

Brain potentials predict learning, transmission and modification of an artificial symbolic system

Massimo Lumaca^{1,*} and Giosuè Baggio²

¹SISSA International School for Advanced Studies, Trieste and ²Language Acquisition and Language Processing Lab Department of Language and Literature, Norwegian University of Science and Technology, Trondheim

*Correspondence should be addressed to Massimo Lumaca, SISSA International School for Advanced Studies, via Bonomea, 265 - 34136 Trieste. E-mail: mlumaca@sissa.it.

Abstract

It has recently been argued that symbolic systems evolve while they are being transmitted across generations of learners, gradually adapting to the relevant brain structures and processes. In the context of this hypothesis, little is known on whether individual differences in neural processing capacity account for aspects of ‘variation’ observed in symbolic behavior and symbolic systems. We addressed this issue in the domain of auditory processing. We conducted a combined behavioral and EEG study on 2 successive days. On day 1, participants listened to standard and deviant five-tone sequences: as in previous oddball studies, a mismatch negativity (MMN) was elicited by deviant tones. On day 2, participants learned an artificial signaling system from a trained confederate of the experimenters in a coordination game in which five-tone sequences were associated to affective meanings (emotion-laden pictures of human faces). In a subsequent game with identical structure, participants transmitted and occasionally changed the signaling system learned during the first game. The MMN latency from day 1 predicted learning, transmission and structural modification of signaling systems on day 2. Our study introduces neurophysiological methods into research on cultural transmission and evolution, and relates aspects of variation in symbolic systems to individual differences in neural information processing.

Key words: cultural transmission; neural predictors; MMN; auditory processing; signaling games

Introduction

Human symbolic systems such as language and music evolve while they are being acquired and transmitted by learners from subsequent generations. It has recently been hypothesized that the brain co-determines the direction of cultural evolution: symbol systems would gradually adapt to the constraints imposed by the relevant neural circuits, becoming easier to learn or process by inheriting specific structural and semantic properties (Dehaene and Cohen 2007; Christiansen and Chater, 2008). This view has found some support in research on language (Kirby *et al.*, 2008), arithmetic and reading (Dehaene *et al.*, 2015; Hannagan *et al.*, 2015), where it has been invoked to account for universal aspects of such systems. However, little is known about ‘variation’: would subtle yet possibly critical differences in brain structure and function across individuals

account for some of the forms of variation (e.g. across individuals or over time) observed in cultural symbolic systems?

Here we provide the first experimental evidence that symbol systems reflect individual cognitive processing capabilities when they are culturally learned and transmitted. We set out to test whether individual differences in neural information processing, as revealed by event-related potentials (ERPs), would predict how individuals learn, transmit and modify an artificial symbolic system. Our methodology involved two elements: (i) a laboratory model of cultural transmission, and (ii) a neural marker of processing efficiency for structured sequences.

We used signaling games as a model of cultural learning and transmission (Lewis, 1969; Skyrms, 2010; Moreno and Baggio, 2015; Nowak and Baggio, 2016). In signaling games, the goal for the sender and receiver is to converge on a mapping of signals to meanings via iterated signaling rounds (see ‘Methods’

Received: 29 March 2016; Revised: 28 May 2016; Accepted: 3 August 2016

© The Author (2016). Published by Oxford University Press. For Permissions, please email: journals.permissions@oup.com

section). Our participants played two signaling games: first, as receivers, they learned a mapping of signals (five-tone sequences) to meanings (simple and compound emotions); then, as senders, they transmitted the system learned in the first game (Figure 2). Artificial languages are generally assumed to be relevant to understand natural languages. Similarly, artificial tone systems can be used to study actual musical systems. Moreover, music may be seen as a cultural ‘symbolic’ system (Cross, 2005): it has often been used throughout its history to convey cognitive or affective meaning (Juslin and Sloboda, 2001). Our tone sequences were indeed endowed with affective meanings, referring to facial expressions of emotions: a link between facial expression and music-evoked emotions is supported by neural and behavioral research (Hsieh et al., 2012; Lense et al., 2014; Palmer et al., 2013).

As a neural marker of auditory processing we used the mismatch negativity (MMN). The MMN is an ERP component peaking between 100 and 220 ms from the onset of a deviant event disrupting a standard sequence (Näätänen et al., 1978). Its elicitation therefore depends on whether a memory trace for the standard sequence has been formed (Alho et al., 1990). Its latency and amplitude are correlated with pattern discrimination abilities in individuals (Näätänen et al., 1993). The MMN has been seen as a marker of processing efficiency (Tervaniemi 2001; Gottselig et al., 2004), and as a means to investigate individual differences in auditory perception (Tervaniemi et al., 1997).

