

Dynamic acoustic salience evokes motor responses

Citation for published version (APA):

Schultz, B. G., Brown, R. M., & Kotz, S. A. (2021). Dynamic acoustic salience evokes motor responses. *Cortex*, 134, 320-332. <https://doi.org/10.1016/j.cortex.2020.10.019>

Document status and date:

Published: 01/01/2021

DOI:

[10.1016/j.cortex.2020.10.019](https://doi.org/10.1016/j.cortex.2020.10.019)

Document Version:

Publisher's PDF, also known as Version of record

Document license:

Taverne

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.



Research Report

Dynamic acoustic salience evokes motor responses

Benjamin G. Schultz¹, Rachel M. Brown¹ and Sonja A. Kotz^{*}

Basic & Applied NeuroDynamics Laboratory, Faculty of Psychology & Neuroscience, Department of Neuropsychology & Psychopharmacology, Maastricht University, the Netherlands

ARTICLE INFO

Article history:

Received 17 December 2019

Reviewed 15 March 2020

Revised 25 June 2020

Accepted 8 October 2020

Action editor Stephen Jackson

Published online 18 November 2020

Keywords:

Audio-motor integration

Acoustic salience

Involuntary motor responses

sEMG

ABSTRACT

Audio-motor integration is currently viewed as a predictive process in which the brain simulates upcoming sounds based on voluntary actions. This perspective does not consider how our auditory environment may trigger **involuntary action** in the absence of prediction. We address this issue by examining the relationship between acoustic salience and involuntary motor responses. We investigate how acoustic features in music contribute to the perception of salience, and whether those features trigger involuntary peripheral motor responses. Participants with little-to-no musical training listened to musical excerpts once while remaining still during the recording of their muscle activity with surface electromyography (sEMG), and again while they continuously rated perceived salience within the music using a slider. We show cross-correlations between 1) salience ratings and acoustic features, 2) acoustic features and spontaneous muscle activity, and 3) salience ratings and spontaneous muscle activity. Amplitude, intensity, and spectral centroid were perceived as the most salient features in music, and fluctuations in these features evoked involuntary peripheral muscle responses. Our results suggest an involuntary mechanism for audio-motor integration, which may rely on brainstem-spinal or brainstem-cerebellar-spinal pathways. Based on these results, we argue that a new framework is needed to explain the full range of human sensorimotor capabilities. This goal can be achieved by considering how predictive and reactive audio-motor integration mechanisms could operate independently or interactively to optimize human behavior.

© 2020 Elsevier Ltd. All rights reserved.

1. Introduction

The human nervous system closely integrates perception and action. This integration allows the body to respond to and

align itself with a dynamically changing environment. Although neural and cognitive models have increased our understanding of perception-action coupling in voluntary actions, understanding of this coupling in *involuntary* action remains limited. In particular, it is not known whether highly

^{*} Corresponding author. Basic & Applied NeuroDynamics Laboratory, Maastricht University, Universiteitssingel 40, 6229 ER, Maastricht, the Netherlands.

E-mail addresses: sonja.cimon@gmail.com, sonja.kotz@maastrichtuniversity.nl (S.A. Kotz).

¹ First authorship is shared between these authors.

<https://doi.org/10.1016/j.cortex.2020.10.019>

0010-9452/© 2020 Elsevier Ltd. All rights reserved.

complex and continuous stimuli, such as speech or music, can evoke involuntary muscle responses. For example, the startle response is an involuntary motor response to a stimulus that stands out relative to neighboring stimuli and, consequently, is perceived as salient. This reflex normally occurs to extremely sudden sounds, and is mediated by a brainstem-spinal pathway. If a muscle response can also be evoked by prominent changes embedded within complex and continuous auditory streams such as speech or music, it would imply that a similarly fast and direct auditory-brainstem pathway might also contribute to complex listening tasks. The present experiment examined whether involuntary motor responses are evoked by changes in acoustic features within complex and naturalistic musical excerpts, and whether these same acoustic features are perceived as salient. We aimed to establish a three-fold link between 1) involuntary muscle responses and acoustic features within naturalistic stimuli, 2) involuntary muscle responses and perceived salience, and 3) acoustic features and perceived salience.

The first aim of this study was to examine whether involuntary muscle responses can be evoked by changes within highly complex auditory stimuli. The role of involuntary audio-motor integration has received relatively little theoretical and empirical attention, relative to voluntary audio-motor integration. Recent neural and cognitive models for sensorimotor integration emphasize how voluntary movements can filter or enhance perception through prediction (Blakemore, Frith, & Wolpert, 1999; Morillon & Schroeder, 2015; Wolpert, Ghahramani, & Jordan, 1995). Actions could be used to focus attention on predictable points in time, potentially through oscillatory neural activity that entrains attention-driven perceptual gating to external stimuli (Morillon & Baillet, 2017; Morillon, Hackett, Kajikawa, & Schroeder, 2015; Morillon, Schroeder, & Wyart, 2014; Schroeder & Lakatos, 2009; Schroeder, Lakatos, Kajikawa, Partan, & Puce, 2008; Schroeder, Wilson, Radman, Scharfman, & Lakatos, 2010). This perspective helps us understand how we align our movements and/or attention with regularities, or predictable patterns, such as temporal structures within speech or music. What is not yet understood is whether complex external stimuli that lack a clearly predictable structure may still engage the motor system (Bendixen, SanMiguel, & Schröger, 2012). Organisms may not always have the capacity or information available to predict accurately, and we must be able to respond to changes that are unpredictable.

Neural pathways for startle and orienting responses are useful in this case to respond to unexpected changes in the environment. The vertebrate motor system is known to show an involuntary response to ambient changes in sound. Sounds that are highly intense and infrequent relative to other ambient sounds evoke a startle reflex (Brown et al., 1991a, 1991b; Davis, Gendelman, Tischler, & Gendelman, 1982). The caudal brainstem has been identified as the origin of this reflex, which is thought to propagate from the auditory nerve to the bulbopontine reticular formation to the motor periphery. This response allows humans and other vertebrates to detect and automatically respond to a sudden, unexpected change in the environment quickly (<150 ms; Brown et al., 1991a). Importantly, even sound changes that are not highly intense can evoke responses in sensory receptor organs. The orienting

response, such as turning the eyes or head toward the source of a sound (Johnson & Lubin, 1967; Sokolov, 1963), only depends on a change in sound parameters regardless of the magnitude of those changes. The orienting response occurs immediately following a change (e.g., a sound on or a sound off), and it disappears when the feature that changed maintains the same state. The orienting response involves a combination of neuronal firing, autonomic and muscle responses, and may also rely on the brainstem (Sokolov, 1963), potentially the superior colliculus (Goodale & Murison, 1975). Finally, the ability to quickly adjust or adapt ongoing behavior to changes in ambient sounds may further rely on the cerebellum (Sokolov, Miall, & Ivry, 2017), potentially through fast cerebellar-to-frontal cortex transmission (Schwartz & Kotz, 2013, 2016). Thus, humans possess subcortical machinery whereby features of sound can innervate the muscles without conscious control.

The utility of these reflexive pathways for survival is clear, such as responding efficiently to potential threats (Sokolov, 1963). Beyond their use for survival, these phylogenetically older pathways may also enable organisms to respond to dynamic and complex acoustic stimuli such as spoken language or music. The question remains, in which contexts can involuntary audio-motor innervation occur? Do only highly salient and rare sounds innervate the muscles or can changes along a range of acoustic features innervate the muscles? The few studies that have examined spontaneous human movement to sounds demonstrated that certain features of a musical texture, such as the entrance of different instruments, can influence the amount and quality of spontaneous overt (observable) movement (Burger, Thompson, Luck, Saarikallio, & Toiviainen, 2014; Hurley, Martens, & Janata, 2014; Janata, Tomic, & Haberman, 2012). However, these studies sought to document alignment of overt movement with general aspects of musical structure, for example, a regular beat. It remains an open question whether features of sound that do not correspond to a predictable structure within a complex sound stream such as music, can evoke involuntary, subthreshold muscle activity. Music and speech contain a wealth of acoustic features that often converge on a predictable temporal structure (Jones, 2001; Rothermich, Schmidt-Kassow, & Kotz, 2012), yet they also contain unpredictable events that may also engage the motor system. Given that the motor system is flexibly attuned not only to the buildup of regularities over time (Morillon, Schroeder, Wyart, & Arnal, 2016), but also to instantaneous relative changes in stimulus features (Sokolov, 1963), the motor system may be sensitive to a range of features of sound even within complex listening contexts. Involuntary audio-motor integration could operate across many or all sound processing contexts, even contexts that require detection of and alignment to patterns, such as, tapping along with a beat in music (cf. Repp, 2005; Repp & Su, 2013) or coordinating responses to a conversation partner (e.g., Schultz et al., 2016). We address this issue by examining whether a range of acoustic changes in music trigger involuntary motor responses.

