.3 years, 13 females), who reported no prior experience in music training except one with a one-year experience in learning accordion and two with limited experience of playing piano or keyboard at young ages (these three participants took part in the study given their limited music experience and no music training in the last ten years, but their data were excluded in further analysis).

2.2 Materials

There were three experimental conditions, that is, three genres of music – classical/tonal, impressionist/pantonal, atonal – and three control conditions – their respective scrambled versions. In order to inspect more global and salient violations of tonal syntax, we adopted a method used in Levitin and Menon (2003), in which scrambled versions of musical pieces were included as baseline conditions to disrupt the musical structure, in other words the overall relationship between adjacent notes. Each of the three experimental conditions contained 40 phrases, selected from representative Western composers' masterpieces, as listed in Table 1. The phrases were reconstructed using Sibelius software to be synchronous, to have a similar number of notes (32 ± 2 notes), and similar intensity. Only the relative positions of the notes, or the internal organizational structure of the phrase, was preserved. By doing so, the low-level acoustic features such as tempo, loudness, and timbre were balanced across music genres and leave the music syntax intact. The mean duration available under aCC-BY-NC-ND 4.0 International license. (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made bioRxiv preprint doi: <https://doi.org/10.1101/2022.10.14.512259>; this version posted October 17, 2022. The copyright holder for this preprint Tonality decides how much we can appreciate the music. 5144 of the phrases was $6.2 (\pm 0.4)$ s. Scrambled versions were made by shuffling all of the notes within each of the original phrases, so that the relative pitch of adjacent notes was disrupted. The scrambled phrases were then rated by three professional musicians independently to ensure that the inner original organizational structures had been destroyed while the same notes were kept. To increase relative loudness of pitch in the noisy scanner environment, dynamic range compression was applied on all the pieces using the compressor effect of Audacity (Farbood, 150 2015). In addition to these stimuli, a 250-Hz pure tone (660 ms duration) was used as probe stimulus. Five such trials were included, inserted evenly between other trials, within each scanning session/

run, to ensure that participants were attending to the task. 154

2.3 Procedure 155

During fMRI scanning, participants were required to listen carefully to each phrase presented 156 (they were not informed that some are original and others are scrambled), and to press a button 157 with their right index finger when they hear the pure tone (which had been presented to them 158 outside before scanning). The same task was performed twice in the scanner. Each session/run 159 contained 125 trials: 20 trials for each of the six conditions plus five pure-tone probe-detection 160 trials. Each session/run started with a fixation of 10 s, and then all trials were presented in a 161 random order. Between each phrase, a 2-4-6 s blank interval was presented (see Figure 1A). 162 Stimuli were presented using E-Prime 2.0 software. 163 After scanning, participants listened to all phrases again, classified each piece into four 164 categories (classical/tonal, impressionist, atonal music, and random notes), rated the level of 165 confidence in his/their decision (from 1 = least confident to 5 = most confident), and familiarity 166 with the phrase (from 1 = least familiar to 5 = most familiar; see Figure 1B). 167

2.4 Data Acquisition 168

Whole-brain images were collected on a 3T Siemens Trio MR scanner, with a 32- channel head 169 coil. First, an anatomical image was obtained using a T1-weighted MPRAGE sequence (TR=170 2530 ms, TE = 2.34 ms, image matrix = 256 * 256, FoV = 256 mm, flip angle = 7° 171, voxel size=171 1*1*1mm, 192 slices). Functional MRI images were acquired using a T2*-weighted gradient 172 echo EPI sequence covering the whole brain (TR= 2400 ms, TE = 30 ms, image matrix =64*64, 173 FoV = 192 mm, flip angle = 81° 174, voxel size = 3*3*3mm, slice thickness = 3mm, 40 slices, 174 interleaved acquisition). Stabilization cushions were used to minimize head motion, and ear 175 plugs were worn by participants to reduce noise from the scanner during operation. Auditory 176 stimuli were presented using RT-300 (Resonance Technology, Canada). Behavioral data were 177 collected outside the MRI environment after scanning. available under aCC-BY-NC-ND 4.0 International license. (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made bioRxiv preprint doi: <https://doi.org/10.1101/2022.10.14.512259>; this version posted October 17, 2022. The copyright holder for this preprint Tonicity decides how much we can appreciate the music. 6178

Figure 1. Experimental procedure for (A) fMRI and (B) behavioral tasks. 179

2.5 Behavioral Data Analysis 180

Two-way mixed design ANOVA with Tukey's HSD comparison tests were performed separately 181 for the genre classification, confidence rating, and familiarity rating, with group (musician, non 182 musician) and musical syntax (classical, impressionist, atonal, random notes) as independent 183 factors. Data from two musicians were excluded because their accuracy for "randomnotes" were 184 outliers (0% and 2.5%). Note that for each participant and each genre, familiarity score was 185 calculated based on ratings for all phrases, and confidence score only took into account the 186 correctly classified trials. 187

2.6 Functional Imaging Data Analysis 188

Functional MRI data preprocessing and statistical analysis was carried out using SPM8 189 (www.fil.ion.ucl.ac.uk/spm). After slice-timing correction, the functional images were realigned 190 for headmotion correction. The functional and co-registered anatomical images were spatially 191 normalized to MNI space, and then smoothed using a Gaussian kernel with full width at half available under aCC-BY-NC-ND 4.0 International license. (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made bioRxiv preprint doi: <https://doi.org/10.1101/2022.10.14.512259>; this version posted October 17, 2022. The copyright holder for this preprint Tonicity decides how much we can appreciate the music. 7192

maximum (FWHM) of 5 mm. Head movements were checked for each subject using Artifact 193 Detection Tools (ART; www.nitrc.org/projects/artifact_detect) package. Time points 194 (scans/volumes) with motion outliers (≥ 2 mm) or outliers in global signal intensity (≥ 5 SD) were 195 recorded for nine participants. 196 Data from each participant were then analyzed

using a general linear model (GLM), with three music genre conditions (classical, impressionist, and atonal), three scrambled conditions, and the probe condition. Head movement parameters were included for each participant as regressors, and the above mentioned time points with motion or intensity outliers were omitted by including a single regressor for each in GLM. Familiarity scores from participants' behavioral ratings were included as parametric modulators for each condition to dissociate familiarity effects from main effects. We first examined whether there were significant differences between any two scrambled conditions (out of the three scrambled conditions) using a 2 (group) * 3 (condition) flexible factorial model at the group level. Given no significant main effect or interaction was found for the scrambled conditions, the three scrambled conditions were combined into one, referred to as the random notes condition (matching the music genre classification in behavioral analysis). A 2 (group) * 4 (musical syntax: classical, impressionist, atonal, random notes) flexible factorial model was used in further analysis. Given previous discoveries on the functional role of bilateral IFG in music processing, bilateral IFG (pars triangularis and pars opercularis) anatomical ROIs were selected from MarsBaR ROIs. Percent signal change relative to global brain signal was computed using MarsBar, to further investigate how the brain reacted to different music genres in musicians and non-musicians. To further derive the synchronous function of cortical regions that processed different music genres in musicians and non-musicians separately, informational connectivity analysis (Coutanche & Thompson-Schill, 2013) was conducted. The whole brain was segmented into 162 regions of interest (ROIs) based on Automated Anatomical Labeling (AAL) template (Tzourio-Mazoyer et al., 2002; Schmahmann et al., 1999). Four ROIs were excluded in further analysis because they have not been fully covered in certain participants while scanning. For each ROI, a representational dissimilarity matrix (RDM) of all 240 musical trials was computed based on β values extracted from all voxels for each participant. Then, for each ROI pair, the correlation coefficient was calculated between the two RDMs of the ROI pair and then transformed to Fisher's z values indicating representational similarity of general musical sentences processing between brain regions. After that, the correlation analysis was then performed separately for musicians and non-musicians to investigate the relationship between the z values of each region pair and the behavioral overall genre classification accuracy (representing each participant's general musical genre sensitivity). Informational connectivity analysis allows us to inspect the highly stimuli-dependence neural processing between brain regions, which offers a higher-order explanation than univariate analysis.

3 Results

3.1 Behavioral Results

available under aCC-BY-NC-ND 4.0 International license. (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made bioRxiv preprint doi: <https://doi.org/10.1101/2022.10.14.512259>; this version posted October 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Tonality decides how much we can appreciate the music. ANOVA on classification accuracy showed a main effect of group, $F(1,136)=25.058$, $p<0.001$, a main effect of musical syntax $F(3,136)=19.135$, $p<0.001$, and an interaction between group and musical syntax ($F(3,136)=4.837$, $p<0.01$). Tukey's HSD post-hoc test indicated that classical music and impressionist music were easier to identify than atonal music (HSD=24.118, $p<0.001$, HSD=20.881, $p<0.001$, respectively classical and impressionist) and random notes (HSD=15.662, $p<0.001$, HSD=12.425, $p<0.01$, respectively for classical and impressionist). The musician group classified classical and impressionist music better than atonal music (HSD=23.594, $p<0.001$, HSD=28.292, $p<0.001$, respectively for classical and impressionist) and random notes (HSD=22.552, $p<0.001$, HSD=27.25, $p<0.001$, respectively for classical and impressionist). The non-musician group was found

to have better knowledge only of classical 243 compared to atonal genre (HSD=24.583, $p < 0.001$). Within music genres, a significant group difference was only found for impressionist music classification, with musicians outperforming non-musicians (HSD=28.083, $p < 0.001$; see Figure 2A). For familiarity ratings, ANOVA showed only a significant main effect of musical syntax, $F(3,140)=21.91$, $p < 0.001$. Post-hoc tests showed that classical musical phrases were rated as significantly more familiar than atonal musical phrases (HSD=0.56, $p < 0.05$), and significantly more familiar than random notes (HSD=0.638, $p < 0.01$; see Figure 2B). For confidence ratings, ANOVA showed significant main effects of groups, $F(3,142)=9.079$, $p < 0.01$, and of musical syntax, $F(3,140)=23.657$, $p < 0.001$. Musicians were overall more confident than non-musicians in their genre classifications (HSD=0.896, $p < 0.001$). Confidence was significantly higher when classifying classical music comparing to atonal music (HSD=0.727, $p < 0.01$) and random notes (HSD=0.676, $p < 0.01$; see Figure 2C). Figure 2. Behavioral results for musicians and non-musicians for (A) percentage correct genre classification, (B) familiarity ratings (1, least familiar~5, most familiar) in musicians and non-musicians, and (C) confidence ratings (1, least confident~5, most confident).

3.2 Functional Imaging Results

The group-level factorial analysis showed a significant interaction between group and musical available under aCC-BY-NC-ND 4.0 International license. (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made bioRxiv preprint doi: <https://doi.org/10.1101/2022.10.14.512259>; this version posted October 17, 2022. The copyright holder for this preprint

Tonality decides how much we can appreciate the music. syntax, which involved activation in right postcentral areas, left supplementary motor area (SMA), left middle temporal gyrus (MTG), left hippocampus, and bilateral superior frontal gyrus (SFG). The main effect of musical syntax was observed in bilateral superior temporal regions, bilateral IFG pars triangularis extending to left insula, bilateral superior medial frontal areas, bilateral precentral gyrus, right SFG, right middle frontal gyrus (MFG), left angular gyrus, right supramarginal gyrus, left SMA, and bilateral cerebellum. The main effect of group was observed in bilateral cerebellum, bilateral precentral gyrus, right SFG, right superior temporal pole, bilateral inferior temporal gyrus, left amygdala, right STG, and bilateral IFG pars opercularis (all p 's < 0.001 , α sim corrected; see Table 2). Overall, classical/tonal music (compared to random notes) involved significant activation in bilateral STG, left inferior frontal regions (including pars triangularis, pars opercularis, and pars orbitalis), right inferior frontal regions (including pars opercularis and insula), bilateral precentral gyrus, bilateral SMA, and bilateral cerebellum. Impressionist music (compared to random notes) involved significant activation in bilateral STG, right superior temporal pole, right MTG, left IFG pars opercularis and pars triangularis, right IFG pars triangularis, left supramarginal gyrus, right hippocampus, right precentral gyrus, right SMA, and left cerebellum. When contrasting classical over impressionist music processing, classical condition involved greater activation in right IFG pars opercularis and left insula compared to impressionist; the reverse contrast involved more right IFG pars triangularis, right precentral gyrus, bilateral STG, bilateral superior temporal pole, and left SMA. When comparing to atonal music, classical music showed more activation in bilateral STG and MTG, right IFG pars triangularis and pars opercularis, left IFG pars opercularis and insula, bilateral precentral gyrus, bilateral SMA, and bilateral cerebellum; impressionist music showed more activation in bilateral STG and MTG, left IFG pars triangularis and pars orbitalis, right IFG pars triangularis, bilateral SMA, bilateral putamen, and bilateral cerebellum. Atonal music involved more activation in bilateral MTG than classical music, with no areas showing greater activation compared to impressionist music (all p 's < 0.001 , α sim