To test whether individual differences in auditory processing, as measured by MMN peak latencies and amplitudes, predict learning, transmission and structural modification of signaling systems, we designed two experimental sessions on 2 successive days: on day 1, participants underwent standard auditory MMN acquisition; on day 2, they played first as receivers (learners) then as senders (transmitters) in two successive signaling games (see ‘Study design’ section). The MMN stimuli used on day 1 and the signals used on day 2 were five-tone sequences. If MMN amplitudes or latencies predict learning, transmission and reorganization of signaling systems, this would suggest a direct link between individual differences in neural processing and variation in cultural symbolic systems.

Methods

Participants

Twenty-one right-handed native speakers of Italian (9 female; mean age 23.71) with no hearing or neurological disorders participated in the study: 3 (1 female) had an excess of movement artifacts in the EEG data; the remaining 18 were included in the final data analysis (mean age 23.71). To reduce expertise effects on auditory processing (Fujioka et al., 2004; Seppänen et al., 2007), we included only musically naïve individuals.

Study design

Each subject participated in two sessions on 2 consecutive days (24 h apart). The EEG session took place on day 1. At the start of the session, we assessed participants’ working memory with a digit span test (Orsini et al., 1987) and a complex span tests (Foster et al., 2015). We recorded EEG data while tone sequences were delivered in three blocks with high-, low- and 0-entropy sequences (see ‘Stimuli’ section) using loudspeakers (80 dB). During EEG recording, participants watched a silent subtitled movie in a 8 × 11 cm frame in the center of the screen. They

were informed there would be questions about the movie at the end of the session, in order to divert their attention from the stimuli. On day 2, the participant played two signaling games, first as receiver (Game 1), then as sender (Game 2) (see ‘Transmission design’ section; Figure 2), with a confederate of the experimenters; sender and receiver roles were fixed. In Game 1, the participant (receiver) learned a signaling system where the signals were five 5-tone sequences and the meanings were three basic and two compound emotions. In Game 2, the participant (sender) had to transmit the signaling system learned in Game 1, as they recalled it, to the receiver. The signaling games took place in a quiet room with two screens aligned back-to-back, with a curtain between them, preventing players from looking at or communicating with each other. Written instructions were given to the participant before the beginning of each session (Supplementary Materials sections 1.1.1 and 1.1.2). Players used a full-size keyboard and headphones.

Bohlen-Pierce scale

Sequences were built using tones of the Bohlen-Pierce scale (Mathews et al., 1988). The equal tempered Bohlen-Pierce scale is based on 13 logarithmically even divisions of a tritone (3:1 frequency ratio). The frequency (F) of each note is: $F = k * 3^{(n/13)}$; k is the reference pitch frequency; n is the number of steps on the scale. This scale was used instead of the Western diatonic scale to prevent participants from exploiting prior intuitions about musical harmony or melody, which may affect perception of tone sequences. To avoid carry-over effects of the stimuli learned on day 1 (during the EEG session) on those used on day 2 in the signaling games, we set different n values to build the EEG stimuli and the materials used by the confederate in the signaling games (see below).

Day 1—EEG experiment

Stimuli. Tone sequences of different complexity (high-entropy, low-entropy and 0-entropy or control) were presented in three blocks. In each block, three types of tone sequences were used: standards (80% frequency), contour deviants (‘C-deviants’, 10%) and interval deviants (‘I-deviant’, 10%); for control stimuli, a single deviant was used (Figure 1). A session consisted of 8 sub-blocks with 250 stimuli each, and the control block included 4 sub-blocks. The order of the sessions (high-entropy, low-entropy and control) was counterbalanced across participants. The complexity of tone sequences was defined as the complexity of their contour transforms (‘contour entropy’) where ‘+’ stands for upward intervals, ‘-’ for downward intervals, and ‘=’ for repeated tones. We produced high-entropy standard stimuli (-+-+; Shannon Entropy = 1), low-entropy standard stimuli (+++-; SE = 0.81), and 0-entropy stimuli with repeated tones (====; SE = 0). In C-deviants, the fourth tone violated the surface structure of standard sequences, but not interval size; vice versa for I-deviants on the fifth position. The deviant was always the same as the first tone in a sequence (Tervaniemi et al., 2001). Each sequence consisted of five 50-ms sinusoidal tones from the frequencies 440, 521, 567, 617, 730.6 Hz (low register), separated by 50-ms of silence (interstimulus interval, ISI 750 ms). Sequences were transposed over trials at three registers of the Bohlen-Pierce scale (lowest frequencies: 440, 478, 567 Hz).