The second aim of the study was to examine the link between acoustic features that may evoke involuntary motor responses and the perception of salience. The human auditory system is highly attuned to change. This change sensitivity is

captured by the concept of salience. Salience is the perception of a stimulus as distinct, prominent, or conspicuous relative to neighboring stimuli. Salience is evoked by a change or contrast, often sudden or sharp, along a particular stimulus feature (Borji, Sihite, & Itti, 2013; Parkhurst, Law, & Niebur, 2002). Visual salience and the ways in which visual features highlight points in space are well-documented (Borji et al., 2013; Gottlieb, Kusunoki, & Goldberg, 1998; Itti, 2006; Parkhurst et al., 2002; Thompson & Bichot, 2004). Acoustic salience is far less studied, despite its utility in highlighting points in time with high precision (Benoit et al., 2014; Dalla Bella, Benoit, Farrugia, Schwartz, & Kotz, 2015; Ellis & Jones, 2009; Fernandez-Del-Olmo et al., 2013; Rodger & Craig, 2016). The neural signatures of acoustic salience are well documented. For instance, the mismatch negativity (MMN) response is a comparative evoked potential that is understood to be a measure of pre-attentive change detection (Schröger, 1997; Schröger & Winkler, 1995). However, we still know little about how particular acoustic features contribute to the perception of salience, particularly within naturalistic stimuli that combine multiple features. Here, we aim to isolate particular acoustic features within music and identify those that contribute to the perception of salience, as well as to involuntary motor responses.

Salience can be derived from multiple acoustic features. Acoustic features such as amplitude and intensity (perceived as loudness), frequency (perceived as pitch), spectral properties of individual sounds (timbre), and interactions between multiple sounds (harmony) can distinguish an individual event at a particular point in time from an immediately preceding event(s), such as, when a breaking object produces a sudden crash while you are listening to someone speaking (Caclin et al., 2006; Caclin, McAdams, Smith, & Winsberg, 2005). The abrupt changes in loudness and timbre features are potential sources of salience that can signal an event at a particular point in time (Chon & McAdams, 2012; McAdams, Winsberg, Donnadieu, De Soete, & Krimphoff, 1995). Here we focus on acoustic features that can define how sound unfolds over time to examine motor responses to instantaneous acoustic events that do not necessarily depend on predictable patterns in acoustic events. We further investigate possible influences of predictable acoustic patterns (e.g., beat structures and harmonic context) on motor responses. We examined amplitude and intensity, timbre, and harmonic features (see *Method* for full descriptions). Amplitude and intensity are known to contribute to the startle response (Brown et al., 1991a; Davis et al., 1982) and, in the context of music, these features correspond to perceived arousal levels conveyed by the music (Dean, Bailes, & Schubert, 2011). Acoustic salience based on timbre has received little attention. Here, we focus on spectral measures that capture global features of the sound spectrum, and have been found to influence the discriminability of different timbres (Caclin et al., 2005; McAdams et al., 1995) and to elicit an MMN response when there is a change along this dimension (Caclin et al., 2006). We therefore examined whether amplitude, intensity, timbre, and harmonic features both evoke involuntary motor responses and contribute to the perception of salience. If so, salient events within complex sounds such as music may evoke involuntary motor responses. This is a crucial missing piece to

understanding the full sensorimotor capacity of the human nervous system. In addition, to examine whether involuntary motor responses to, and perceived salience of, these features was dependent on predictability, we also investigated involuntary motor responses to high-level acoustic features that represent predictability in harmony (harmonic change) and timing (beat and downbeat). We hypothesized that changes in intensity, amplitude, timbre, and harmonic features should evoke spontaneous muscle activity and increased salience ratings, regardless of the degree of harmonic or temporal predictability.

The overall goal of the present study was to establish whether there is a direct link between salient acoustic features and involuntary motor responses within the context of a continuous and dynamic sound stream, in this case music. We investigated the dynamic link between continuous changes in acoustic features, salience perception, and involuntary responses in the motor periphery. First, we address whether perceived salience could be a mechanism behind involuntary motor responses. Second, we investigate which acoustic features elicit changes in salience perception as music unfolds. Third, we determine whether changes in particular acoustic features within a complex sound stream evoke involuntary muscle responses while passively listening to music. Thus, we investigate how the peripheral motor system dynamically responds to salient acoustic features. Participants with little to no musical training listened to musical excerpts once while remaining still during the recording of their muscle activity using surface electromyography (sEMG), and again while they continuously rated perceived salience within the music using a slider. The slider measurement provided a means to capture salience perception in a dynamic manner, allowing sensitivity to dynamically unfolding acoustic signals that may evoke salience. Musical excerpts were 40-sec segments of commercially available musical recordings of various genres, tempi, and temporal regularity (i.e., beat salience). We applied time series analysis with dynamic time warping (DTW) to examine the correspondence between various dynamic acoustic features, salience ratings, and muscle activity. We hypothesized that 1) continuous ratings of perceived salience correspond to spontaneous motor activity, 2) continuous ratings of salience correlate with acoustic features as they unfold in time, that is, that certain acoustic features are perceived as salient, and 3) sEMG activity correlates with acoustic features, that is, that certain acoustic features evoke motor responses.

2. Method

We report how we determined our sample size, all data exclusions (if any), all inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study.

2.1. Participants

Forty-six young healthy adults from Maastricht University, the Netherlands, participated (fourteen males, age range

17–32 years, $M = 21.98$, $SD = 3.58$). Participants had normal hearing and none reported having any cognitive or neuro-motor disorder. There was a variety of musical expertise; 18 participants reported no musical training, 11 reported 1–5 years of musical training, and 17 participants reported more than five years of musical training. Two participants were excluded due to equipment failure and another two were excluded due to a corrupted hard disk drive. Participants gave written informed consent and received vouchers as compensation (VVV vouchers; 10€ per hour). The study was approved by the ethics committee of the Faculty of Psychology and Neuroscience at Maastricht University.

2.2. Stimuli

Stimuli consisted of twelve commercially available songs, with excerpts selected based on variability in intensity, pitch (consonance and dissonance), and timbre. To achieve this, songs were selected, and several sections were extracted (e.g., the bridge, chorus, breakdown, outro) before the acoustic features were assessed. Most songs were selected based on ratings of “groove” from Janata et al. (2012) and some were selected based on beat structures that were potentially difficult to abstract. We further obtained groove ratings in a pilot study performed during stimulus selection (see Table 1). “Groove” refers to the degree to which music elicits the desire to move, and is often associated with a strong underlying beat, or temporal regularity. Songs across the spectrum of groove ratings were chosen to ensure muscle activity occurs across a range of “groovy-ness”. Moreover, the beat strength in the acoustic signal (inter-beat interval CV) and variability of beat entrainment in the tapping task (inter-tap interval CV) encompassed a broad range (see Table 1). One excerpt of 40 sec was extracted from each song. Thus, the final stimuli consisted of 12 40-sec musical excerpts.

2.3. Acoustic features

Acoustic features were extracted from musical excerpts using Sonic Visualizer. *Amplitude* and *intensity* were measures relating to the perception of loudness. The global spectral features (i.e., timbre) examined were *spectral centroid*, *spectral*

crest, and *spectral variance*. *Spectral centroid* refers to the weighted mean of spectrum energy, in other words the amount of power in the frequency range where the most power is concentrated. *Spectral crest* refers to the ratio between maximum spectrum power and mean spectrum power. This describes whether there are one or more frequency bands with relatively high-energy peaks within the spectrum and, thus, whether there are one or more frequency bands that stand out. *Spectral variance* refers to the variability of spectral power, or the variation in the energy peak amplitudes across the frequency bands. We examined two aspects of harmony: *Inharmonicity*, or the degree of momentary dissonance within a short window, and *harmonic change*, or the degree to which harmonicity changed given what has been previously experienced within the excerpt. We included *harmonic change* to capture acoustic changes that were context-dependent and to account for predictable harmonic patterns within the music. In order to account for temporal predictability, we quantified the temporal regularity of the sounds as they unfold in time by applying a beat-tracking algorithm (Böck & Schedl, 2011). This algorithm uses recurrent neural networks to calculate the probability of a *beat* or a *downbeat* (a.k.a. a strong beat) at each specified time point (Essens & Povel, 1985). The variation in harmonic change and temporal regularity across stimuli allowed us to link motor responses to acoustic features that were not dependent on predictable harmonic patterns or the beat (see Table 1).