corrected; see Figure 3A-C). 287 Simple effects were further analyzed using t-tests to investigate how the processing of musical 288 structure was modulated by musicianship. For classical music processing, musicians showed 289 greater activation in right STG, right IFG pars triangularis, right superior medial frontal gyrus, 290 right inferior parietal gyrus, and bilateral SMA, whereas bilateral anterior cingulate cortex (ACC) 291 were more activated in non-musicians (all p 's < 0.001, alphasim corrected; see Table 2). available under aCC-BY-NC-ND 4.0 International license. (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made bioRxiv preprint doi: <https://doi.org/10.1101/2022.10.14.512259>; this version posted October 17, 2022. The copyright holder for this preprint Tonicity decides how much we can appreciate the music. 10292 Figure 3. Brain activation for musical syntax processing (all results alphasimcorrected at 293 p <0.001, unless otherwise stated). (A) Classical music compared to randomnotes; (B) 294 impressionist music compared to random notes; (C) Comparisons among musical genres: a) 295 classical compared to impressionist music (alphasim corrected at p <0.01 for illustration); b) 296 impressionist compared to classical music (alphasim corrected at p <0.01 for illustration); c) 297 classical compared to atonal music; d) impressionist compared to atonal music; e) atonal 298 compared to classical music; f) atonal compared to impressionist music; (D) difference between 299 groups for classical music: a) musicians compared to non-musicians; b) non-musicians compared 300 to musicians; c) percent signal change in left and right IFG; (E) differences between groups for 301 impressionist music: a) musicians compared to non-musicians; b) non-musicians compared to 302 musicians. 303 When processing impressionist music, musicians showed more activation in left cerebellum 304 (Vermis 9) compared to non-musicians; non-musicians showed more activation in bilateral 305 hippocampal gyrus, bilateral postcentral gyrus, MTG, SFG, insula, precuneus, and middle 306 occipital lobe in the left hemisphere (all p 's < 0.001, alphasim corrected; see Table 2). 307 Lastly, for atonal music, no significant differences were found between musicians and non 308 musicians (p 's < 0.001, alphasim corrected). available under aCC-BY-NC-ND 4.0 International license. (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made bioRxiv preprint doi: <https://doi.org/10.1101/2022.10.14.512259>; this version posted October 17, 2022. The copyright holder for this preprint Tonicity decides how much we can appreciate the music. 11309 For the ROI analysis on bilateral IFG (see Figure 3D(c)), left pars triangularis showed no 310 significant effects of music genre ($F(3) = 0.842$, $p = 0.472$) or group ($F(1) = 0.000$, $p = 0.984$), or 311 their interaction ($F(1,3) = 0.218$, $p = 0.884$). A significant musical syntax main effect ($F(1,3) = 312 0.048$, $p < 0.01$) was found for right pars triangularis, specifically, both classical ($t = 2.76$, $p < 313 0.01$, 95% CI = [0.0112, 0.0222]) and impressionist music ($t = 3.28$, $p < 0.01$, 95% CI = 314 [0.0683, 0.0745]) have greater signal change than atonal music. For both left and right pars 315 opercularis, there were significant group differences (left: $F(1,3) = 4.61$, $p < 0.05$; right: $F(1,3) = 316 4.65$, $p < 0.05$), with percent signal change in musicians greater than in non-musicians (left: $t = 317 2.15$, $p < 0.05$, 95% CI = [0.0014, 0.0322]; right: $t = 2.18$, $p < 0.05$, 95% CI = [0.0553, 0.0459]). 318 Informational connectivity between right Heschl's gyrus and right superior temporal pole was 319 positively correlated with behavioral classification accuracy in musicians ($r = 0.69$, FDR 320 corrected at $q = 0.005$); informational connectivity between right IFG pars orbitalis and left 321 pallidum was also positively correlated with behavioral classification accuracy in musicians ($r = 322 0.79$, FDR corrected at $q = 0.005$). Informational connectivity between cerebellum (cerebellar 323 vermis 7, VER7) and both left and right STG was negatively correlated with behavioral accuracy 324 in non-musicians (left STG: $r = -0.89$, $q = 0.0001$; right STG: $r = -0.74$, $q = 0.0001$) (see Figure 325 4). available under aCC-BY-NC-ND 4.0 International license. (which was not certified by

peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made bioRxiv preprint doi: <https://doi.org/10.1101/2022.10.14.512259>; this version posted October 17, 2022. The copyright holder for this preprint Tonality decides how much we can appreciate the music. 12326 Figure 4. Upper panel: illustration for computing informational connectivities between ROIs for 327 all the participants. Lower panel: correlations between informational connectivities and 328 behavioral classification accuracy in musicians and non-musicians, for (A) connectivity between 329 right Heschl's gyrus and right superior temporal pole, (B) connectivity between right IFG pars 330 orbitalis and left pallidum, (C) connectivity between the left superior temporal gyrus and 331 cerebellum, and (D) connectivity between right superior temporal gyrus and cerebellum. 332 4 Discussion 333 The present study investigated the neural mechanisms underlying tonality and musical syntax 334 processing, as well as the role of music training on such processing. Musicians and non 335 musicians listened to phrases from classical, impressionist, and atonal music genres inside an 336 MRI scanner, and performed a classification task outside the scanner. The results elucidated the 337 on-line processing mechanisms of musical syntax across different genres, and showed how 338 musicianship impacted the neural response to different musical syntax. available under aCC-BY-NC-ND 4.0 International license. (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made bioRxiv preprint doi: <https://doi.org/10.1101/2022.10.14.512259>; this version posted October 17, 2022. The copyright holder for this preprint Tonality decides how much we can appreciate the music. 13339 4.1 Musical syntax, tonality, and musicianship 340 For overall processing of hierarchical structure in music, neural response was observed in 341 bilateral temporal lobes, IFG, postcentral gyrus, and cerebellum. This finding indicates the 342 engagement of the dorsal stream in decoding musical syntax, where auditory information is 343 transformed to motor actions, and that this engagement is stronger among musicians than non 344 musicians in the presence of tonality, as discussed later. 345 For Western classical music perception, musicians and non-musicians achieved equally high 346 accuracy in behavioral classification, but with a higher confidence rating in musicians, 347 suggesting that musicians took advantage of their expertise to analyze the musical notes. 348 Bilateral anterior superior temporal areas, bilateral left inferior frontal regions extending to 349 bilateral precentral gyrus, insula, SMA, and cerebellum were engaged in the processing of tonal 350 music, in line with previous studies (Koelsch et al., 2002a, 2013; Tillmann et al. 2006, Sammler 351 et al., 2013, Farbood et al., 2015). However, whereas previous studies suggested that people 352 perceive musical syntax implicitly, regardless of music training (Koelsch et al., 2000, Bigand et 353 al., 2006), our results showed that music experience modulated neural activation in classical 354 tonal music processing, though non-musicians and musicians performed equally well in 355 behavioral classifications. Specifically, differences between musicians and non-musicians in 356 neural activation were observed in a right-lateralized front-parieto-temporal network, covering 357 right STG, right IFG pars triangularis and superior medial frontal gyrus, right inferior parietal 358 gyrus, and bilateral SMA. Together with previous studies showing the role of right IFG in 359 musical syntax processing (Cheung et al., 2018) and structural brain changes in right fronto 360 temporal regions linked to music training (Sato, Kirino, & Tanaka, 2015; James et al., 2014), the 361 present findings suggest that the left fronto-temporal neural network plays an important role in 362 musical syntactic processing in a domain-general and experience-independent way, and that the 363 right fronto-temporal cortical areas contribute to musical syntactic processing in a musicianship 364 modulated way. 365 For impressionist music, musicians showed significantly higher accuracy in behavioral 366 classification, as well as stronger activation in left cerebellum than non-musicians. A closer look 367 at the neural basis among musicians and non-musicians showed that bilateral STG and bilateral 368

IFG pars triangularis were engaged in both groups, whereas right IFG was significantly recruited only among musicians. These results suggest that the minor disruption of tonality rules in impressionist music could weaken the functions of the left IFG in resolving musical syntax. The right IFG, on the other hand, still played an important role in musical syntax processing, particularly with music training. Together with the results of classical music processing, these results indicate that music experience has an impact on the neural response to syntactic processing of tonal music – both classical tonal and impressionist (reduced tonality). For atonal music, there were no differences between musicians and non-musicians in either neural activation or behavioral classification performance. Furthermore, atonal music could not be differentiated from random notes, either neurally or behaviorally, even among musicians. This is likely due to a lack of pitch-center relationship in atonal music, leading to an absence of structural information processing. Given that previous studies on atonal music suggested that available under a CC-BY-NC-ND 4.0 International license. (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made bioRxiv preprint doi: <https://doi.org/10.1101/2022.10.14.512259>; this version posted October 17, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Tonality decides how much we can appreciate the music. Familiarity has some effects on induced emotional responses to atonal music (Daynes, 2000), or that listeners can learn to detect or expect the avoidance of pitch repetition (Krumhansl, Sandell & Sergeant, 1987; Ockelford & Sergeant, 2012), it would be of interest for future studies to investigate whether atonal music is processed differently among musicians with more varied experiences and those with expertise in atonal music, such as the composers and conductors who have developed a positive taste for atonal music.

4.2 Cortical and subcortical neural networks for musical syntax processing

The IFG has been deemed to be a storage buffer required to process sequences with regular structure (Fitch and Martins, 2014). Within the IFG, left pars triangularis, a part of Broca's area, has been suggested to be involved in domain-general processing, playing a crucial role in sequence regularities, and particularly being the site of a buffer zone for syntactic computations (Sammler et al., 2011; Fitch and Martins, 2014). Previous studies have further put forward a shared resource system for domains of both language and music, seated in Broca's area (Patel, 2003; Fedorenko et al., 2009). The role of right IFG is less clear, though some studies have suggested that the right inferior frontal area is crucial for processing specific musical syntax (Maess et al., 2001), or is sensitive to music training (Oechslin et al., 2013; Koelsch et al., 2002b). In the present study, the left pars triangularis was engaged in the syntactic processing of classical music equally for musicians and non-musicians. The right pars triangularis and pars opercularis, on the other hand, were involved to a greater extent among musicians compared to non-musicians in the syntactic processing of both classical and impressionist music. Percent signal change of different subregions of bilateral IFG further showed that right pars triangularis was sensitive to tonal differences, and that both left and right pars opercularis were sensitive to music experience differences. We therefore suggest a more precise division of labor of bilateral IFG regions in music processing: the left IFG pars triangularis carries out on-line unit relationship computations independently of music genre and music experience; the right IFG pars triangularis detects tonality and adjusts to tonal varieties, partly dependently of music experience; both left and right pars opercularis are modulated by music experience, with the right pars opercularis more dominantly so. We also found an involvement of right anterior temporal regions, together with right frontal regions, in musical syntactic processing, especially among musicians. Furthermore, informational connectivity results revealed that higher behavioral classification accuracy among musicians was accompanied by stronger functional

cooperation between right Heschl's gyrus and right superior temporal pole. According to previous findings, temporal resolution is better in left auditory cortices, whereas spectral resolution is better in right auditory cortices (Zatorre et al., 2002). Therefore, our results suggest that right temporal regions are more engaged in musicians to achieve better performance in detecting precise changes in frequency. Together with above-mentioned results on frontal regions, the present findings suggest that a right fronto-temporal network is crucial in allowing musicians to outperform non-musicians in musical syntactic processing. The neural processing of musical syntax engages not only cortical structures but also subcortical structures, such as basal ganglia, which has been found to be activated in the processing of available under aCC-BY-NC-ND 4.0 International license. (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made bioRxiv preprint doi: <https://doi.org/10.1101/2022.10.14.512259>; this version posted October 17, 2022. The copyright holder for this preprint Tonality decides how much we can appreciate the music. Musical beats and music-related emotions (Frisch et al., 2003; Kung et al., 2013). In the present study, neural recruitment of pallidum and the cerebellum was found for processing tonal music in musicians. Results of the informational connectivity analysis showed that strong connectivity between the right IFG and left pallidum was positively correlated with music classification performance in musicians. Given that the sensorimotor territory of the globus pallidus internus known to be the main output of basal ganglia, and that basal ganglia plays an important role in the storage and expression of learned sequential skills (Hikosaka et al., 2002; Doyan et al., 2009), the current finding of pallidum activation and its connection with right IFG is especially interesting. Furthermore, both globus pallidus and cerebellum are the most effective sites for deep brain stimulation (DBS) in reducing motor impairments (Tewari, Fremont, & Khodakhah, 2017) and a recent study suggested that the basal ganglia and the cerebellum are interconnected at the subcortical level (Bostan & Strick, 2019). Therefore, our findings suggest that this cortico-subcortical network facilitates the perception of musical sequences, especially for the musicians, given their intensive training in music performance. The left cerebellum was also found to be significantly more engaged in musicians compared to non-musicians in the processing of impressionist music. Among non-musicians, connectivity between the cerebellum and bilateral STG was negatively correlated with classification performance. A previous study has suggested that experience-dependent changes in cerebellum could contribute to motor sequence learning (Doyan et al., 2002), given that the motor network is important for production and perception of music (Schubotz et al. 2000), our results for the musicians suggest that the engagement of cerebellum facilitates motor sequence and musical sequence perception in turn. Further studies are needed to clarify the role of cerebellum-STG connectivity in music processing among non-musicians. A cortico-subcortical network involving the putamen, SMA, and PMC has been proposed to be engaged in the analysis of temporal sequences and in auditory-motor interactions (Grahn & Rowe, 2009). The present study verified the engagement of these proposed regions, and in addition allowed us to have a more refined understanding of the functions of different regions. Furthermore, this cortical-subcortical connectivity is shown to be functionally correlated with behavioral performance in music genre classification and neural musical syntax processing among musicians.

4.3

Appreciation of tonality in music from a scientific perspective Western classical (tonal) music has been widely appreciated due to its consonance and stability. In the present study, musicians showed stronger and more widespread neural responses to classical music compared to non-musicians. Non-musicians, though with relatively less activation than musicians, still showed stronger neural responses to classical music than to

impressionist or atonal music. The higher accuracy in classifying classical musical phrases
 457 among non-musicians can be seen as evidence of implicit knowledge of musical structure
 even 458 among those with minimal musical expertise. Furthermore, as described by Tonal
 Pitch Space 459 (TPS) theory (Lerdahl, 1988), tension and relaxation of chords unfolding
 over time in classical 460 music provide listeners with a musical context in which to generate
 reliable expectations. available under aCC-BY-NC-ND 4.0 International license. (which was
 not certified by peer review) is the author/funder, who has granted bioRxiv a license to
 display the preprint in perpetuity. It is made bioRxiv preprint doi: [https://doi.org/](https://doi.org/10.1101/2022.10.14.512259)
 10.1101/2022.10.14.512259; this version posted October 17, 2022. The copyright holder for
 this preprint Tonality decides how much we can appreciate the music. 16461 Impressionist
 music, on the other hand, is well-known for feelings of ambiguity and intangibility, 462 like
 impressionist paintings. This music genre places the listener in a reduced tonality context,
 463 which causes difficulty in integrating harmonics. In our study, although impressionist
 music and 464 classical music both engaged similar fronto-temporal regions, they each
 involved specific 465 regions as well. Furthermore, the differences between musicians and
 non-musicians in both 466 behavioral and neural responses suggest that the processing of
 impressionist music especially 467 involved frontal regions of the right hemisphere, and that
 impressionist music processing 468 benefited from musicianship more so than classical music
 processing. 469 Lastly, the atonal genre stands opposite to tonality. Its disordered structure
 and unexpected 470 musical context may well be perceived as scrambled pieces, resulting in
 poor performance in 471 differentiating atonal phrases from random notes, and in a lack of
 significant differences in 472 neural responses between atonal phrases and random notes,
 regardless of the level of music 473 experience. There are only a few studies on tonality in
 neuroscience. Among them, Proverbio et 474 al. (2015) suggested that atonal music decreased
 non-musicians' heart rates and increased their 475 blood pressure, possibly reflecting an
 increase in alertness and attention, and thus appeared to be 476 perceived as being more
 agitating and less joyful than tonal music. The present study provides 477 complementary
 results regarding the absence of "syntactic" processing in atonal music 478 perception, and
 questions the "meaning" of atonal music. 479 Overall, by studying varying music genres and
 corresponding aesthetic experiences, findings in 480 the present study allow us to gain a better
 understanding of neural mechanisms underlying 481 musical syntax processing, namely how
 it varies across levels of tonality, and how it is 482 modulated (or not) by music experience,
 and also lend strong support to music theory. 483 5 Funding 484 This work was supported by
 Art Project of National Social Science Foundation of China (No. 485 16BD050), National
 Natural Science Foundation of China (No. 31771210), National Social 486 Science Major
 Project of China (No. 17ZDA323), and Science and Technology Commission of 487
 Shanghai Municipality (No. 19JC1410101). 488 6 Acknowledgments 489 We appreciate
 Shuai Wang's suggestions and help on informational connectivity analysis. 490 7 Conflict of
 Interest 491 The authors declare that the research was conducted in the absence of any
 commercial or 492 financial relationships that could be construed as a potential conflict of
 interest. 493 8 Data Availability Statement 494 The raw datasets for this study can be found
 in the OSF repository <https://osf.io/4fejw/>. available under aCC-BY-NC-ND 4.0 International
 license. (which was not certified by peer review) is the author/funder, who has granted
 bioRxiv a license to display the preprint in perpetuity. It is made bioRxiv preprint doi: [https://](https://doi.org/10.1101/2022.10.14.512259)
 doi.org/10.1101/2022.10.14.512259; this version posted October 17, 2022. The copyright
 holder for this preprint Tonality decides how much we can appreciate the music. 17495
 References 496 Bianco R, Novembre G, Keller PE, Kim S-G, Scharf F, Friederici AD,
 Villringer A, Sammler D. 497 2016. Neural networks for harmonic structure in music
 perception and action. *NeuroImage*. 142: 498 454–464. 499 Bigand E, Poulin-Charronnat B.
 2006. Are we "experienced listeners"? A review of the musical 500 capacities that do not

depend on formal musical training. *Cognition*. 100(1): 100–130. 501 Bostan AC, Strick PL. 2018. The basal ganglia and the cerebellum: nodes in an integrated network. *Nature reviews. Neuroscience*. 19(6): 338–350. 503 Brattico E, Bogert B, Jacobsen T. 2013. Toward a neural chronometry for the aesthetic experience of music. *Frontiers in Psychology*. 4: 206. 505 Cheung VKM, Meyer L, Friederici AD, Koelsch S. 2018. The right inferior frontal gyrus processes nested non-local dependencies in music. *Scientific Reports*. 8(1): 1–12. 507 Coutanche MN, Thompson-Schill SL. 2013. Informational connectivity: identifying synchronized discriminability of multi-voxel patterns across the brain. *Frontiers in Human Neuroscience*. 7: 15. 510 Daynes H. 2010. Listeners' perceptual and emotional responses to tonal and atonal music. *Psychology of Music*. 39(4): 468–502. 512 Day-O'Connell, J. 2009. Debussy, Pentatonicism, and the Tonal Tradition. *Music Theory Spectrum*. 31(2): 225–261. 514 Doyon J, Song AW, Karni A, Lalonde F, Adams MM, Ungerleider LG. 2002. Experience dependent changes in cerebellar contributions to motor sequence learning. *Proceedings of the National Academy of Sciences*. 99: 1017–1022. 517 Doyon J, Bellec P, Amsel R, Penhune V, Monchi O, Carrier J, Lehericy S, Benali H. 2009. Contributions of the basal ganglia and functionally related structures to motor learning. *Behavior and Brain Research*. 199: 61–75. 520 Farbood MM, Heeger DJ, Marcus G, Hasson U, Lerner Y. 2015. The neural processing of hierarchical structure in music and speech at different timescales. *Frontiers in Neuroscience*. 9: 522 157. 523 Fedorenko E, Patel A, Casasanto D, Winawer J, Gibson E. 2009. Structural integration in language and music: Evidence for a shared system. *Memory & Cognition*. 37(1): 1–9. 525 Feng L, Wang X. 2017. Harmonic template neurons in primate auditory cortex underlying complex sound processing. *Proceedings of the National Academy of Sciences*. 114(5): E840–E848. available under aCC-BY-NC-ND 4.0 International license. (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made bioRxiv preprint doi: <https://doi.org/10.1101/2022.10.14.512259>; this version posted October 17, 2022. The copyright holder for this preprint Tonality decides how much we can appreciate the music. 18528 Fitch WT, Martins MD. 2014. Hierarchical processing in music, language, and action: Lashley revisited. *Annals of the New York Academy of Sciences*. 1316(1): 87–104. 530 Frisch S, Kotz SA, von Cramon DY, Friederici AD. 2003. Why the P600 is not just a P300: the role of the basal ganglia. *Clinical Neurophysiology*. 114(2): 336–340. 532 Grahn JA, Rowe JB. 2009. Feeling the Beat: Premotor and Striatal Interactions in Musicians and Nonmusicians during Beat Perception. *Journal of Neuroscience*. 29(23): 7540–7548. 534 Hikosaka O, Nakamura K, Sakai K, Nakahara H. 2002. Central mechanisms of motor skill learning. *Current Opinion in Neurobiology*. 12: 217–222. 536 James CE, Oechslin MS, Van De Ville D, Hauert C-A, Descoux C, Lazeyras F. 2013. Musical training intensity yields opposite effects on grey matter density in cognitive versus sensorimotor networks. *Brain Structure and Function*. 219(1): 353–366. 539 Koelsch S, Rohrmeier M, Torrecuso R, Jentschke S. 2013. Processing of hierarchical syntactic structure in music. *Proceedings of the National Academy of Sciences*. 110(38): 15443–15448. 541 Koelsch S, Gunter TC, v. Cramon DY, Zysset S, Lohmann G, Friederici AD. 2002a. Bach Speaks: A Cortical “Language-Network” Serves the Processing of Music. *NeuroImage*. 17(2): 956–966. 544 Koelsch S, Gunter T, Friederici AD, Schröger E. 2000. Brain Indices of Music Processing: “Nonmusicians” are Musical. *Journal of Cognitive Neuroscience*. 12(3): 520–541. 546 Koelsch S, Schmidt B, Kansok J. 2002b. Effects of musical expertise on the early right anterior negativity: An event-related brain potential study. *Psychophysiology*. 39(5): 657–663. 548 Koelsch S, Siebel WA. 2005. Towards a neural basis of music perception. *Trends in Cognitive Sciences*. 9(12): 578–584. 550 Krumhansl CL, Sandell GJ, Sergeant DC. 1987. The perception of tone hierarchies and mirror forms in twelve-tone serial

music. *Music Perception*. 5(1): 31–78. 552 Kunert R, Willems RM, Casasanto D, Patel AD, Hagoort P. 2015. Music and Language Syntax⁵⁵³ Interact in Broca's Area: An fMRI Study. *PLOS ONE*. 10(11): e0141069. 554 Kung S-J, Chen JL, Zatorre RJ, Penhune VB. 2013. Interacting Cortical and Basal Ganglia⁵⁵⁵ Networks Underlying Finding and Tapping to the Musical Beat. *Journal of Cognitive⁵⁵⁶ Neuroscience*. 25(3): 401–420. 557 Lerdahl F, Jackendoff R. 1983. *A Generative Theory of Tonal Music*. Cambridge, MA: MIT⁵⁵⁸ Press. available under aCC-BY-NC-ND 4.0 International license. (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made bioRxiv preprint doi: <https://doi.org/10.1101/2022.10.14.512259>; this version posted October 17, 2022. The copyright holder for this preprint Tonicity decides how much we can appreciate the music. 19559 Lerdahl F. 1988. Tonal Pitch Space. *Music Perception: An Interdisciplinary Journal*. 5(3): 315–560 349. 561 Levitin DJ, Menon V. 2003. Musical structure is processed in “language” areas of the brain: a⁵⁶² possible role for Brodmann Area 47 in temporal coherence. *NeuroImage*. 20(4): 2142–2152. 563 Maess B, Koelsch S, Gunter TC, Friederici AD. 2001. Musical syntax is processed in Broca's⁵⁶⁴ area: an MEG study. *Nature Neuroscience*. 4(5): 540–545. 565 Musso M, Weiller C, Horn A, Glauche V, Umarova R, Hennig J, Schneider A, Rijntjes M. 2015. ⁵⁶⁶ A single dual-stream framework for syntactic computations in music and ⁵⁶⁷ language. *NeuroImage*. 117: 267–283. 568 Oechslin MS, Van De Ville D, Lazeyras F, Hauert C-A, James CE. 2013. Degree of musical⁵⁶⁹ expertise modulates higher order brain functioning. *Cerebral Cortex*. 23(9): 2213–2224. 570 Patel AD. 2003. Language, music, syntax and the brain. *Nature Neuroscience*. 6(7): 674–681. 571 Patel AD, Iversen JR, Wassenaar M, Hagoort P. 2008. Musical syntactic processing in ⁵⁷² agrammatic Broca's aphasia. *Aphasiology*. 22(7-8): 776–789. 573 Proverbio AM, Manfrin L, Arcari LA, De Benedetto F, Gazzola M, Guardamagna M, Lozano⁵⁷⁴ NV, Zani A. 2015. Non-expert listeners show decreased heart rate and increased blood pressure⁵⁷⁵ (fear bradycardia) in response to atonal music. *Frontiers in Psychology*. 6. 576 Sammler D, Koelsch S, Friederici AD. 2011. Are left fronto-temporal brain areas a prerequisite⁵⁷⁷ for normal music-syntactic processing? *Cortex*. 47(6): 659–673. 578 Sammler D, Novembre G, Koelsch S, Keller PE. 2013. Syntax in a pianist's hand: ERP⁵⁷⁹ signatures of “embodied” syntax processing in music. *Cortex*. 49(5): 1325–1339. 580 Sato K, Kirino E, Tanaka S. 2015. A Voxel-Based Morphometry Study of the Brain of ⁵⁸¹ University Students Majoring in Music and Nonmusic Disciplines. *Behavioural Neurology*, 2015: 582 10p. 583 Schmahmann JD, Doyon J, McDonald D, Holmes C, Lavoie K, Hurwitz AS, 584 Kabani N, Toga A, Evans A, Petrides M. 1999. Three-Dimensional MRI Atlas of the Human⁵⁸⁵ Cerebellum in Proportional Stereotaxic Space. *NeuroImage*. 10(3): 233–260. 586 Schubotz RI, Friederici AD, von Cramon DY. 2000. Time perception and motor timing: A⁵⁸⁷ common cortical and subcortical basis revealed by event-related fMRI. *NeuroImage*. 11: 1-12. 588 Tewari A, Fremont R, Khodakhah K. 2017. It's not just the basal ganglia: cerebellum⁵⁸⁹ as a target for dystonia therapeutics. *Movement Disorders*, 32(11), 1537–1545. available under aCC-BY-NC-ND 4.0 International license. (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made bioRxiv preprint doi: <https://doi.org/10.1101/2022.10.14.512259>; this version posted October 17, 2022. The copyright holder for this preprint Tonicity decides how much we can appreciate the music. 20590 Tillmann B, Koelsch S, Escoffier N, Bigand E, Lalitte P, Friederici AD, von Cramon DY. 2006. ⁵⁹¹ Cognitive priming in sung and instrumental music: Activation of inferior frontal ⁵⁹² cortex. *NeuroImage*. 31(4): 1771–1782. 593 Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, 594 Joliot M. 2002. Automated anatomical labeling of activations in SPM using a macroscopic ⁵⁹⁵ anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*. 15: 273–289. 596 Zatorre RJ, Belin P, Penhune VB.