2.4. Procedure

Participants were seated in a quiet, sound shielded room. Participants were fitted with bipolar sEMG electrodes on their lower and upper arm (flexor carpi radialis and biceps brachii, respectively) and to their palm (abductor pollicis brevis) to measure electrical activity produced by muscle responses. Participants received stimuli through two speakers (JBL 25-AV-LS). The concept of salience was explained to participants but no explicit instruction was given as to which features were to be rated in the music. Specifically, salience was defined as “how much sounds stand out” and told to change their ratings throughout each song depending on how their perception of salience increased or decreased. Participants

Table 1 – Descriptive factors of stimuli including start time of excerpts (duration 40 sec), beats-per-minute (BPM), coefficient of variation (CV) of inter-beat intervals derived from MadMom, CV of inter-tap intervals (ITI) produced by participants, and mean groove ratings obtained from a pilot study during stimulus selection.

Artist (Song)	Excerpt time (sec)	Excerpt BPM	Beat CV	ITI CV	Mean groove (SD)
The Beastie Boys (Sabrosa)	65.8–105.8	157.89	.017	.200	75.51 (24.35)
The Bee Gees (You Should be Dancing)	25.5–85.5	122.45	.011	.226	85.71 (8.25)
Britney Spears (Toxic)	10.5–50.5	142.86	.014	.181	79.59 (13.94)
Django Reinhardt & Stephane Grappelli (Minor Swing)	57.7–97.7	200.00	.382	.165	77.55 (7.64)
Herbie Hancock (Rock it)	30.4–70.4	111.11	.01	.203	69.39 (15.27)
Marvin Gaye (Sexual Healing)	71.4–110.4	95.24	.008	.161	83.67 (26.64)
Radiohead (Just)	162.3–202.3	171.43	.484	.194	67.35 (21.37)
Snarky Puppy (Lingus)	420.1–460.1	166.67	.629	.276	46.94 (17.91)
Stevie Wonder (Superstition)	225.3–265.3	103.45	.019	.134	81.63 (10.80)
Stitt Rollins Gillespie (Eternal Triangle)	589.7–629.7	150.00	.030	.183	61.22 (17.91)
The Beach Boys (God Only Knows)	106.5–146.5	117.65	.025	.155	53.06 (17.91)
The Glenn Miller Orchestra (In the Mood)	88.3–128.3	176.47	.412	.272	73.47 (12.85)

listened to each excerpt three times. First, participants listened to each excerpt while sitting still, and after each excerpt they rated their familiarity with the music. Second, participants listened to the music excerpts again while continuously rating their perceived salience of the music using a slider. Third, participants listened a third time to each musical excerpt and tapped along with the music according to any beat or temporal regularity they could detect. The tasks were performed in this order to a) ensure that implicit motor responses were not attenuated/driven by expectations from recent presentations and, b) to reduce reaction times in rating salience for changes in acoustic features.

2.5. Behavioral and EMG data analysis

Time series analysis with DTW was used to examine the correspondence between acoustic features, salience ratings, and muscle responses for each stimulus and participant (Levitin, Nuzzo, Vines, & Ramsay, 2007). DTW allowed comparing relationships between variables that may have non-linear temporal relationships with acoustic features (e.g., muscle responses and continuous subjective ratings (Bishop, Bailes, & Dean, 2013)). To ensure comparability, DTW was performed with the same parameters for all pairwise combinations of muscle responses and salience ratings with acoustic features. Data were smoothed using a moving average within a 50 msec window; then DTW was performed to align local patterns using the smallest Euclidean distance cost within a time window of 2000 msec. The final dependent variable was the Fisher transformed cross-correlation coefficient at lag 0 after DTW for each pair of responses and/or features (see Fig. 1).

The Fisher transformation was used to ensure cross-correlation coefficients followed a normal distribution. Chance cross-correlation coefficients were obtained through jackknifing procedures whereby the response was paired with the acoustic features of all other songs and underwent the same procedure including smoothing and DTW. The mean Fisher-transformed cross-correlation coefficient of these mismatched pairs represented chance levels. If the Fisher-transformed cross-correlation coefficients were not significantly above chance, they were subjected to Bayes factor tests to examine evidence for the null hypothesis. Bayes factor values between 1 and 3 represent anecdotal evidence for the null, 3–10 represents substantial evidence for the null, 10–30 represents strong evidence for the null, 30–100 represents very strong evidence for the null, and over 100 represents extreme evidence for the null (see Jeffreys, 1961).

To test the hypothesis that salience ratings correspond to spontaneous motor activity, the relationship between continuous motor activity in three sEMG electrodes (hands, lower arm, upper arm) and continuous salience ratings were compared to estimated chance levels using a linear mixed-effects model (LMEM). The Fisher-transformed cross-correlation coefficients were fitted to LMEMs separately for the relationships between 1) salience ratings and sEMG activity, 2) salience ratings and acoustic features, and 3) sEMG activity and acoustic features. For the model for salience ratings and sEMG activity, the sEMG electrode position (3; hand, lower arm, upper arm) was included as a fixed factor. For the model

for salience ratings and acoustic features, the fixed factor was Feature (any of the following that were above chance: amplitude, intensity, spectral centroid, spectral crest, spectral variance, inharmonicity, harmonic change, beat probability, and downbeat probability). For the model for sEMG activity and acoustic features, both Feature and sEMG electrode position were included as fixed factors. Participant and Song were included as random effects for every LMEM. Models were selected based on the lowest Akaike's Information Criteria (AIC) value when including the interaction term or only main effects.

3. Results

3.1. The relationship between perceived salience and motor activity

To test the hypothesis that salience ratings correspond to spontaneous motor activity, the relationship between continuous motor activity in three sEMG electrodes (hands, lower arm, upper arm) and continuous salience ratings were compared to estimated chance levels using a LMEM. All relationships were significantly above chance ($ps < .002$, $OR = 10.81$; see Fig. 2). This relationship significantly decreased in strength from the hand to the lower arm and from the lower arm to the upper arm ($ps < .04$, $OR = 148.34$). These results support the hypothesis that subjective ratings of salience significantly correspond to spontaneous motor activity.

3.2. Revealing which acoustic features influence salience perception

To assess which acoustic features are perceived as salient, the relationship between salience ratings and acoustic features were compared to estimated chance levels using a LMEM. Continuous salience ratings significantly correlated with changes in amplitude, intensity, spectral centroid, spectral crest, harmonic change, and inharmonicity above chance levels ($ps < .004$). There was strong evidence for the null hypothesis for spectral variance ($BF_{01} = 13.88 \pm .0002$), and also for beat ($BF_{01} = 14.07 \pm .0002$) and downbeat ($BF_{01} = 13.55 \pm .0002$) probability, and these did not demonstrate relationships with salience ratings above chance levels ($ps > .36$).

To examine which features showed the strongest relationship with salience ratings, features that related to salience ratings significantly above chance were subjected to the LMEM. There was a main effect of Feature [$F(5, 61.46) = 23.28$, $p < .001$, $OR = 1155.96$] and planned comparisons yielded no significant differences between amplitude, intensity, and spectral centroid ($ps > .76$), and these three features were significantly more correlated with salience ratings than spectral crest ($ps < .001$), harmonic change ($ps < .001$), and inharmonicity ($ps < .02$; see Fig. 3a). Inharmonicity showed stronger relationships with salience ratings than harmonic change ($p < .009$) but not spectral crest ($p = .1$), and there was no significant difference between harmonic change and spectral crest ($p = .97$). Overall, these results support the

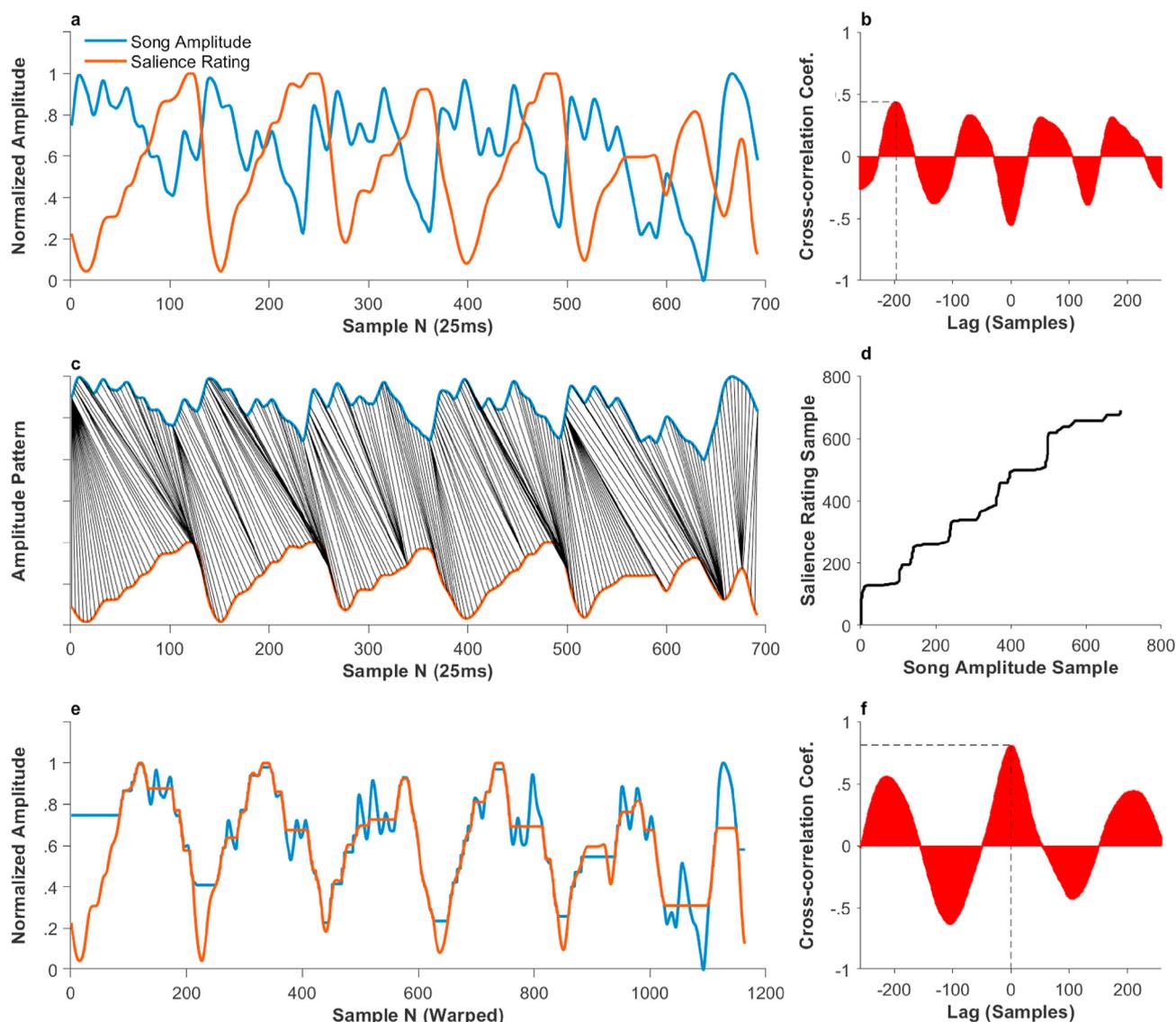


Fig. 1 – Example of the dynamic time warping procedure. a) Original time series. b) Cross-correlation coefficients between the two timeseries. Non-warping methods often take the maximum cross-correlation coefficient within some range (dotted black lines). c) Warping paths (black lines) between the two timeseries. d) Alignment path of relative samples. e) Warped data. f) Cross-correlation coefficients of the between the two warped time series. We used the cross-correlation coefficient at lag zero (dotted black lines), even if this was not the maximum cross-correlation coefficient.

hypothesis that amplitude, intensity, and spectral centroid are perceived as the most salient features in music.

3.3. Revealing which acoustic features evoke motor responses

To assess which acoustic features evoked involuntary motor responses, the relationships between sEMG recordings and acoustic features were compared to estimated chance levels using a LMEM. The best model fit contained no interactions between Baseline, Feature, and sEMG channel and, therefore, data were averaged across sEMG channels. sEMG activity significantly correlated with changes in amplitude, intensity, spectral centroid, spectral crest, inharmonicity, and spectral variance above chance levels ($ps < .008$). Other acoustic

features did not display relationships with sEMG activity above chance levels ($ps > .15$), and there was strong evidence for the null hypothesis for harmonic change ($BF_{01} = 16.66 \pm 0$), and beat ($BF_{01} = 23.30 \pm 0$) and downbeat ($BF_{01} = 15.10 \pm 0$) probability.

To test which acoustic features showed the strongest relationships with motor activity, features that related to sEMG activity significantly above chance were subjected to the LMEM. As above, the best model fit contained no interactions between Feature and EMG channel and, therefore, only main effects are reported. The significant main effect of sEMG channel [$F(2, 196.67) = 115.22, p < .001, OR = 194.81$] indicated that there were stronger relationships between acoustic features and sEMG activity for the hand compared to the lower and upper arm ($ps < .001$), and for the lower arm compared to

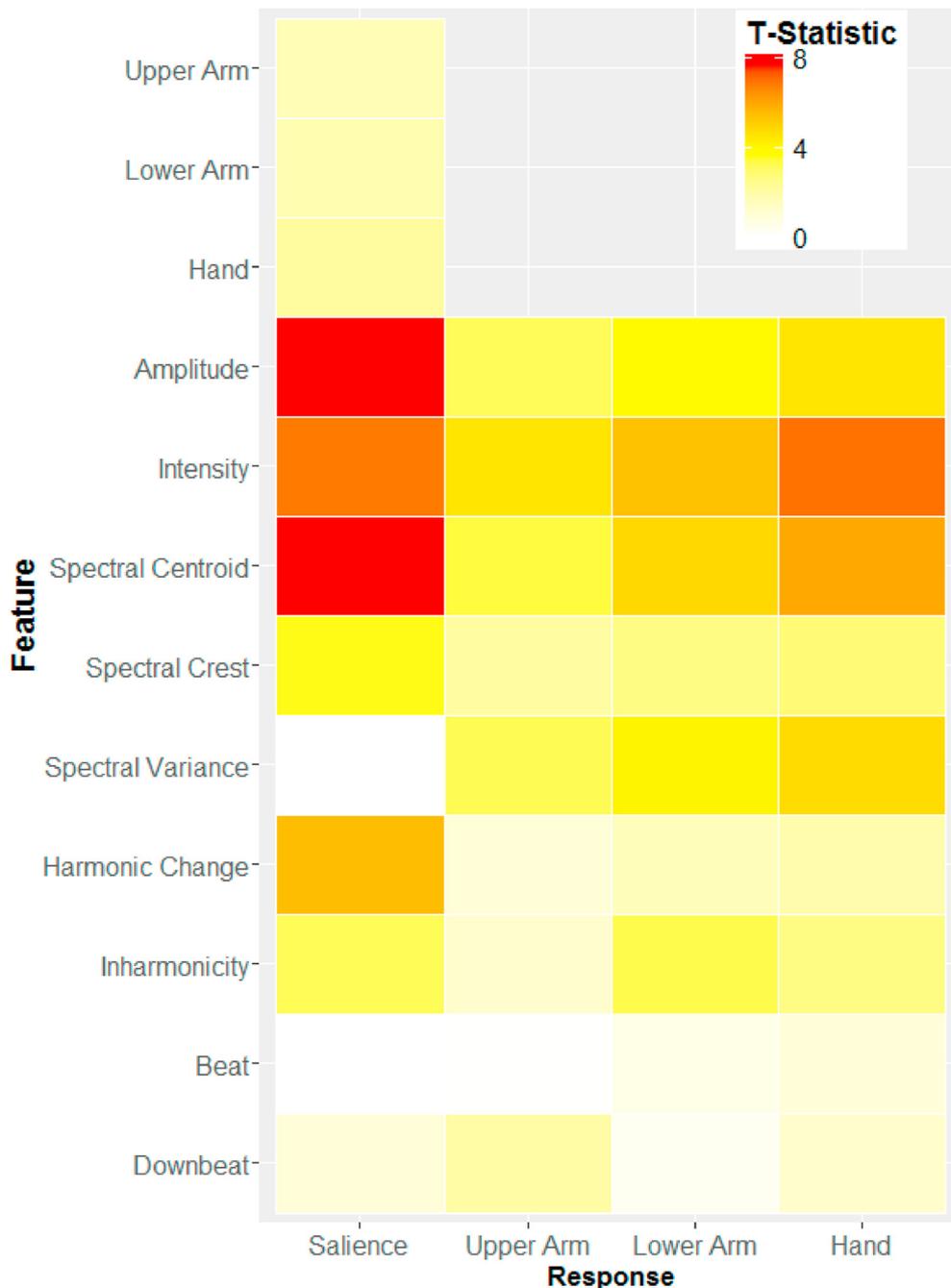


Fig. 2 – T-statistics for comparisons between Fisher-transformed cross-correlation coefficients and approximations of chance via jackknifing. T-statistics that were less than zero were set to zero, and greater than 8 were set to 8 to improve scale visibility.

the upper arm ($p < .001$). There was a significant main effect of Feature [$F(5, 58.30) = 29.51, p < .001, OR = 2250.54$]; pairwise comparisons indicated that intensity, spectral centroid, and spectral variance did not significantly differ from each other ($ps > .99$) and showed the largest relationships with sEMG activity compared to amplitude ($ps < .03$), inharmonicity ($ps < .001$), and spectral crest ($ps < .001$; see Fig. 3b). Amplitude demonstrated the next highest relationship with sEMG activity and was significantly greater than spectral crest ($p < .001$) but not inharmonicity ($p = .1$). Inharmonicity demonstrated a stronger relationship with sEMG activity than spectral crest

($p = .01$). These results support the hypothesis that amplitude, intensity, and spectral centroid evoke motor activity and further show that motor activity also relates to changes in inharmonicity, spectral crest, and spectral variance.