2002. Structure and function of auditory cortex: music and speech. Trends in Cognitive Sciences. 6(1): 37–46. 598 Table 1. List of sources for the three genres of musical materials. Genre Composer Catalogue Number of phrases Classical Bach The Well-Tempered Clavier 20 BMV1043 Brahms Hungarian Dances 20 Symphony No.4 Impressionist Debussy Estampes, Images, La Mer 20 Prélude à l'après-midi d'un faune Ravel Miroirs, Gaspard de la nuit 20 Ma mère l'Oye available under aCC-BY-NC-ND 4.0 International license. (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made bioRxiv preprint doi: <https://doi.org/10.1101/2022.10.14.512259>; this version posted October 17, 2022. The copyright holder for this preprint Tonicity decides how much we can appreciate the music. 21 Atonal Schoenberg The Book of the Hanging Gardens 20 String Quartets No. 3, Piano Suite Webern String Quartet, Variations 20 599 Table 2. Activation results of main effects of musical syntax and group, and simple effects of 600 musicianship on classical and impressionist music processing (all alphas corrected at $p < 0.001$). Regions (aal) Cluster Size Z x(mm) y(mm) z(mm) Musical syntax main effect Temporal_Sup_L 282 7.04 -51 5 -5 Temporal_Mid_L 3.95 -54 -10 -17 Hippocampus_L 74 3.58 -27 -19 -20 Frontal_Sup_Medial_L 18 4.22 -6 59 13 Frontal_Inf_Tri_L 16 4.45 -36 23 -2 Insula_L 5 4.40 -39 17 4 Frontal_Mid_L 25 4.64 -24 23 37 Postcentral_L 3.60 -54 -13 37 Supp_Motor_Area_L 571 4.00 -6 2 64 Cerebellum_6_L 4.32 -30 -67 -23 Frontal_Mid_R 34 4.10 30 41 43 Frontal_Inf_Tri_R 4.58 51 32 19 Frontal_Sup_R 186 3.80 27 -7 61 Postcentral_R 377 4.05 54 -19 37 available under aCC-BY-NC-ND 4.0 International license. (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made bioRxiv preprint doi: <https://doi.org/10.1101/2022.10.14.512259>; this version posted October 17, 2022. The copyright holder for this preprint Tonicity decides how much we can appreciate the music. 22 Pallidum_R 4 3.91 18 8 4 Cerebellum_Crus1_R 8 3.92 27 -85 -29 Group main effect (musician > non-musician) Cerebellum_Crus2_L 648 Inf -12 -82 -32 Precentral_L 43 Inf -21 -16 70 Parietal_Sup_L 37 6.41 -21 -67 40 Temporal_Mid_L 9 4.03 -63 -43 10 Temporal_Inf_L 34 7.51 -39 -43 -11 Frontal_Inf_Orb_L 5 4.74 -33 23 -11 Temporal_Inf_R 13 5.68 57 -46 -11 Frontal_Sup_R 6 7.20 15 47 22 Frontal_Inf_Orb_R 7 6.56 24 14 -11 Postcentral_R 8 6.20 48 -19 58 Cerebellum_6_R 12 5.80 24 -52 -26 Temporal_Pole_Sup_R 6 4.97 42 11 -20 Classical syntax: musician > non-musician Supp_Motor_Area_L 36 3.77 -6 17 46 Supp_Motor_Area_R 3.75 6 17 46 Frontal_Inf_Tri_R 8 3.41 45 20 4 Temporal_Sup_R 6 3.4 66 -22 4 Frontal_Sup_Medial_R 3.32 3 26 52 available under aCC-BY-NC-ND 4.0 International license. (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made bioRxiv preprint doi: <https://doi.org/10.1101/2022.10.14.512259>; this version posted October 17, 2022. The copyright holder for this preprint Tonicity decides how much we can appreciate the music. 23 Classical syntax: non-musician > musician Cingulum_Ant_L 29 4.42 -3 32 1 Cingulum_Ant_R 3.71 0 26 -5 Impressionist syntax: musician > non-musician Vermis_9 13 4.52 0 -58 -32 Impressionist syntax: non-musician > musician Hippocampus_L 30 4.82 -30 -19 -20 Temporal_Mid_L 88 4.27 -51 -67 19 Frontal_Sup_L 17 3.57 -21 38 40 Postcentral_L 77 3.92 -57 -10 34 Frontal_Mid_R 29 3.88 27 29 34 Postcentral_R 76 4.33 54 -19 34 available under aCC-BY-NC-ND 4.0 International license. (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made bioRxiv preprint doi: <https://doi.org/10.1101/2022.10.14.512259>; this version posted October 17, 2022. The copyright holder for this preprint

Syntax in Action Has Priority over Movement Selection in Piano Playing: An ERP Study
Roberta Bianco¹, Giacomo Novembre², Peter E. Keller², Florian Scharf¹, Angela D.

Friederici¹, Arno Villringer¹, and Daniela Sammler¹ Abstract ■ Complex human behavior is hierarchically organized. Whether or not syntax plays a role in this organization is currently under debate. The present ERP study uses piano performance to isolate syntactic operations in action planning and to demonstrate their priority over nonsyntactic levels of movement selection. Expert pianists were asked to execute chord progressions on a mute keyboard by copying the posture of a performing model hand shown in sequences of photos. We manipulated the final chord of each sequence in terms of Syntax (congruent/incongruent keys) and Manner (conventional/ unconventional fingering), as well as the strength of its predictability by varying the length of the Context (five-chord/twochord progressions). The production of syntactically incongruent compared to congruent chords showed a response delay that was larger in the long compared to the short context. This behavioral effect was accompanied by a centroparietal negativity in the long but not in the short context, suggesting that a syntax-based motor plan was prepared ahead. Conversely, the execution of the unconventional manner was not delayed as a function of Context and elicited an opposite electrophysiological pattern (a posterior positivity). The current data support the hypothesis that motor plans operate at the level of musical syntax and are incrementally translated to lower levels of movement selection. ■ INTRODUCTION To facilitate everyday interactions and communication, the brain constantly screens the environment for regularities, forms predictions about upcoming events, and accordingly “pre-engages” potentially relevant neural or cognitive processes (Tenenbaum, Kemp, Griffiths, & Goodman, 2011; Bubic, von Cramon, & Schubotz, 2010; Wilson & Knoblich, 2005). This ability, which does not require deliberate effort or awareness, might be a general function shared by different cognitive domains and pivotal for survival (Perruchet & Pacton, 2006). Understanding how simple elements are planned and perceived in temporally ordered and coherently structured sequences constitutes a central question in comparative studies across music, language, and action domains (Fitch & Martins, 2014; Tillmann, 2012). The specifically human ability of the brain to variably combine discrete meaningful units into rule-based hierarchical sequences is what is referred to as “syntactic processing” and has been defined as core aspect of language and communication (Friederici, 2011; Hauser, Chomsky, & Fitch, 2002; Lashley, 1951). Over the past years, similarities in the syntactic organization of language and Western music have been increasingly demonstrated (Rohrmeier & Koelsch, 2012; Katz & Pesetsky, 2011; Koelsch, 2005; Patel, 2003). Experimental studies have shown similar neural correlates for syntactic operations in language and music perception (Patel, Gibson, Ratner, Besson, & Holcomb, 1998; Sammler, Koelsch, & Friederici, 2011; Fedorenko, Patel, Casasanto, Winawer, & Gibson, 2009; Koelsch, Gunter, Wittfoth, & Sammler, 2005; Maess, Koelsch, Gunter, & Friederici, 2001) and also in the processing of complex action (Clerget, Winderickx, Fadiga, & Olivier, 2009; Fazio et al., 2009) inviting the hypothesis that syntactic processing might be a general “supramodal” key capability of the human brain (Fadiga, Craighero, & D’Ausilio, 2009; Slevc, Rosenberg, & Patel, 2009; Patel, 2003). Although analogies with the domain of action, in terms of hierarchical and combinatorial organization (Pulvermüller, 2014; Guerra-Filho & Aloimonos, 2012; Pastra & Aloimonos, 2012; Pulvermüller & Fadiga, 2010) remain conceptually controversial, they might be empirically tenable if shifted from actions to action planning (Moro, 2014a, 2014b). Therefore, in this study, we aimed to explore syntax-related mechanisms operating during action motor planning. Piano performance in the Western classical music tradition provides an ideal test bed for exploring syntax in the action domain. First, playing chord progressions from this tradition is the direct motoric translation of musical syntax, a theoretically established hierarchical system of rules governing music structure (Rohrmeier, 2011). Second, it affords the possibility to investigate different hierarchical

¹ Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, ² Western Sydney University, Sydney,

Australia © 2015 Massachusetts Institute of Technology *Journal of Cognitive Neuroscience* 28:1, pp. 41–54 doi:10.1162/jocn_a_00873 stages in action planning (Keller, 2012; Uithol, van Rooij, Bekkering, & Haselager, 2012; Haggard, 2008; Shaffer, 1981) from lower nonsyntactic levels of movement selection to higher levels of syntax-based action plans. Sammler, Novembre, et al. (2013) and Novembre and Keller (2011) showed that expert pianists—due to intense practice—have motorically learned syntactic regularities governing musical sequences and therefore generate motor predictions based on their acquired long-term syntactic knowledge. In a priming paradigm, expert pianists were asked to imitate silent videos of a handplaying chord sequences. The last chord was either syntactically congruent or incongruent with the preceding musical context. Despite the absence of musical sounds, both studies revealed slower imitation times for syntactically incongruent chords as well as motor facilitation (i.e., faster responses) for the syntactically congruent chords. In terms of ERPs (Sammler, Novembre, Koelsch, & Keller, 2013), the imitation of the incongruent chords elicited an early negativity, which was associated with the perception of the syntactic violation (Koelsch, 2009), and a later posterior negativity, indexing the reprogramming (Leuthold & Jentsch, 2002) of an anticipated motor act (i.e., the congruent chord) primed by the syntactic structure of the musical sequence. In line with models of incremental planning of serial actions (Palmer & Pfordresher, 2003), the authors argued that, during imitation of musical sequences, motor predictions of trained musicians are strongly based on long-term music-syntactic knowledge, as abstract structuring principles translate into a “grammar of musical action.” However, piano performance not only requires the planning of which chord to play according to the preceding music-syntactic context but also the selection of a specific fingering for an optimal and smooth execution of the musical sequence. Notably, through intensive musical training, frequently occurring musical patterns (i.e., scales, chord progressions) are associated with conventional fingering configurations that are automatically activated during execution of these patterns (Gellrich & Parncutt, 1998; Sloboda, Clarke, Parncutt, & Raekallio, 1998; Clarke, Parncutt, Raekallio, & Sloboda, 1997). From this perspective, it may be suggested that motor pattern familiarity, beyond syntactic knowledge, has a role in motor predictions when playing common chord progressions. This assumption finds support in the facilitated imitation of overlearned (Koeeneke, Lutz, Herwig, Ziemann, & Jäncke, 2006; Hund-Georgiadis & von Cramon, 1999) and complex actions that belong to one’s motor repertoire (Aglioti, Cesari, Romani, & Urgesi, 2008; Calvo-Merino, Glaser, Grèzes, Passingham, & Haggard, 2005). To what extent action planning operates at the level of musical syntax or at the level of common transitions of fingering configurations (i.e., the manner) is addressed here. In the present ERP study, we aimed at untangling two hierarchical stages of musical action planning related to (i) selecting a syntax-based motor program (relative to the musical goal) versus (ii) setting the parameters of this program (the specific movement selection defining the manner of execution). Evidence for the dissociation between program selection and parameter setting has been gleaned from theoretical and empirical work. For example, the framework of “generalized motor programs” (Rosenbaum, Kenny, & Derr, 1983; Keele & Summers, 1976; Schmidt, 1975; Lashley, 1951) posits that action plans consist of core motor programs whose specific movement parameters are only chosen at the time of their use. Furthermore, it has been shown that the performance of advanced pianists is based on abstract conceptual plans and is independent of the specific movement requirements (Palmer & Meyer, 2000). Similarly, a dissociation between a more general and higher versus a more specific and lower level of action processing finds support in the “hierarchical organization of goal-directed actions” theory (see Grafton, 2009). Along these lines, the priority of the goal of an action over the means used to achieve it has been extensively demonstrated in behavioral imitation (Wohlschläger, Gattis, & Bekkering, 2003; Bekkering, Wohlschläger, & Gattis, 2000), neuroimaging (Hamilton & Grafton, 2006; Chaminade,