4. Discussion

The present study examined whether involuntary motor responses are evoked by changes in acoustic features within complex, continuous, and naturalistic musical excerpts, and

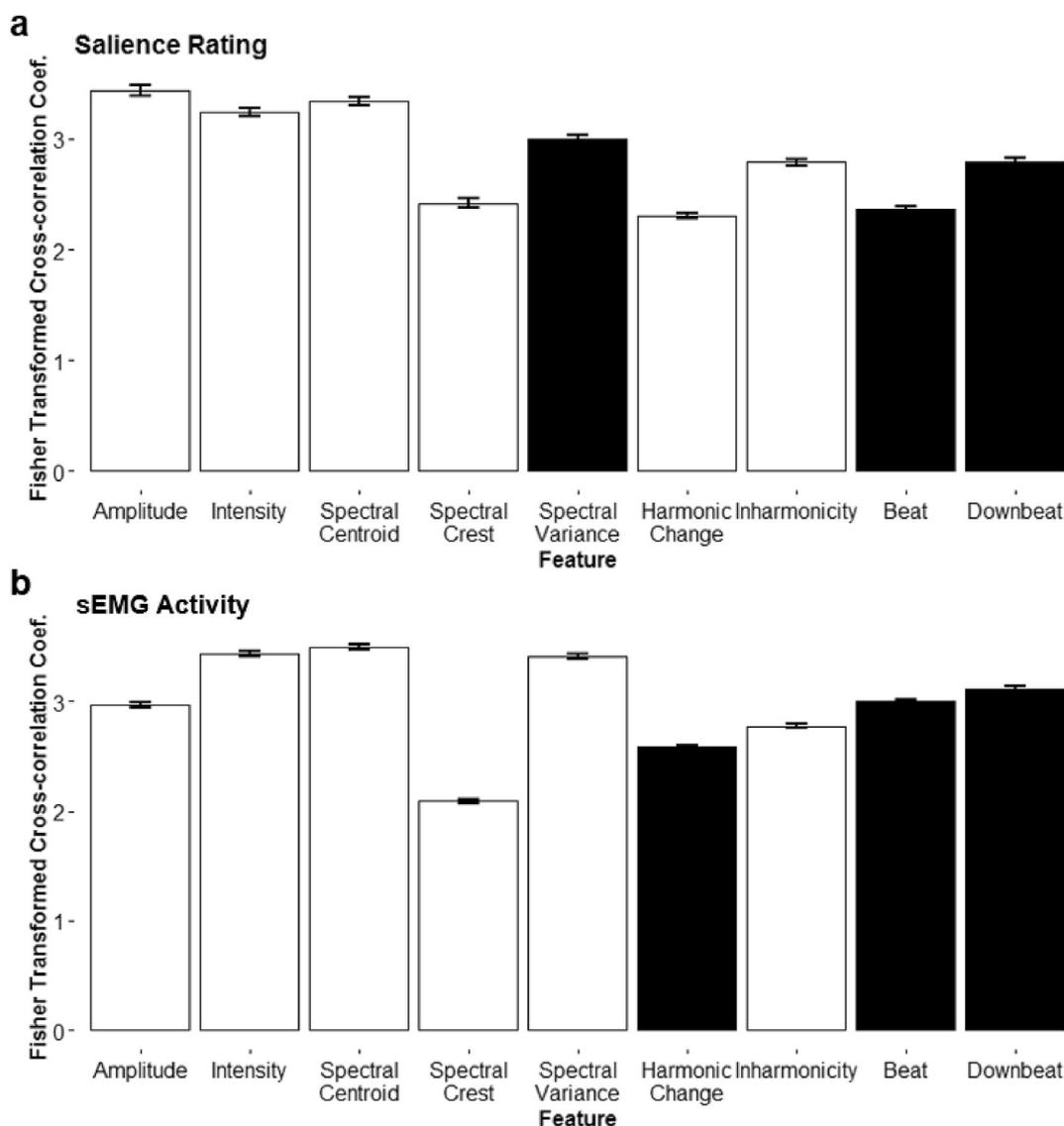


Fig. 3 – Means of the Fisher transformed cross-correlation coefficient for salience ratings (top panel) and motor responses pooled over sEMG electrodes (bottom panel). Black bars are used for features that did not produce Fisher transformed cross-correlation coefficients above chance levels. Error bars represent standard error of the mean.

whether these same acoustic features are perceived as salient. We show that changes in several acoustic features within music corresponded to changes in involuntary motor responses and to perceived acoustic salience. Amplitude, intensity, and spectral centroid were perceived as the most salient features in music, and they correlated with peripheral muscle responses. We also show a gradient of sensitivity in the motor periphery to salient features from the hand to the lower arm to the upper arm. This gradient may reflect increased audio-motor coupling due to motor dexterity in the case of hand muscles compared to arm muscles, or it could reflect a prepotent response to move one's hand by tapping or clapping to music. Neither salience ratings nor sEMG activity correlated with changes in beat and downbeat probability, indicating that auditory salience responses are not governed by beat-based regularity but, instead, by dynamic changes in acoustic features that correspond to amplitude, intensity, and

spectral energy. Taken together, these results suggest that amplitude, intensity, and spectral centroid are salient acoustic features that evoke involuntary motor responses.

In addition to evoking involuntary motor responses, changes in amplitude, intensity, spectral centroid, spectral crest, and inharmonicity were also perceived as salient. Inharmonicity and spectral crest showed lower correlations with motor responses and salience than other features, and spectral variance corresponded to motor responses, but not to perceived salience. The perception of amplitude and intensity changes as salient is consistent with the type of stimuli that normally evoke acoustic startle responses, namely an intensely loud sound (e.g., 120 dB, compared to a normal level of 70 dB; Davis et al., 1982; Fernandez-Del-Olmo et al., 2013) and with orienting responses to sound intensity (Brown et al., 1991a; Johnson & Lubin, 1967). Perceived loudness is also associated with emotional responses to

sounds (Olsen & Stevens, 2013). For instance, increased subjective or physiological arousal responses correlated with gradual increases or decreases in acoustic intensity of natural sounds and music (Bradley & Lang, 2000; Dean et al., 2011; Tajadura-Jiménez, Väljamäe, Asutay, & Västfjäll, 2010). In addition, spectral features, mainly spectral centroid, corresponded with ratings of perceived salience. This is consistent with the role of spectral centroid in discriminating timbres and the sensitivity of the MMN response to changes along spectral centroid (Caclin et al., 2005, 2006; McAdams et al., 1995). Changes in these features likely stand out for listeners by marking particular points in time. Instantaneous points of salience, when repeated often enough, could also be a precursor to temporal pattern perception (Ellis & Jones, 2009). This raises the possibility that the perception of temporal patterns could arise, at least in part by salience at each instantaneous point, and the temporal grouping of repeated points of salience. Recurring salience could provide a basis for adapting to and eventually encoding temporal patterns. This encoding could occur in the absence of attention, or resonance could eventually engage dynamic, directed attention to points in time once a temporal pattern is recognized (Jones & Boltz, 1989). Response to salience could, therefore, provide the origins for dynamic attending. This could have occurred to some extent in the present experiment. It is possible that salience perception was partially mediated by participants' attention to, or prediction of, patterns in the music. Even so, it is unlikely that attention or prediction accounts for the relationship between salience and involuntary motor responses, because temporal regularities and harmonic context within the musical excerpts did not significantly correlate with salience or involuntary motor responses. Our findings suggest that the relationship between salience responses and involuntary muscle responses most likely reflects pre-attentive process, even while salience could have additionally evoked or modulated higher-level cognitive processes simultaneously.