Meltzoff, & Decety, 2002; Koski et al., 2002), and brain stimulation studies (Lago & Fernandez-del-Olmo, 2011; Cattaneo, Caruana, Jezzini, & Rizzolatti, 2009). Therefore, we transferred this hierarchical concept of action planning to music by focusing on predictions at the two levels of the motor hierarchy (goal and manner; see also Novembre & Keller, 2011). We reasoned that motor predictions of expert pianists should concern the musical goal (Syntax) rather than the finger movement selection (Manner), which should be specified only at the time of execution. We asked expert pianists to watch and execute as fast and accurately as possible chord sequences played by a performing pianist's hand presented in a series of photos on a computer screen. Moreover, to negate exogenously driven auditory predictive processes, no sound was used. Piano performance (RTs and errors) and an EEG were recorded. In a 2×2 factorial design, we manipulated the last chord of the sequences in terms of the identity of the target keys (Syntax congruent/incongruent) to address the syntactic level of action planning and in terms of fingering (Manner correct/incorrect) to address the level of movement selection. To induce different strengths of syntax/manner-based predictions, pianists were presented with five-chord or two-chord sequences (long/short Context). The execution of the long compared to the short context was expected to provide more information, hence lead to a stronger prediction of the last chord to be executed. Crucially, the manipulation of the manner, while keeping the syntax congruent and vice versa, allowed us to dissociate behavioral and neural patterns elicited by the execution of the syntactic violation (Syntax) from those triggered by a general violation of movement patterns (Manner). Additionally, the 2×2 factorial design permitted us to investigate syntax-related mechanisms on top of the concurrent manner violation to test whether, in musical action planning, high levels of syntactic operations are prioritized over movement parameter specification. First, we hypothesized that motor predictions in expert pianists are driven by music-syntactic knowledge more than by motor pattern familiarity. This should be reflected in a stronger priming effect of the long context on the musical goal (Syntax) than and irrespective of the specific movement selection (Manner). Specifically, we expected the execution of the syntactically congruent/ incongruent chords to be facilitated/impeded more strongly in the long than in the short contexts, whereas no such effect should occur during the execution of the manner correct/incorrect chords. Second, in terms of neural correlates, we predicted specific response-related patterns evoked by the syntax violation, different from those associated with the processing of the manner violation. To this end, we specifically focused on the syntax-related early and late negativity described by Sammler, Novembre, et al. (2013) and manner-related effects in the same time windows.

METHODS Participants Twenty-six pianists (16 women) aged 20–33 years (mean = 25.15; SD = 3.55) were included in the analysis. Eight more pianists were tested but excluded because of an insufficient number of valid trials (cutoff = 50% of valid trials). The included pianists possessed between 12 and 27 years of classical music training (mean years of training = 18.21; SD = 3.92) and had started to play the piano on average at 6.04 years (SD = 2.73). All participants were naive with regard to the purpose of the study. They gave written informed consent to take part in this experiment and received monetary compensation for participation. The study was approved by the local ethics committee. Stimuli Stimuli were photos showing a male pianist's right hand playing sequences of chords on the piano (Yamaha Clavinova CLP150; Yamaha Music Europe GmbH, Rellingen, Germany). To maximally address action planning processes, we used photos rather than videos (cf. Sammler, Novembre, et al., 2013; Novembre & Keller, 2011) obtaining more precise onset times of target chord presentation. In two sessions, we presented a total of 72 sequences that were all different from each other in terms of melodic contour. All sequences were composed according to the rules of classical harmony. The first chord always represented the tonic. The second chord could be tonic, mediant, or

subdominant. Chords at the third position were subdominant, dominant, or dominant six–four chords, and chords at the fourth position were dominant seventh chords. At the last position, the target chord of each sequence was manipulated in terms of Syntax (syn) and Manner (man) in a 2×2 factorial design. Thirty-six sequences were conventional in terms of both Syntax and Manner (syn congruent/man correct), 12 were violated in terms of Syntax (syn incongruent/man correct), 12 in terms of Manner (syn congruent/man incorrect), and 12 in terms of both factors (syn incongruent/man incorrect). More precisely, syn congruent/man correct (ScMc) sequences ended with a tonic (a chord typically used to resolve a musical sequence) played with conventional fingering; syn incongruent/man correct (Si Mc) sequences ended with a Neapolitan chord (a minor subdominant with a diminished sixth instead of a fifth, rarely used in classical harmony to resolve a musical sequence) played with conventional fingering; the syn congruent/man incorrect (ScMi) sequences ended on a syntactically congruent tonic chord but played with an unconventional fingering; and finally the double violation (syn incongruent/man incorrect, Si Mi) was constituted by a Neapolitan chord played with an unconventional fingering. The fingering adopted by the model hand was chosen by a piano teacher with 24 years of experience conforming to the fingering taught in classical piano lessons. This was aimed to achieve smoothness and movement economy between chord transitions (i.e., 124, 125, 135, where 1 represents the thumb; 2 represents the index; and 3, 4, and 5 indicate the middle, the ring, and the little finger, respectively). Conversely, the manner manipulation of the target chord consisted of a fingering that was anatomically awkward and highly unlikely to be used (i.e., 123, 235, 245). The fingering of the last chord was rated by nine pianists on a scale from 1 (very conventional) to 9 (very unconventional). An ANOVA with the factors Syntax (congruent/incongruent) and Manner (correct/incorrect) on the mean ratings yielded a main effect of Manner [$F(1, 8) = 932.3, p < .001, \eta^2 = .991$], but neither main effect of Syntax [$F(1, 8) < 1, p = .721, \eta^2 = .017$] nor interaction of the two factors [$F(1, 8) = .206, p = .662; \eta^2 = .025$], confirming the motor unconventionality of the chosen fingering, and the independence of the syntax from the manner manipulation. All chords consisted of three keystrokes. Four tonalities with either two or four sharps or flats in the key signatures, that is, D, E, B \flat , and A \flat major, were used with equal probability for each condition to balance the average amount of black and white keys in syntactically congruent and incongruent chords and to thus equate their visual surface structure and difficulty of execution. Red circles were superimposed on top of each pressed key for the whole duration of the photo to facilitate the recognition of the pressed keys. Sequences of two different lengths were created: five-chord sequences (long context) and two-chord sequences (short context). The two-chord sequences were identical to the last two chords of the five-chord sequences; thus, the long and short sequences differed only in the strength of the predictability of the last chord. In addition, comparing Bianco et al. 43 the four conditions across the two contexts allowed us to control for motor differences between conditions, that is, naturally longer hand trajectories from the penultimate to the syntactically incongruent (Si Mc; due to different keys) and the manner incorrect (ScMi, Si Mi; due to hand rotation) chords than to the not manipulated chords (ScMc). Finally, it should be noted that the sequence-final tonic chords naturally share a number of keys with the harmonically related context. To control for the possibility that final tonics may be merely motorically primed by the repeated use of these keys, we also allowed the final correct manner to be partly motorically primed by presenting the respective fingering on average 1.1 times in the context. The balanced repetition of tonic key configuration (1.3 times) and final chord fingering, that is, the similar likelihood of motor priming, discloses any effect that occurs in the syntax but not manner conditions as related to syntactic (not motor) priming. Each sequence started with 2-sec presentation of a preparatory photo showing a stationary hand poised to press the three keys associated with the first chord. Then the following photos were

presented at a rate of 2 sec per photo (total duration: 12-sec long sequences, 6-sec short sequences). Procedure Participants were seated in front of a MIDI (musical instrument digital interface) piano (Yamaha Clavinova CLP150) and watched the photo sequences on a computer monitor (100-Hz refresh rate). Simultaneously, they were required to execute the chords they saw, one by one, with their right hand on the piano, as quickly and accurately as possible, both in terms of the keys (syntax) and in terms of fingering (manner). Note that the piano was muted and no sound was presented with the photos, that is, the experiment took place in total absence of musical sounds. Each trial started with a visual fixation cross of 0.5-sec duration and ended with a black screen for 1.5 sec after the final photo of the stimulus sequence (Figure 1). Participants were invited for two sessions in which they were presented with the same stimuli. Each session consisted of six experimental blocks, one for each of the three violation conditions (SiMc, ScMi, SiMi) and separately for long and short sequences. Each block contained a total of 48 trials: 24 nonviolated trials (ScMc) intermixed with 24 trials of the respective violation condition. Block order was counterbalanced across participants and alternated between blocks with long and short sequences. To acquaint participants with the unusual and challenging task and increase accuracy, in the first session, they all received a training of 24 trials (50% nonviolated) for all the violation conditions in both long and short context. Different tonalities were used for the training (G, F, D \flat , and B major) than in the main experiment. At the end of the experiment, participants filled out a questionnaire to assess how much they relied on auditory imagery, motor imagery, and/or theoretical knowledge of western harmony to do the task. Their piano expertise was estimated as the sum of training hours per day across all years of piano lessons. Stimulus presentation and response registration were controlled by Presentation software (Version 14.9, Neurobehavioral System, Inc., Berkeley, CA). Through a (custombuilt) MIDI interface, the MIDI piano key values were converted into a serial signal compatible with Presentation software. This allowed us to compute the RTs of the keystrokes in relation to the onset of the target chord photo. Moreover, a video camera placed above the keyboard recorded the pianist's hand from an aerial view, allowing us to detect (offline) trials in which the pianists used a different fingering from that shown in the photos. EEG Data Acquisition The EEG recordings were acquired from 61 Ag/AgCl electrodes (Fpz, Fp1, Fp2, AFz, AF3, AF4, AF7, AF8, Fz, F3, F4, F5, F6, F7, F8, F1, FC2, FCz, FC3, FC4, FC5, FC6, FT7, FT8, FC1, F2, Cz, C1, C2, C3, C4, C5, C6, T7, T8, CPz, CP3, CP4, CP5, CP6, TP7, TP8, P1, P2, Pz, P3, P4, P5, P6, P7, P8, CP1, CP2, POz, PO3, PO4, PO7, PO8, O1, O2, Oz) according to the international 10–20 system (Sharbrough et al., 1991). The left mastoid (M1) served as reference. Three additional electrodes were placed on the sternum as common ground, on the right mastoid bone (M2), and on the tip of the nose for offline re-referencing. The EOG was recorded by two bipolar montages, one with electrodes located above and below the left eye and the other with two electrodes placed on the outer canthus of each eye. Signals were amplified using a 24-bit Brainvision QuickAmp 72 amplifier (Brain Products GmbH, Gilching, Germany) with input impedance below 5 k Ω and digitized at a 500-Hz sampling rate. Figure 1. Experimental design. In the total absence of musical sound, pianists executed chord progressions with their right hand by copying (as fast and accurately as possible) the posture of a performing model hand shown in sequences of photos. The target chord of each progression was manipulated in terms of keys (congruent/ incongruent Syntax) and fingering (conventional/unconventional Manner) in a 2 \times 2 factorial design and was presented at the end of a five- or two-chord sequence (long/short Context) to induce different strengths of predictability. 44 Journal of Cognitive Neuroscience Volume 28, Number 1 Behavioral Data Analysis RTs and execution errors of the last chord of each trial were analyzed in accordance with Novembre and Keller (2011). The minimum requirement for including a participant's data into the analysis was correct responses on 50% of trials. Trials were considered valid