Harmonic change corresponded to perceived salience but did not correspond with involuntary motor responses. Features of harmony have been shown to contribute to reported and physiological emotional responses to music. Unlike loudness which tends to engage emotional arousal (Dean et al., 2011), harmony appears to modulate emotional valence. Consonant music tends to evoke pleasant feelings while dissonant music tends to evoke unpleasant feelings as well as paralimbic brain regions (Blood, Zatorre, Bermudez, & Evans, 1999; Gomez & Danuser, 2007). Our results suggest that inharmonicity, or instantaneous dissonance, drives both salience perception and muscle responses during music listening, but harmonic change is only related to salience perception. One explanation for this result may be that harmonic context drives top-down appraisal of emotional valence but not involuntary motor responses. This result further attests to the sensitivity of involuntary auditory-motor transmission to instantaneous acoustic features.

We demonstrated that several of the acoustic features that corresponded to perceived salience (amplitude, intensity, spectral centroid, spectral crest, and inharmonicity) also

corresponded to involuntary responses of the peripheral motor system. An involuntary innervation of muscles by a salient acoustic event points to a potential mechanism for sensorimotor coupling that is distinct from voluntary entrainment and the effortful alignment of movements in time with an external temporal pattern (Morillon & Baillet, 2017; Morillon et al., 2014). Motor responses to patterns of salience over an extended period of time may give rise to a continuous, dynamic motor and auditory response correspondence, or resonance. Resonance theory describes how two intrinsically oscillating mechanical systems, such as a metronome or a swinging pendulum, will eventually spontaneously synchronize if there is a transfer of energy from one system to the other (Iwasaki & Zheng, 2006). Resonance behavior has been simulated mathematically and demonstrated in both biological (Ott & Antonsen Jr, 2017) and non-living (Iwasaki & Zheng, 2006) systems, and is hypothesized to be a potential means for the alignment of external stimuli with internal neural oscillations. We here provide a potential basis for a mechanism that enables resonance between acoustic salience fluctuations and motor response fluctuations. The transfer of energy that synchronizes external acoustic events with internal neural fluctuations may occur through a brainstem-spinal pathway mediating orienting reflexes, or a brainstem-cerebellar-spinal pathway that quickly adjusts actions following significant stimulus changes (Brown et al., 1991a; Davis et al., 1982; Schwartz & Kotz, 2013; Sokolov, 1963). We hypothesize that the physical transfer of energy from acoustic events to motor responses occurs through the fast transmission of signals from the auditory nerve to the motor periphery, via the brainstem (Brown et al., 1991b) and/or cerebellum (Schwartz & Kotz, 2013), rather than via a brainstem-cortical-spinal loop. Over time, this unidirectional recurrent transmission could theoretically give rise to a temporal alignment of peripheral motor activity fluctuations with the external acoustic events. This temporal alignment could appear as if motor responses are synchronized or entrained with acoustic events, yet in this type of resonant system there would be no foreknowledge of when the next event should occur. Thus, this hypothetical neural pathway for resonance is assumed to be completely reactive. On this view, the motor system may not only function as a regularity detector (Morillon et al., 2015), but could respond involuntarily to acoustic irregularities and thereby lock on to external stimuli that are not predictable. Such a system could also provide the basis for developing motor routines that dictate voluntary entrainment, and such a system could provide a means for adapting or updating existing motor routines. In other words, the salient acoustic features that evoke motor responses may eventually give rise to voluntary movements to music from tapping your feet through to dancing. We suggest further work to replicate and extend the current results, firstly by strengthening the causal relationship between salient acoustic events and involuntary motor responses, for instance, by creating artificial musical stimuli in which particular acoustic features are manipulated independently. Additional studies can also corroborate whether involuntary audio-motor innervation occurs outside of attention, for instance by requiring participants to perform an additional, attention-demanding task. More work is also

needed to clarify the neural pathways involved using auditory brainstem recording and high-field imaging.

The results reported here suggest a more complete understanding of the link between auditory and motor processing. Ultimately, an optimized motor system must be able to anticipate and react. Current perspectives of sensorimotor integration focus on the capacity for the brain to predict external events or anticipate outcomes from actions, allowing the organism to behave in a fluent, proactive way, aligning optimally with upcoming events (Morillon et al., 2015; Wolpert et al., 1995). This framework offers a way to understand sensorimotor integration that requires prior knowledge, whether it is an action plan, a learned stimulus-response association, or a sequential/temporal pattern that has been encoded into working memory. However, it remains equally important to account for the role of purely stimulus-driven sensorimotor integration to understand the full range of our sensorimotor capacities. An appropriate memory trace or motor routine is not always available or useful for every possible environmental condition we may encounter. Motor routines that are not based on intrinsic properties of the motor system have to be acquired through experience. When we cannot predict, we must react. The present results demonstrate a mechanism that links involuntary motor responses with acoustic events embedded within a complex and continuous auditory stream, in a way that does not depend on an acquired memory trace or motor program. These results provide an initial step toward closing the gap between memory-dependent sensorimotor integration and the acquisition of those memory traces.

Based on these results we suggest that further work be directed toward understanding how voluntary and involuntary sensorimotor integration work together, rather than examine each in isolation. In order to learn and extract meaningful patterns in the environment, we should be able to extract initially unpredictable information before we can detect predictable patterns. Stimulus-driven and memory-driven sensorimotor integration may enable complementary functions. A stimulus-driven auditory-motor system could provide the starting point for acquiring sensorimotor memory traces and thereby provide the basis for memory-driven sensorimotor integration. It may also be that stimulus-driven and memory-driven systems interact such that, for instance, acquired motor routines are involuntarily updated or adjusted (Valls-Solé et al., 2008). Three types of tasks may make use of both stimulus-driven and memory-driven sensorimotor integration, either in a complementary or interactive way: implicit learning, attention orienting, and memory updating (e.g., “relearning”). Each of these tasks may be a fruitful context for investigating how stimulus-driven and memory-driven sensorimotor integration may coordinate. In implicit learning tasks such as the serial reaction time task and classical conditioning tasks, involuntary motor response to auditory salience may be a precursor to learning, by enabling adaptation to external stimuli via resonance. Prolonged resonance may then give rise to an implicit memory trace. In addition, auditory-motor resonance could also facilitate attention orienting to salient events. A shift from

resonance to predictive or voluntary action could result from a gradual shifting of attention through repetition, a sudden capture of attention by one or more salient events, or by a shift in task-relevance of external events. Finally, auditory-motor resonance could provide a means for updating or adapting existing memory traces through error-correction. Potential neural bases could be gradual or sudden synchrony between subcortical and cortical neural populations, such as those that are thought to originate from the thalamus (Saalmann, 2014). Thalamo-cortical pathways such as cerebellar-motor or cerebellar-prefrontal pathways would be suitable for adapting to salient auditory events and signaling cortical regions to encode motor sequences, engage attention, or update memory traces. These questions could be addressed using a combination of behavioral, electrophysiological, and high field imaging approaches to determine how subcortical-cortical signaling and neural synchrony change with implicit learning, attention orienting, or memory updating.

Finally, a system for involuntary audio-motor integration provides an alternate route for motor training and potentially helping to address movement disorders. While training regimes that introduce a predictable temporal structure have proven helpful for walking movements in Parkinson’s disease (Benoit et al., 2014; Dalla Bella et al., 2015), our results suggest that unstructured, salient cues can also promote motor responses. A fast, subcortical pathway for audio-motor transmission may provide the means to stimulate non-repetitive, ballistic movements such as reaching, or through subcortico-cortico communication described above, facilitate the learning of repetitive movements. We further demonstrate that specific acoustic features promote motor responses and influence subjective salience perception. This evidence provides a direction for designing an acoustic environment for training as well as a means for tailoring that environment to individual differences in subjective salience perception. This versatile approach has the potential to innovate training programs either on its own or in combination with existing protocols.