when three conditions were satisfied: (1) both the last and the second last chord had to be correctly imitated in terms of keys and fingering (for the error analysis we included the trials correctly imitated in the second last but incorrectly in the last chord), (2) the keystrokes within a chord had to be synchronous (i.e., no more than 150 msec should intervene between the first and the last of the three keystrokes), and (3) mean RTs of the three keystrokes after the onset of the target chord photo had to stay within 3000 msec (cf. Drost, Rieger, Brass, Gunter, & Prinz, 2005). RTs were calculated by subtracting the time of execution of the target chord (i.e., mean of the times of three keystrokes composing the chord) from the onset time of the last photo showing the target chord. The fingering performed by each pianist was analyzed through offline inspection of the video recordings in which the fingers employed by the participants were compared with the fingers presented in the stimulus photos. Statistical evaluation of the RT data was done using three-way repeated-measures ANOVA with the factors Syntax (congruent/incongruent), Manner (correct/ incorrect), and Context (long/ short). Errors were analyzed with an analogous ANOVA, but with the additional withinsubject factor Error type (key/fingering error). EEG Data Analysis Data analysis was carried out using EEGLAB toolbox 9.01 (Delorme & Makeig, 2004) implemented in MATLAB 7.7. The EEG data were offline re-referenced to the algebraic mean of the mastoids and were 0.3-Hz high-pass filtered (fir, 5854 points, Blackman window). Strong muscle artifacts, electrode drifts, or technical artifacts were manually rejected. Independent component analysis was used for linear decomposition of the continuous data to remove the contributions of artifact sources (slow drifts, eye blink/movement, and muscle artifacts) on the scalp sensors. After 25-Hz low-pass filtering (fir, 110 points, Blackman window), epochs of -200 to 1500 msec, time-locked to the onset of the photo of the last chord, were extracted from the data. Epochs were rejected whenever signal voltages exceeded $\pm 80 \mu\text{V}$ in one or more electrodes. Nonrejected trials were averaged separately for each condition and baseline-corrected (-200 msec before the onset of the target photo). Only correct trials according to the behavioral analysis were included in the ERP statistical analysis (i.e., mean number of trials \pm SD for ScMc, SiMc, ScMi, SiMi in the long context: 122.9 ± 10 , 38.1 ± 5.4 , 37.7 ± 4.8 , 35.8 ± 5.5 ; in the short context: 131.3 ± 6 , 41.3 ± 4.4 , 40.4 ± 4.7 , 38.3 ± 6.2). Figure 2. (A) Mean RTs during imitation of syntactically incongruent (dashed line) and congruent chords (solid line; left) and during imitation of manner incorrect (dashed line) and correct chords (solid line; right) in the long and short context. (B) Number of key errors (left) and fingering errors (right) during imitation of syntactically incongruent and congruent chords in the long and short context. Error bars indicate 1 SEM. *** $p < .001$, ** $p < .01$, * $p < .05$.

Bianco et al. 45 Effects of Syntax, Manner, and Context were analyzed time-locked to the onset of the last (target) photo of the sequence. Mean amplitudes were computed separately for each condition over nine ROIs and for three specific time windows. The ROIs comprised (i) left anterior (F3, F5, F7, FC3, FC5, FT7, AF3), (ii) left central (C3, C5, T7, CP3, CP5, TP7), (iii) left posterior (P3, P5, P7, PO3, PO7), (iv) middle anterior (F1, FZ, F2, FC1, FCZ, FC2, AFZ), (v) middle central (C1, CZ, C2, CP1, CPZ, CP2), (vi) middle posterior (P1, PZ, P2, POZ), (vii) right anterior (F4, F6, F8, FC4, FC6, FT8, AF4), (viii) right central (C4, C6, T8, CP4, CP6, TP8), (ix) right posterior (P4, P6, P8, PO4, PO8). Three time windows (i) from 210 to 520 msec, (ii) from 520 to 800 msec, and (iii) from 800 to 1200 msec were selected by visual inspection of the ERPs and topography plots shown in Figures 2 and 3. As objective and external criteria independent from the data, borders of the time windows were set to time points at which either a change in polarity or in topography was found, assuming that different map topographies and polarities directly indicate different underlying generators, that is, different cognitive processes (Michel et al., 2004). The same time windows were used in the analysis of the two conditions. Statistical analysis of mean amplitude values was carried out by means of five-way ANOVAs with the repeated-measures

factors Syntax (congruent/incongruent) \times Manner (correct/incorrect) \times Context (long/short) \times Laterality (left/middle/right) \times AntPost (anterior/central/posterior), separately for each time window. RESULTS Behavioral Data RTs Statistical values of the $2 \times 2 \times 2$ repeated-measures ANOVA with the factors Syntax (congruent/incongruent), Manner (correct/incorrect), and Context (long/short) are reported in Table 1. These results revealed main effects of Syntax and Manner, indicating that imitation of the syntactically incongruent as well as manner violated chords was generally slower compared to chords that contained no such violations. No main effect of Context was found, showing that RTs for the imitation of target chords was comparable between long and short sequences. Notably, a highly significant Syntax \times Context interaction showed that more in the long than in the short context the execution of syntactically congruent chords was faster compared to incongruent chords. Conversely, no interaction between Manner \times Context was found (Figure 2A). This finding suggests that the harmonic structure of the musical context rather than the motor pattern familiarity drove the motoric prediction of the target chord and that the Figure 3. Effect of Syntax. ERPs evoked by syntactically incongruent (dotted line) compared to congruent (solid line) chords in the long (left) and short (right) context across all trials. The three time windows are shaded according to their polarity (red for positivity, blue for negativity). Topography maps for each statistical time window (lower row) depict the difference potentials of syntactically incongruent minus congruent chords (arrows below indicate the interaction between Syntax and Context). *** $p < .001$, ** $p < .01$, * $p < .05$. 46 Journal of Cognitive Neuroscience Volume 28, Number 1 prediction concerned the musical goal (Syntax) rather than the movement selection (Manner). A three-way interaction of Syntax \times Manner \times Context suggested a reciprocal influence of syntax and manner processing in relation to the Context. We calculated separate ANOVAs for the manner correct and incorrect trials with the factors Syntax and Context and for the syntax congruent and incongruent trials with the factors Manner and Context. This analysis yielded a Syntax \times Context interaction in both manner correct [$F(1, 25) = 71.99$, $p < .001$, $\eta^2 = .742$] and, although weaker, in the manner incorrect trials [$F(1, 25) = 5.98$, $p = .022$, $\eta^2 = .193$], whereas a Manner \times Context interaction was found only in the syntax congruent trials [$F(1, 25) = 4.505$, $p = .044$, $\eta^2 = .153$] and not in the syntax incongruent trials [$F(1, 25) = 1.649$, $p = .211$, $\eta^2 = .062$]. This indicates that while the syntactic prediction effect (Syntax \times Context interaction), although weaker, held in presence of the concurrent manner violation, the manner was facilitated (Manner \times Context) only when the pianists' syntactic prediction was fulfilled (syntax congruent trials). These data suggest that the syntax of the context primes primarily the motor program of the musical goal, which in turn may trigger information about optimal movement parameters for its execution. In other words, movement selection is facilitated only when the higher plan on the musical goal is confirmed. Error Analysis Errors in terms of pressed keys and fingering were counted separately. Trials with both error types were excluded from the analysis (as in Novembre & Keller, 2011). Key and fingering errors are assumed to reflect distinct cognitive processes associated with the musical goal (Syntax) and the specific movement used in the execution (Manner), respectively. A $2 \times 2 \times 2 \times 2$ repeated-measures ANOVA with the factors Syntax, Manner, Context, and Error type (key/ fingering errors; for statistical values, see Table 2) revealed that —overall—less errors were committed during execution of syntactically congruent compared to incongruent chords (main effect of Syntax) as well as during the imitation of manner correct compared to incorrect chords (main effect of Manner). Also, less errors were committed in the short compared to the long context (main effect of Context), whereas there was no significant difference between number of key and fingering mistakes (no significant main effect of Error type). The interactions of Syntax \times Error type and Manner \times Error type revealed that Syntax and Manner conditions were associated with greater amount of key and

fingering errors, respectively. Importantly, a Syntax \times Context \times Error type interaction indicated that key errors, but not fingering errors, were more prevalent in the long than in the short context during the execution of syntactically incongruent chords, irrespective of the manner. Indeed, follow-up ANOVAs with the factors Syntax and Context, calculated separately for the key and the fingering errors (Figure 2B), yielded a significant Syntax \times Context interaction for the key errors [$F(1, 25) = 7.164, p = .013, \eta^2 = .223$], but not for the fingering errors [$F(1, 25) = 2.599, p = .122, \eta^2 = .093$]. These data indicate that the harmonic structure of the context strongly affected the motor program of the musical goal, irrespective of the specific movement selection. Additionally, similar to what was observed in the RTs, we found that during the execution of the manner incorrect chords more fingering errors were committed in the long than in the short context but only when the syntax was congruent [Manner \times Context interaction on the fingering errors across syntactically congruent trials: $F(1, 25) = 9.120, p = .006, \eta^2 = .267$; across syntactically incongruent trials: $F(1, 25) = 1.161, p = .292, \eta^2 = .044$]. This finding confirms that the selection of which fingers to use was facilitated in the long context only when the musical goal matched the (syntactic) predictions. Table 2. Results of the ANOVA on Number of Errors with the Factors Syntax \times Manner \times Context \times Error Type

Effect	df	F	p	η^2
S	1, 25	39.896	.000	.615
M	1, 25	20.907	.000	.455
C	1, 25	5.779	.024	.188
Et	1, 25	3.276	.082	.116
S \times Et	1, 25	9.028	.006	.265
M \times Et	1, 25	32.146	.000	.563
S \times C \times Et	1, 25	8.868	.006	.262

Bold values indicate significant results ($p < .05$). Partial eta squared: $\eta^2 > .5$, large effect size; $\eta^2 > .3$, medium effect size; $\eta^2 \leq .1$, small effect size (Bortz & Döring, 2003). S = Syntax; M = Manner; C = Context; Et = Error type. Table 1. Results of the ANOVA on RTs with the Factors Syntax \times Manner \times Context

Effect	df	F	p	η^2
S	1, 25	80.16	.000	.762
M	1, 25	133.65	.000	.842
C	1, 25	1.522	.229	.057
S \times C	1, 25	52.56	.000	.678
M \times C	1, 25	.604	.445	.024
S \times M \times C	1, 25	4.78	.038	.160

Bold values indicate significant results ($p < .05$). Partial eta squared: $\eta^2 > .5$, large effect size; $\eta^2 > .3$, medium effect size; $\eta^2 \leq .1$, small effect size (Bortz & Döring, 2003). S = Syntax; M = Manner; C = Context. Bianco et al. 47 EEG Data We were interested in distinguishing neurophysiological correlates of (i) the higher level of syntax-based motor programming (prediction of the musical goal, reflected by a Syntax \times Context interaction) and (ii) the lower level of specific movement selection (prediction of a conventional optimal movement, reflected by a Manner \times Context interaction). To this end, we analyzed (Table 3) the effects of (i) Syntax (Figure 3), (ii) Manner (Figure 4), and (iii) their interaction separately in three time windows (see EEG Data Analysis). Nonparametric clusterbased permutation tests with standard parameters in Fieldtrip (www.fieldtriptoolbox.org) revealed qualitatively similar results to those described below. The analysis of the effects of Syntax and Manner revealed different neural signatures between 520 and 800 msec, in line with our hypothesis that the planning of a musical goal (Syntax) and the specific movements (Manner) rely on different mechanisms. Table 3. Results of the ANOVAs with the Factors Syntax \times Manner \times Context \times Laterality \times AntPost for Each Time Window

Effect	df	1st tw: 210...520 msec	2nd tw: 520...800 msec	3rd tw: 800...1200 msec		
Effect of Syntax	S	1, 25 5.733 .024 .187 1.019 .322 .039 <1 .352 .035	S	1, 25 1.578 .221 .059 1.673 .208 .063 <1 .650 .008	S	2, 50 <1 .417 .033 3.965 .029 .137 10.894 .000 .304
S \times L	S \times L	2, 50 1.481 .238 .056 1.082 .159 .073 <1 .475 .029	S \times A	2, 50 8.225 .007 .248 1.543 .227 .058 1.017 .369 .039	S \times A \times C	2, 50 <1 .912 .001 <1 .835 .003 <1 .557 .023
S \times A \times L \times C	S \times A \times L \times C	4, 100 2.015 .120 .075 <1 .660 .021 2.977 .051 .103	S \times A \times L \times C	4, 100 1.083 .360 .042 2.886 .035 .103 1.685 .179 .063	Effect of Manner	
M	M	1, 25 1.630 .213 .061 29.014 .000 .537 <1 .402 .028	M \times C	1, 25 <1 .487 .008 1.715 .202 .064	M \times L	2, 50 24.202 .000 .492 11.486 .000 .315 9.401 .001
M \times L \times C	M \times L \times C	2, 50 1.064 .347 .041 <1 .520 .025 1.401 .256 .053	M \times A	2, 50 9.592 .003 .277 2.004 .168 .074 <1 .425 .028	M \times A \times C	2, 50 10.279 .003 .291 7.833 .005 .239 1.230

.283 .047 M × A × L 4, 100 9.387 .000 .273 8.595 .000 .256 3.601 .024 .126 M × A × L × C 4, 100 1.165 .328 .045 1.552 .205 .058 1.558 .210 .059 Syntax and Manner Interaction S × M 1, 25 1.164 .291 .044 <1 .869 .001 <1 .964 .000 S × M × C 1, 25 <1 .584 .012 1.586 .219 .060 3.780 .063 .131 S × M × C × L 2, 50 <1 .926 .002 <1 .419 .033 3.039 .068 .108 S × M × A 2, 50 1.962 .172 .073 6.614 .012 .209 <1 .358 .035 S × M × C × A 2, 50 <1 .714 .006 <1 .594 .012 1.128 .304 .043 S × M × L × A 4, 100 <1 .672 .021 1.124 .347 .043 2.778 .039 .100 S × M × L × A × C 4, 100 1.137 .340 .044 1.633 .187 .061 1.249 .297 .048 Bold values indicate the effects due to the difference in strength of potentials. Partial eta squared: $\eta^2 > .5$, large effect size; $\eta^2 > .3$, medium effect size; $\eta^2 \leq .1$, small effect size (Bortz & Döring, 2003). S = Syntax; M = Manner; C = Context; L = Laterality; A = AntPost. 48 Journal of Cognitive Neuroscience Volume 28, Number 1 Effect of Syntax In the first time window (210–520 msec), syntactically incongruent compared to congruent chords elicited a positivity with a central scalp distribution as revealed by a significant main effect of Syntax and an interaction of Syntax × AntPost. This early effect did not differ between long and short context (i.e., no interaction of Syntax × Context), suggesting that it was not related to prediction. Most relevant, in the second time window (520–800 msec), syntactically incongruent compared to congruent chords evoked a centroparietal negativity that was present in the long and not in the short context. The four-way ANOVA showed a significant interaction of Syntax × Context × AntPost × Laterality demonstrating that the negativity was particularly strong in the long context and more enhanced in the midline central regions. In the third time window (800–1200 msec), the four-way ANOVA yielded an interaction of Syntax × Laterality, as well as a marginally significant interaction of Syntax × AntPost × Laterality, indicating a predominantly middle-central negativity that did not differ in amplitude between long and short context. To evaluate whether the negativity between 520 and 800 msec in the long context was influenced by auditory imagery, the difference wave in the middle central ROI (mean = $-0.516 \pm 1.384 \mu\text{V}$) was correlated with the subjective ratings of the extent to which participants actively imagined the sound of the up-coming chord during performance. No significant correlation was found [$r(25) = .261$, $p = .301$, $R^2 = .046$]. Effect of Manner In the first time window (210–520 msec), a significant interaction of Manner × AntPost × Laterality revealed a left middle anterior positivity elicited by the manner incorrect compared with manner correct chords across all trials. This positivity was stronger in the long compared to the short context, as shown by a significant interaction of Manner × Context × AntPost. Follow-up ANOVAs with the factors Manner × Context calculated for each ROI confirmed a left middle anterior distribution of this effect [Manner × Context interaction, middle anterior: $F(1, 25) = 7.920$, $p = .009$; left anterior: $F(1, 25) = 1.188$, $p = .027$; right anterior: $F(1, 25) = 3.793$, $p = .063$; all $ps > .116$ in the other ROIs]. In the second time window (520–800 msec), a main effect of Manner indicated that manner incorrect chords elicited more positive potentials than manner correct chords with a predominately middle to left centroparietal distribution, as confirmed by a Manner × AntPost × Laterality interaction. A Manner × Context × AntPost interaction showed that the positivity in the short context did not extend as far anteriorly as in the long context. This difference in scalp distribution was confirmed by significant Manner × Context interactions in the anterior regions as revealed by follow-up ANOVAs with the factors Figure 4. Effect of Manner. ERPs evoked by target chords played with incorrect (dotted line) compared to correct (solid line) manner in the long (left) and short (right) context across all trials. Time windows of the three time windows are shaded according to their polarity (red for positivity, blue for negativity). Topography maps for each statistical time window (lower row) depict the difference potentials of manner incorrect minus correct chords (arrows below indicate the interaction between Manner and Context). *** $p < .001$, ** $p < .01$, * $p < .05$. Bianco et al. 49 Manner × Context calculated for each ROI [middle anterior: $F(1, 25) = 7.813$, $p = .010$; right anterior:

$F(1, 25) = 10.884, p = .003$; left anterior: $F(1, 25) = 3.844, p = .061$; all $ps > .161$ in the other ROIs]. In the third time window (800–1200 msec), an interaction of Manner \times AntPost \times Laterality indicated a middle centrally distributed negativity elicited by the manner incorrect compared to the manner correct chords that tended to be larger in the short compared to the long context [Manner \times Context: $F(1, 25) = 4.002, p = .056$]. Interaction Effects Finally, we analyzed how far the effects of Syntax and Manner described above reciprocally interact. To this end, we tested for interactions that involved the factors Syntax \times Manner \times Context (and any topographical factor). No such interactions were found in any of the three time windows (see Table 3). Consequently, no further split of the general linear model was performed.

DISCUSSION Action plans are hierarchically organized, with higher levels representing the general goal of an action and lower levels concerning the specific movements required to realize the goal (Uithol et al., 2012). This study aimed to differentiate action planning based on higher-order syntactic structures (Syntax) from lower nonsyntactic processes of movement selection (Manner) in expert pianists. Therefore, behavioral and neural indices of motor prediction were examined during the execution of chords that contained either a syntax or a manner violation and that were primed by long or short musical contexts. We found (i) a strong context-dependent priming effect on the execution of syntactic violations (RTs and errors), indicating that plans of musical goals are made ahead according to the musical context. Crucially, (ii) no contextual priming was observed during the execution of manner violations, unless the syntax was congruent. In line with models of action hierarchy (Grafton & Hamilton, 2007), this suggests a priority of planning the goal of the musical action (Syntax) that in turn can prime the selection of the optimal movement parameters (Manner). Finally, (iii) different electrophysiological signals were elicited by the syntactically incongruent chords (centroparietal negativity) and manner incorrect chords (posterior positivity). These signatures may represent the different levels of action planning, pertaining to higher levels of syntax-based motor plans versus lower levels of movement parameter setting, respectively.

Behavior In line with previous findings (Sammler, Novembre, et al., 2013; Novembre & Keller, 2011), syntactically incongruent chords were executed more slowly and evoked more key mistakes than congruent chords, particularly when primed by a long musical context. Conversely, execution times and number of fingering mistakes were commensurately higher in manner incorrect than correct chords, irrespective of context length. This pattern of results not only excludes an interpretation in terms of mere motor priming (which should have led to similar context effects in syntax and manner; see Methods) but indicates that particularly the syntactic structure of the musical context narrows down the probabilities of chord transitions, thus leading the pianists to (motorically) anticipate the execution of the most likely harmonically coherent chord (Syntax). By contrast, the specific movement parameters seem to be far less strongly determined by the preceding context, despite high familiarity with the types of chord progressions employed in the paradigm. Notably, the context dependency of syntactically incongruent chords was observed irrespective of whether the concurrent manner was correct or incorrect, although it was stronger in the former due to movement familiarity. This shows that, in experts, distal goals are the main drivers of motor predictions regardless of how the goal is realized. At the same time, our data further suggest that the preplanned goal tends to prime the selection of optimal movement parameters required to achieve the goal. We ground this assumption on the observation that manner incorrect chords showed context sensitivity both in terms of execution time and number of fingering errors, but exclusively in syntactically congruent trials (i.e., when the preplanned goal was valid). This suggests that the specific movement selection constitutes a late stage of motor preparation dependent on the action plan concerning the more distal musical goal. In conclusion, motor predictions concerning the musical goal, prior to the manner, are consistent with the framework of “generalized motor

programs” (for a review, see Summers & Anson, 2009), as knowledge structures allow a given class of movements to be executed in different ways, depending on underlying parameter settings. Furthermore, it is reminiscent of imitation studies showing a hierarchical organization of action in which the action goal is prioritized over the short-term selected movements (Wohlschläger et al., 2003; Bekkering et al., 2000). ERPs The execution of the syntax and manner violations elicited different electrophysiological patterns between 520 and 800 msec: We found that the syntax violations evoked a centroparietal negativity in the long and not in the short context (similar to Sammler, Novembre, et al., 2013) whereas the execution of the manner violations elicited a positivity with left posterior scalp distribution. In line with integrated models of hierarchical organized motor plans (Grafton & Hamilton, 2007), we claim that this distinction speaks in favor of a motor program level coding for the goal 50 *Journal of Cognitive Neuroscience* Volume 28, Number 1 structure of an action (Syntax) and a lower motor level for computing the coordinated movement to a goal (Manner). More specifically, the centroparietal negativity elicited by the syntax violation was modulated by the length of the musical context and thus matched the context-dependent effects in RTs and errors. This suggests a response-related nature of the negativity, which may be interpreted as a signal of high-level movement reprogramming following the cancellation of the prepotent response in face of the incongruity to be executed (Sammler, Novembre, et al., 2013; Leuthold & Jentzsch, 2002). Importantly, this interpretation implies that the motor program for a structurally coherent musical goal was present at the moment of the target chord presentation, as it had been preplanned based on the syntactic context. Obviously, longer contexts lead to stronger syntax-based predictions of the musical goals that in turn require more effort in their revision (larger amplitude of the negativity, longer RTs, and a higher number of key errors in the long than in the short context). In a later time window (800–1200 msec), a late centrally distributed negativity, which resembles the contingent negative variation (Walter, Cooper, Aldridge, MacCallum, & Winter, 1964), was elicited in all conditions similarly in the two contexts. This time window immediately preceded and partly overlapped with the execution of the final chord and might reflect the computation of muscle-specific commands common to all conditions (Cunnington, Windischberger, Deecke, & Moser, 2003; Rektor, 2000). Indeed, the contingent negative variation is typically elicited before motor responses and specifically in the final stage of response preparation of externally cued movements (Smith, Johnstone, & Barry, 2006). Crucially, manner violations did not evoke a centroparietal negativity between 520 and 800 msec, as opposed to syntax violations. This discloses the negativity (in line with the behavioral results) as related to syntactic processes not motor priming (see Methods). Furthermore, if one accepts the idea that this negativity reflects the reprogramming of a preplanned motor response, its absence in the manner violations implies that the specific movement for execution had not been programmed at the time of the target chord presentation. This interpretation would be in line with the assumption that the musical goal is planned before movement selection. Instead, the manner violations elicited a left posterior positivity in the time window between 520 and 800 msec in both contexts, speaking in favor of a different nature of the syntax- and manner-related cognitive processes. This effect was preceded by an earlier positivity (210– 520 msec) with anterior scalp distribution. Together, these potentials resemble the P300 complex composed of P3a and P3b, typically elicited by infrequent behaviorally relevant stimuli (Gómez, Flores, Digiacomo, Ledesma, & González-Rosa, 2008) and modulated in amplitude by the probability of the deviant target (Duncan-Johnson & Donchin, 1977, 1982). Interestingly, both the early and late positivities were stronger in the long than in the short context, revealing a context-dependent effect that obviously mismatches with the behavioral data. One explanation might be that, unlike the response-related ERPs in the syntax condition, these positivities rather reflect stimulus-related

processes that are contingent on the different sequential probabilities of the manner violation in long and short stimulus sequences. More precisely, given that the manner violation can be recognized as odd even in single photos (see Methods), its occurrence probability amounts to 10% in task blocks with long sequences and 25% in task blocks with short sequences. It should be noted that the same reasoning does not apply to the syntax violations that were only recognizable as part of the sequence. This amounts to an equal occurrence probability of 50% in both long and short sequences and should not lead to amplitude differences. In line with this, the detection of the syntax violations indeed evoked an early positivity (a P3a) that did not differ between long/ short contexts (for similar results, see also Sammler, Novembre, et al., 2013). Altogether, the perceptual detection of both syntax- and manner-related violations elicited a P3a; however, rather than a motor reprogramming phase as observed in the syntax violation (centroparietal negativity between 520 and 800 msec), the salient fingering manipulation evoked a following P3b that might reflect memory updating processes dependent on the behaviorally relevant stimulus (Polich, 2007). As a final remark, we did not find a syntax-related early anterior negativity as is usually evoked by music-syntactic violations in the auditory domain (i.e., an ERAN; Koelsch, 2009; Koelsch, Gunter, Friederici, & Schröger, 2000) and as was found in our previous study, in which chord progressions were presented as videos (Sammler, Novembre, et al., 2013). This suggests that the early anterior negativities might be specifically tied (i) to the auditory detection of music-syntactic irregularities (but see Gunter, Schmidt, & Besson, 2003) and/or (ii) to the perceptual continuity of the musical input as present in dynamic auditory and video streams but less so in discrete photo series. The comparison of music-syntactic processing in perception and production and the potential beneficial effect of real motion on harmonic priming are interesting topics for future research.