4.1. Limitations

Due to non-linear temporal relationships between human ratings of perceived salience, acoustic features, and sEMG responses, we used DTW. This approach has previously been applied to EEG data to examine its relationship to non-periodic human movement (Chemin, Huang, Mulders, & Mouraux, 2018). To ensure our analysis was not too liberal and reduce the chance of a type I error, we used a conservative baseline using jackknifing techniques that ensured cross-correlation coefficients represented idiosyncratic relationships for each stimulus. To the knowledge of the authors, there are no continuous objective measures of auditory salience perception that have linear temporal relationships with auditory stimuli. If such measures were discovered, it would be interesting to examine the relationship between salience perception and acoustic features without using DTW. This would allow us to examine the temporal delay in perceiving salience and how it changes for various acoustic features, and

determine if the delays are short enough to suggest pre-attentive responses.

5. Conclusion

We provide evidence for auditory-motor resonance, an automatic innervation of peripheral motor activity by acoustic events within a complex stream that then leads to a temporal alignment of acoustic and motor fluctuations. As a purely stimulus-driven mechanism that links external events with internal motor activity, resonance closes the theoretical gap between voluntary, memory-driven behaviors, and the origins of those behaviors. Resonance may also play an important role in voluntary movement in general, as a starting point for goal-driven behavior to develop, in cases where an appropriate memory trace or motor routine is not yet available.

CRediT author statement

Benjamin G. Schultz: Conceptualization, Formal analysis, Methodology, Software, Data curation, Project administration, Resources, Visualization, Writing - Original draft of Methods and Results sections, Reviewing and Editing. **Rachel Brown:** Writing - Original draft of Introduction and Discussion sections, Reviewing and Editing. **Sonja Kotz:** Conceptualization, Supervision, Resources, Reviewing, and Editing.

Open practices

The study in this article earned an Open Data badge for transparent practices. Materials and data for the study are available at <https://doi.org/10.34894/JQAEIR>.

Acknowledgments

This work was supported by funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 707865 to R.M.B. and S.A.K.

Study data, digital study materials, and analysis code is publicly available upon reasonable at <https://doi.org/10.34894/JQAEIR>. The archived audio files contain copyrighted materials and will be shared unconditionally on request by contacting: datamanagement-fpn@maastrichtuniversity.nl.

No part of the study procedures or analyses was pre-registered prior to the research being conducted.

REFERENCES

Bendixen, A., SanMiguel, I., & Schröger, E. (2012). Early electrophysiological indicators for predictive processing in audition: A review. *International Journal of*

- Psychophysiology*, 83(2), 120–131. <https://doi.org/10.1016/j.ijpsycho.2011.08.003>
- Benoit, C.-E., Dalla Bella, S., Farrugia, N., Obrig, H., Mainka, S., & Kotz, S. A. (2014). Musically cued gait-training improves both perceptual and motor timing in Parkinson's disease. *Frontiers in Human Neuroscience*, 8(July), 494. <https://doi.org/10.3389/fnhum.2014.00494>
- Bishop, L., Bailes, F., & Dean, R. T. (2013). Musical expertise and the ability to imagine loudness. *PLoS One*, 8(2), Article e56052. <https://doi.org/10.1371/journal.pone.0056052>
- Blakemore, S. J., Frith, C. D., & Wolpert, D. M. (1999). Spatio-temporal prediction modulates the perception of self-produced stimuli. *Journal of Cognitive Neuroscience*, 11(5), 551–559. <https://doi.org/10.1162/089892999563607>
- Blood, A. J., Zatorre, R. J., Bermudez, P., & Evans, A. C. (1999). Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nature Neuroscience*, 2(4), 382–387. <https://doi.org/10.1038/7299>
- Böck, S., & Schedl, M. (2011). Enhanced beat tracking with context-aware neural networks. In *Proc. of the 14th Int. Conference on Digital Audio Effects* (pp. 1–5).
- Borji, A., Sihite, D. N., & Itti, L. (2013). What stands out in a scene? A study of human explicit saliency judgment. *Vision Research*, 91, 62–77. <https://doi.org/10.1016/j.visres.2013.07.016>
- Bradley, M., & Lang, P. (2000). Affective reactions to acoustic stimuli. *Psychophysiology*, 37, 204–215. Retrieved from [internal-pdf://1469-8986-2590013302/1469-8986.pdf](https://doi.org/10.1162/089892999563607).
- Brown, P., Rothwell, J. C., Thompson, P. D., Britton, T. C., Day, B. L., & Marsden, C. D. (1991a). New observations on the normal auditory startle reflex in man. *Brain: a Journal of Neurology*, 114(4), 1891–1902. <https://doi.org/10.1093/brain/114.4.1891>
- Brown, P., Rothwell, J. C., Thompson, P. D., Britton, T. C., Day, B. L., & Marsden, C. D. (1991b). The hyperreflexias and their relationship to the normal startle reflex. *Brain: a Journal of Neurology*, 114(4), 1903–1928. <https://doi.org/10.1093/brain/114.4.1903>
- Burger, B., Thompson, M. R., Luck, G., Saarikallio, S. H., & Toiviainen, P. (2014). Hunting for the beat in the body: On period and phase locking in music-induced movement. *Frontiers in Human Neuroscience*, 8(November), 903. <https://doi.org/10.3389/fnhum.2014.00903>
- Caclin, A., Brattico, E., Tervaniemi, M., Näätänen, R., Morlet, D., Giard, M.-H., et al. (2006). Separate neural processing of timbre dimensions in auditory sensory memory. *Journal of Cognitive Neuroscience*, 18(12), 1959–1972. <https://doi.org/10.1162/jocn.2006.18.12.1959>
- Caclin, A., McAdams, S., Smith, B. K., & Winsberg, S. (2005). Acoustic correlates of timbre space dimensions: A confirmatory study using synthetic tones. *The Journal of the Acoustical Society of America*, 118(1), 471–482. <https://doi.org/10.1121/1.1929229>
- Chemin, B., Huang, G., Mulders, D., & Mouraux, A. (2018). EEG time-warping to study non-strictly-periodic EEG signals related to the production of rhythmic movements. *Journal of Neuroscience Methods*, 308, 106–115.
- Chon, S. H., & McAdams, S. (2012). Investigation of timbre saliency, the attention-capturing quality of timbre. *The Journal of the Acoustical Society of America*, 131(August), 3433. <https://doi.org/10.1121/1.4708879>
- Dalla Bella, S., Benoit, C.-E., Farrugia, N., Schwartz, M., & Kotz, S. A. (2015). Effects of musically cued gait training in Parkinson's disease: Beyond a motor benefit. *Annals of the New York Academy of Sciences*, 1337(1), 77–85. <https://doi.org/10.1111/nyas.12651>
- Davis, M., Gendelman, D. S., Tischler, M. D., & Gendelman, P. M. (1982). A primary acoustic startle circuit: Lesion and stimulation studies. *The Journal of Neuroscience: The Official*