Conclusion In line with the notion of action hierarchy, we distinguished syntax-related motor programs operating at high levels of action planning from lower levels of specific movement selection. Using a priming paradigm involving the execution of chord progressions, we showed that expert pianists make motor predictions concerning the musical goal (Syntax) rather than the manner of execution (Manner). Building on previous findings (Sammler, Novembre, et al., 2013; Novembre & Keller, 2011), our results provide further evidence for motor planning Bianco et al. 51 based on long-term music-syntactic knowledge (i.e., a grammar of action based on musical harmony) and for the priority of the motor plan related to the distal goal over the specific ways to achieve it. Although our EEG data do not give specific information on the generators underlying syntax and manner processing, their different electrophysiological and behavioral patterns may indicate different mechanisms in the planning of the musical goal and the movement used to achieve it. We suggest that, during production of musical sequences, motor predictions of the musical goal are driven by the harmonic structure of the musical context recognized through internalized syntactic knowledge of pianists. Critically, given a certain predictable context, the motor program of the distal musical goal might operate at high levels of the action control hierarchy and be incrementally translated to lower levels of movement kinematics at the very late stage of motor preparation. This weighing of action features (i.e., a weak, thus flexible, preselection of the optimal movement associated to the goal) would constitute an advantage in terms of more efficient performance and interactions with unexpected external changes. Finally, the notion that, through years of intensive motor practice, syntactic rules are motorically acquired, that is, a translation of musical syntax into a “grammar of action,” might speak for a training-dependent motor plasticity toward an emergent syntax-based motor control. Whether this phenomenon occurs in other human actions associated with syntactic structures, such as speech, is an intriguing prospect for future investigations.

Acknowledgments We are grateful to the pianists who participated in this experiment and to B. Pace who lent his hand for the stimulus preparation. We also thank S. Gutekunst for

technical support. This research was funded by Max Planck Society. Reprint requests should be sent to Roberta Bianco, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstraße 1a, 04103 Leipzig, Germany, or via e-mail: bianco@cbs.mpg.de.

REFERENCES Aglioti, S. M., Cesari, P., Romani, M., & Urgesi, C. (2008). Action anticipation and motor resonance in elite basketball players. *Nature Neuroscience*, 11, 1109–1116. Bekkering, H., Wohlschläger, A., & Gattis, M. (2000). Imitation of gestures in children is goal-directed. *Quarterly Journal of Experimental Psychology*, 53, 153–164. Bortz, J. J., & Doering, N. (2003). *Forschungsmethoden und Evaluation: Fuer Human- und Sozialwissenschaftler*. Berlin: Springer. Bubic, A., von Cramon, D. Y., & Schubotz, R. I. (2010). Prediction, cognition and the brain. *Frontiers in Human Neuroscience*, 4, 25. Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: An fMRI study with expert dancers. *Cerebral Cortex*, 15, 1243–1249. Cattaneo, L., Caruana, F., Jezzini, A., & Rizzolatti, G. (2009). Representation of goal and movements without overt motor behavior in the human motor cortex: A transcranial magnetic stimulation study. *Journal of Neuroscience*, 29, 11134–11138. Chaminade, T., Meltzoff, A. N., & Decety, J. (2002). Does the end justify the means? A PET exploration of the mechanisms involved in human imitation. *Neuroimage*, 15, 318–328. Clarke, E., Parncutt, R., Raekallio, M., & Sloboda, J. (1997). Talking fingers: An interview study of pianists' views on fingering. *Musicae Scientiae*, 1, 87–107. Clerget, E., Winderickx, A., Fadiga, L., & Olivier, E. (2009). Role of Broca's area in encoding sequential human actions: A virtual lesion study. *NeuroReport*, 20, 1496–1499. Cunnington, R., Windischberger, C., Deecke, L., & Moser, E. (2003). The preparation and readiness for voluntary movement: A high-field event-related fMRI study of the Bereitschafts-BOLD response. *Neuroimage*, 20, 404–412. Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134, 9–21. Drost, U. C., Rieger, M., Brass, M., Gunter, T. C., & Prinz, W. (2005). When hearing turns into playing: Movement induction by auditory stimuli in pianists. *Quarterly Journal of Experimental Psychology Section A*, 58, 1376–1389. Duncan-Johnson, C. C., & Donchin, E. (1977). On quantifying surprise: The variation of event-related potentials with subjective probability. *Psychophysiology*, 14, 456–467. Duncan-Johnson, C. C., & Donchin, E. (1982). The P300 component of the event-related brain potential as an index of information processing. *Biological Psychology*, 14, 1–52. Fadiga, L., Craighero, L., & D'Ausilio, A. (2009). Broca's area in language, action, and music. *Annals of the New York Academy of Sciences*, 1169, 448–458. Fazio, P., Cantagallo, A., Craighero, L., D'Ausilio, A., Roy, A. C., Pozzo, T., et al. (2009). Encoding of human action in Broca's area. *Brain*, 132, 1980–1988. Fedorenko, E., Patel, A., Casasanto, D., Winawer, J., & Gibson, E. (2009). Structural integration in language and music: Evidence for a shared system. *Memory & Cognition*, 37, 1–9. Fitch, W. T., & Martins, M. D. (2014). Hierarchical processing in music, language, and action: Lashley revisited. *Annals of the New York Academy of Sciences*, 1316, 87–104. Friederici, A. D. (2011). The brain basis of language processing: From structure to function. *Physiological Reviews*, 91, 1357–1392. Gellrich, M., & Parncutt, R. (1998). Piano technique and fingering in the eighteenth and nineteenth centuries: Bringing a forgotten method back to life. *British Journal of Music Education*, 15, 5–23. Gómez, C. M., Flores, A., Digiaco, M. R., Ledesma, A., & González-Rosa, J. (2008). P3a and P3b components associated to the neurocognitive evaluation of invalidly cued targets. *Neuroscience Letters*, 430, 181–185. Grafton, S. T. (2009). Embodied cognition and the simulation of action to understand others. *Annals of the New York Academy of Sciences*, 1156, 97–117. Grafton, S. T., & Hamilton, A. F. D. C. (2007). Evidence for a distributed hierarchy of action representation in the brain. *Human Movement Science*, 26, 590–616. Guerra-Filho, G., & Aloimonos, Y. (2012). The syntax of

human actions and interactions. *Journal of Neurolinguistics*, 25, 500–514. Gunter, T. C., Schmidt, B. H., & Besson, M. (2003). Let's face the music: A behavioral and electrophysiological exploration of score reading. *Psychophysiology*, 40, 742–751. Haggard, P. (2008). Human volition: Towards a neuroscience of will. *Nature Reviews Neuroscience*, 9, 934–946. Hamilton, A. F. D. C., & Grafton, S. T. (2006). Goal representation in human anterior intraparietal sulcus. *Journal of Neuroscience*, 26, 1133–1137. Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298, 1569–1579. Hund-Georgiadis, M., & von Cramon, D. Y. (1999). Motorlearning-related changes in piano players and non-musicians revealed by functional magnetic-resonance signals. *Experimental Brain Research*, 125, 417–425. Katz, J., & Pesetsky, D. (2011). The identity thesis of language and music. ling.auf.net/lingBuzz/000959. Keele, S. W., & Summers, J. J. (1976). The structure of motor programs. In G. E. Stelmach (Ed.), *Motor control: Issues and trends* (pp. 109–142). New York: Academic Press. Keller, P. E. (2012). What movement force reveals about cognitive processes in music performance. In A. Mornell (Ed.), *Art in motion II* (pp. 115–153). Frankfurt, Germany: Peter Lang. Koelsch, S. (2005). Neural substrates of processing syntax and semantics in music. *Current Opinion in Neurobiology*, 15, 207–212. Koelsch, S. (2009). Music-syntactic processing and auditory memory: Similarities and differences between ERAN and MMN. *Psychophysiology*, 46, 179–190. Koelsch, S., Gunter, T., Friederici, A. D., & Schröger, E. (2000). Brain indices of music processing: “Nonmusicians” are musical. *Journal of Cognitive Neuroscience*, 12, 520–541. Koelsch, S., Gunther, T. C., Wittfoth, M., & Sammler, D. (2005). Interaction between syntax processing in language and in music: An ERP study. *Journal of Cognitive Neuroscience*, 17, 1565–1577. Koeneke, S., Lutz, K., Herwig, U., Ziemann, U., & Jäncke, L. (2006). Extensive training of elementary finger tapping movements changes the pattern of motor cortex excitability. *Experimental Brain Research*, 174, 199–209. Koski, L., Wohlschläger, A., Bekkering, H., Woods, R. P., Dubeau, M.-C., Mazziotta, J. C., et al. (2002). Modulation of motor and premotor activity during imitation of targetdirected actions. *Cerebral Cortex*, 12, 847–855. Lago, A., & Fernandez-del-Olmo, M. (2011). Movement observation specifies motor programs activated by the action observed objective. *Neuroscience Letters*, 493, 102–106. Lashley, K. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior* (pp. 112–131). New York: Wiley. Leuthold, H., & Jentzsch, I. (2002). Spatiotemporal source localisation reveals involvement of medial premotor areas in movement reprogramming. *Experimental Brain Research*, 144, 178–188. Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca's area: An MEG study. *Nature Neuroscience*, 4, 540–545. Michel, C. M., Murray, M. M., Lantz, G., Gonzalez, S., Spinelli, L., & Grave de Peralta, R. (2004). EEG source imaging. *Clinical Neurophysiology*, 115, 2195–2222. Moro, A. (2014a). On the similarity between syntax and actions. *Trends in Cognitive Sciences*, 18, 109–110. Moro, A. (2014b). Response to Pulvermüller: The syntax of actions and other metaphors. *Trends in Cognitive Sciences*, 18, 221. Novembre, G., & Keller, P. E. (2011). A grammar of action generates predictions in skilled musicians. *Consciousness and Cognition*, 20, 1232–1243. Palmer, C., & Meyer, R. K. (2000). Conceptual and motor learning in music performance. *Psychological Science*, 11, 63–68. Palmer, C., & Pfordresher, P. Q. (2003). Incremental planning in sequence production. *Psychological Review*, 110, 683–712. Pastra, K., & Aloimonos, Y. (2012). The minimalist grammar of action. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, 367, 103–117. Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, 6, 674–681. Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive*

Neuroscience, 10, 717–733. Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences*, 10, 233–238. Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118, 2128–2148. Pulvermüller, F. (2014). The syntax of action. *Trends in Cognitive Sciences*, 18, 219–220. Pulvermüller, F., & Fadiga, L. (2010). Active perception: Sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience*, 11, 351–360. Rektor, I. (2000). Parallel information processing in motor systems: Intracerebral recordings of readiness potential and CNV in human subjects. *Neural Plasticity*, 7, 65–72. Rohrmeier, M. (2011). Towards a generative syntax of tonal harmony. *Journal of Mathematics and Music*, 5, 35–53. Rohrmeier, M. A., & Koelsch, S. (2012). Predictive information processing in music cognition. A critical review. *International Journal of Psychophysiology*, 83, 164–175. Rosenbaum, D. A., Kenny, S. B., & Derr, M. A. (1983). Hierarchical control of rapid movement sequences. *Journal of Experimental Psychology: Human Perception and Performance*, 9, 86–102. Sammler, D., Koelsch, S., Ball, T., Brandt, A., Grigutsch, M., Huppertz, H. J., et al. (2013). Co-localizing linguistic and musical syntax with intracranial EEG. *Neuroimage*, 64, 134–146. Sammler, D., Koelsch, S., & Friederici, A. D. (2011). Are left fronto-temporal brain areas a prerequisite for normal music-syntactic processing? *Cortex*, 47, 659–673. Sammler, D., Novembre, G., Koelsch, S., & Keller, P. E. (2013). Syntax in a pianist's hand: ERP signatures of “embodied” syntax processing in music. *Cortex*, 49, 1325–1339. Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review*, 82, 225–260. Shaffer, L. H. (1981). Performances of Chopin, Bach, and Bartok: Studies in motor programming. *Cognitive Psychology*, 13, 326–376. Sharbrough, F., Chatrian, G.-E., Lesser, R. P., Lueders, H., Nuwer, M., & Picton, T. W. (1991). American Electroencephalographic Society guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology*, 8, 200–202. Slevc, L. R., Rosenberg, J. C., & Patel, A. D. (2009). Making psycholinguistics musical: Self-paced reading time evidence for shared processing of linguistic and musical syntax. *Psychonomic Bulletin & Review*, 16, 374–381. Sloboda, J. A., Clarke, E. F., Parncutt, R., & Raekallio, M. (1998). Determinants of finger choice in piano sight-reading. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 185–203. Bianco et al. 53 Smith, J. L., Johnstone, S. J., & Barry, R. J. (2006). Effects of pre-stimulus processing on subsequent events in a warned go/nogo paradigm: Response preparation, execution and inhibition. *International Journal of Psychophysiology*, 61, 121–133. Summers, J. J., & Anson, J. G. (2009). Current status of the motor program: Revisited. *Human Movement Science*, 28, 566–577. Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *Science*, 331, 1279–1285. Tillmann, B. (2012). Music and language perception: Expectations, structural integration, and cognitive sequencing. *Topics in Cognitive Science*, 4, 568–584. Uithol, S., van Rooij, I., Bekkering, H., & Haselager, P. (2012). Hierarchies in action and motor control. *Journal of Cognitive Neuroscience*, 24, 1077–1086. Walter, W. G., Cooper, R., Aldridge, V. J., MacCallum, W. C., & Winter, A. L. (1964). Contingent negative variation: An electric sign of sensori-motor association and expectancy in the human brain. *Nature*, 203, 380–384. Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, 131, 460–473. Wohlschläger, A., Gattis, M., & Bekkering, H. (2003). Action generation and action perception in imitation: An instance of the ideomotor principle. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, 358, 501–515. 54 *Journal of Cognitive Neuroscience* Volume 28, Number 1