- Journal of the Society for Neuroscience*, 2(6), 791–805. <https://doi.org/7086484>.
- Dean, R. T., Bailes, F., & Schubert, E. (2011). Acoustic intensity causes perceived changes in arousal levels in music: An experimental investigation. *PLoS One*, 6(4), 1–8. <https://doi.org/10.1371/journal.pone.0018591>
- Ellis, R. J., & Jones, M. R. (2009). The role of accent salience and joint accent structure in meter perception. *Journal of Experimental Psychology: Human Perception and Performance*, 35(1), 264–280. <https://doi.org/10.1037/a0013482>
- Essens, P. J., & Povel, D.-J. (1985). Metrical and nonmetrical representations of temporal patterns. *Perception & Psychophysics*, 37(1), 1–7. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/3991313>.
- Fernandez-Del-Olmo, M., Bello, O., Lopez-Alonso, V., Marquez, G., Sanchez, J. A., Morenilla, L., et al. (2013). The effects of startle and non-startle auditory stimuli on wrist flexion movement in Parkinson's disease. *Neuroscience Letters*, 548(June 2013), 56–60. <https://doi.org/10.1016/j.neulet.2013.05.069>
- Gomez, P., & Danuser, B. (2007). Relationships between musical structure and psychophysiological measures of emotion. *Emotion*, 7(2), 377–387. <https://doi.org/10.1037/1528-3542.7.2.377>
- Goodale, M. A., & Murison, R. C. C. (1975). The effects of lesions of the superior colliculus on locomotor orientation and the orienting reflex in the rat. *Brain Research*, 88(2), 243–261. [https://doi.org/10.1016/0006-8993\(75\)90388-1](https://doi.org/10.1016/0006-8993(75)90388-1)
- Gottlieb, J. P., Kusunoki, M., & Goldberg, M. E. (1998). The representation of visual salience in monkey parietal cortex. *Nature*, 391(6666), 481–484. <https://doi.org/10.1038/35135>
- Hurley, B. K., Martens, P. A., & Janata, P. (2014). Spontaneous sensorimotor coupling with multipart music. *Journal of Experimental Psychology: Human Perception and Performance*, 40(4), 1679–1696. <https://doi.org/10.1037/a0037154>
- Itti, L. (2006). Quantitative modelling of perceptual salience at human eye position. *Visual Cognition*, 14(4–8), 959–984. <https://doi.org/10.1080/13506280500195672>
- Iwasaki, T., & Zheng, M. (2006). Sensory feedback mechanism underlying entrainment of central pattern generator to mechanical resonance. *Biological Cybernetics*, 94(4), 245–261. <https://doi.org/10.1007/s00422-005-0047-3>
- Janata, P., Tomic, S. T., & Haberman, J. M. (2012). Sensorimotor coupling in music and the psychology of the groove. *Journal of Experimental Psychology: General*, 141(1), 54–75. <https://doi.org/10.1037/a0024208>
- Johnson, L., & Lubin, A. (1967). The orienting reflex during waking and sleeping. *Electroencephalography and Clinical Neurophysiology*, 1962, 11–21. [https://doi.org/10.1016/0013-4694\(67\)90004-1](https://doi.org/10.1016/0013-4694(67)90004-1)
- Jones, M. R. (2001). Temporal expectancies, capture, and timing in auditory sequences. *Advances in Psychology*, 133, 191–229. [https://doi.org/10.1016/S0166-4115\(01\)80011-X](https://doi.org/10.1016/S0166-4115(01)80011-X)
- Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review*, 96(3), 459–491. <https://doi.org/10.1037/0033-295X.96.3.459>
- Levitin, D. J., Nuzzo, R. L., Vines, B. W., & Ramsay, J. O. (2007). Introduction to functional data analysis. *Canadian Psychology/Psychologie Canadienne*, 48(3), 135–155. <https://doi.org/10.1037/cp2007014>
- McAdams, S., Winsberg, S., Donnadieu, S., De Soete, G., & Krimphoff, J. (1995). Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes. *Psychological Research*, 58(3), 177–192. <https://doi.org/10.1007/BF00419633>
- Morillon, B., & Baillet, S. (2017). Motor origin of temporal predictions in auditory attention. *Proceedings of the National Academy of Sciences*, 114(42), E8913–E8921. <https://doi.org/10.1073/pnas.1705373114>, 201705373.
- Morillon, B., Hackett, T. A., Kajikawa, Y., & Schroeder, C. E. (2015). Predictive motor control of sensory dynamics in auditory active sensing. *Current Opinion in Neurobiology*, 31, 230–238. <https://doi.org/10.1016/j.conb.2014.12.005>
- Morillon, B., & Schroeder, C. E. (2015). Neuronal oscillations as a mechanistic substrate of auditory temporal prediction. *Annals of the New York Academy of Sciences*, 1337(1), 26–31. <https://doi.org/10.1111/nyas.12629>
- Morillon, B., Schroeder, C. E., & Wyart, V. (2014). Motor contributions to the temporal precision of auditory attention. *Nature Communications*, 5, 5255. <https://doi.org/10.1038/ncomms6255>
- Morillon, B., Schroeder, C. E., Wyart, V., & Arnal, L. H. (2016). Temporal prediction in lieu of periodic stimulation. *Journal of Neuroscience*, 36(8), 2342–2347. <https://doi.org/10.1523/JNEUROSCI.0836-15.2016>
- Olsen, K. N., & Stevens, C. J. (2013). Psychophysiological response to acoustic intensity change in a musical chord. *Journal of Psychophysiology*, 23, 16–26.
- Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, 42(1), 107–123. [https://doi.org/10.1016/S0042-6989\(01\)00250-4](https://doi.org/10.1016/S0042-6989(01)00250-4)
- Rodger, M. W. M., & Craig, C. M. (2016). Beyond the metronome: Auditory events and music may afford more than just interval durations as gait cues in Parkinson's disease. *Frontiers in Neuroscience*, 10(JUN), 1–5. <https://doi.org/10.3389/fnins.2016.00272>
- Rothermich, K., Schmidt-Kassow, M., & Kotz, S. A. (2012). Rhythm's gonna get you: Regular meter facilitates semantic sentence processing. *Neuropsychologia*, 50(2), 232–244. <https://doi.org/10.1016/j.neuropsychologia.2011.10.025>
- Saalman, Y. B. (2014). Intralaminar and medial thalamic influence on cortical synchrony, information transmission and cognition. *Frontiers in Systems Neuroscience*, 8(May), 1–8. <https://doi.org/10.3389/fnsys.2014.00083>
- Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, 32(1), 9–18. <https://doi.org/10.1016/j.tins.2008.09.012>
- Schroeder, C. E., Lakatos, P., Kajikawa, Y., Partan, S., & Puce, A. (2008). Neuronal oscillations and visual amplification of speech. *Trends in Cognitive Sciences*, 12(3), 106–113. <https://doi.org/10.1016/j.tics.2008.01.002>
- Schroeder, C. E., Wilson, D. A., Radman, T., Scharfman, H., & Lakatos, P. (2010). Dynamics of active sensing and perceptual selection. *Current Opinion in Neurobiology*, 20(2), 172–176. <https://doi.org/10.1016/j.conb.2010.02.010>
- Schröger, E. (1997). On the detection of auditory deviations: A pre-attentive activation model. *Psychophysiology*, 34(3), 245–257. <https://doi.org/10.1111/j.1469-8986.1997.tb02395.x>
- Schröger, E., & Winkler, I. (1995). Presentation rate and magnitude of stimulus deviance effects on human pre-attentive change detection. *Neuroscience Letters*, 193(3), 185–188. [https://doi.org/10.1016/0304-3940\(95\)11696-T](https://doi.org/10.1016/0304-3940(95)11696-T)
- Schwartz, M., & Kotz, S. A. (2013). A dual-pathway neural architecture for specific temporal prediction. *Neuroscience and Biobehavioral Reviews*, 37(10), 2587–2596. <https://doi.org/10.1016/j.neubiorev.2013.08.005>
- Schwartz, M., & Kotz, S. A. (2016). Contributions of cerebellar event-based temporal processing and preparatory function to speech perception. *Brain and Language*, 161, 28–32. <https://doi.org/10.1016/j.bandl.2015.08.005>
- Sokolov, E. N. (1963). Higher nervous functions: The orienting reflex. *Annual Review of Physiology*, 25(1), 545–580. <https://doi.org/10.1146/annurev.ph.25.030163.002553>
- Sokolov, A. A., Miall, R. C., & Ivry, R. B. (2017). The cerebellum: Adaptive prediction for movement and cognition. *Trends in*

- Cognitive Sciences*, 21(5), 313–332. <https://doi.org/10.1016/j.tics.2017.02.005>
- Tajadura-Jiménez, A., Väljamäe, A., Asutay, E., & Västfjäll, D. (2010). Embodied auditory perception: The emotional impact of approaching and receding sound sources. *Emotion*, 10(2), 216–229. <https://doi.org/10.1037/a0018422>
- Thompson, K. G., & Bichot, N. P. (2004). A visual salience map in the primate frontal eye field. *Progress in Brain Research*. [https://doi.org/10.1016/S0079-6123\(04\)47019-8](https://doi.org/10.1016/S0079-6123(04)47019-8)
- Valls-Solé, J., Kumru, H., & Kofler, M. (2008). Interaction between startle and voluntary reactions in humans. *Experimental Brain Research*, 187(4), 497–507. <https://doi.org/10.1007/s00221-008-1402-0>
- Wolpert, D. M., Ghahramani, Z., & Jordan, M. (1995). An internal model for sensorimotor integration. *Science*, 269(5232), 1880–1882. <https://doi.org/10.1126/science.7569931>
- Repp, B. H., & Su, Y. H. (2013). Sensorimotor synchronization: a review of recent research (2006–2012). *Psychonomic Bulletin & Review*, 20(3), 403–452.
- Repp, B. H. (2005). Sensorimotor synchronization: a review of the tapping literature. *Psychonomic Bulletin & Review*, 12(6), 969–992.
- Schultz, B. G., O'Brien, I., Phillips, N., McFarland, D. H., Titone, D., & Palmer, C. (2016). Speech rates converge in scripted turn-taking conversations. *Applied Psycholinguistics*, 37(5), 1201–1220.
- Ott, E., & Antonsen Jr, T. M. (2017). Frequency and phase synchronization in large groups: Low dimensional description of synchronized clapping, firefly flashing, and cricket chirping. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 27(5), 051101.
- Jeffreys, H. (1961). *Theory of probability* (3rd ed.). New York: Oxford University Press.