Data acquisition. The EEG was acquired using a Biosemi (Amsterdam, The Netherlands) ActiveTwo system with 128 active electrodes, digitized at 1024 Hz, off-line downsampled to 500 Hz and bandpass filtered (0.1–30 Hz). The EEG was

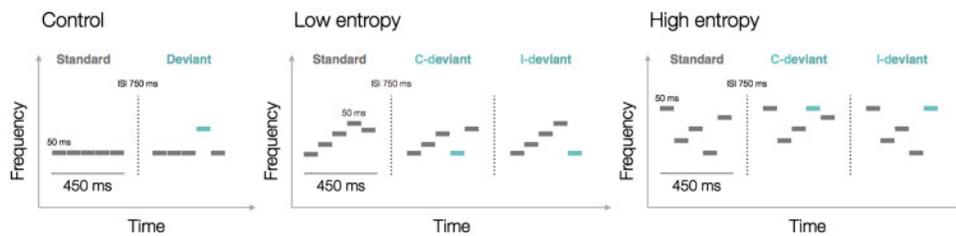


Fig. 1. Schematic illustration of the tone sequences presented to participants on day 1 of the experiment. The EEG was recorded while participants listened to these sequences. Each sequence consisted of five tones with 50 ms duration and 50 ms inter-tone intervals, i.e. 450 ms sequence duration. The ISI (the interval between the offset of the last tone in one sequence and the onset of the first tone in the next sequence) was 750 ms. Standard sequences of different complexity were used: control (monotone) sequences, low- and high-entropy sequences. Deviants were constructed such that the fourth tone violated the melodic contour of the standard sequence ('C-deviant') or the fifth tone violated the interval between the fourth and fifth tones in standard sequences while preserving melodic contour ('I-deviant'). The proportion of standards to deviants was 4:1. A blocked design was used, where the blocks were randomised across participants.

referenced on-line to the average of signals from the left and right mastoid channels. The EOG was recorded by two electrodes above and below the right eye, and by two electrodes placed at the outer canthus of each eye. Data analysis was performed in ERPLab (Lopez-Calderon and Luck, 2014) and EEGlab (Delorme and Makeig, 2004). Epochs were time-locked to the onset of the critical (fourth or fifth) tones for the 'standard' and 'deviant' stimuli, in a time window from -200 (baseline) to 800 ms after tone onset. Pre-processing included rejection of trials with ocular artifacts using a 150 ms step-function for the bipolar horizontal and vertical EOG channels (threshold at $35 \mu\text{V}$; step 10 ms). A 200 ms-wide moving window was slid across the data in all 128 channels: epochs where the peak-to-peak offset exceeded $100 \mu\text{V}$ in any window were rejected. The remaining epochs were averaged separately for each condition (deviant or standard). Mean peak latencies and amplitudes of individual difference waves were calculated in each of four scalp quadrants (posterior, right, anterior, left; 32 channels each), and in each of four 120-ms time windows (0–480 ms after tone onset), selected according to the timing characteristics of the MMN (100–220 ms from deviant tone onset; Näätänen et al., 1978; Paavilainen, 2013) and after inspection of ERP data.

Data analysis. A MMN is as an enhanced negative ERP component in response to deviants relative to standard stimuli, larger over frontal sites at 100–220 ms from stimulus onset (Näätänen et al., 1978; Paavilainen, 2013). We used a four-way ANOVA with mean MMN amplitude as dependent variable in each of three conditions, and four factors: Stimulus Type (standard or deviant), Temporal Window (four levels, as above), Quadrant (four levels, as above) and, nested on the latter, Electrode (32×4 levels), using the Greenhouse-Geisser correction when appropriate. Wilcoxon signed-rank tests were used in *post-hoc* analyses.

To investigate whether individual differences in auditory processing predict learning, transmission and structural modification of signaling systems, we first measured MMN amplitudes and peak latencies in individual difference waves (deviant minus standard stimuli). Next, we computed correlations between MMN peak latencies (or mean amplitudes) in each participant with various measures of behavior in signaling games (see below). *P*-values from the correlation tests were Bonferroni corrected at $=0.05/12 = 0.004$; 12 is the number of times each independent variable (amplitude or peak latency) is correlated with behavioral measures.

Day 2—signaling games

Signaling games. Signaling games are games of incomplete information in which a 'sender' and a 'receiver' develop a mapping

of signals to meanings via repeated interactions (Lewis, 1969; Skyrms, 2010). In the standard form, the sender privately witnesses an 'event' or 'state' and uses a 'signal' to inform the receiver on the identity of that state. Upon receiving the signal, the receiver takes appropriate 'action' to match the state's identity. Communication may succeed or fail depending on whether the state and action match or not. The players receive feedback as to whether the trial was successful. Moreno and Baggio (2015) showed that when sender and receiver roles are fixed within a game, the mapping is largely determined by the sender and learned by the receiver. In these conditions, signaling games are a viable laboratory model of social learning and cultural transmission of artificial symbol systems (for replications and extensions, see Nowak and Baggio, 2016). In our experiment 'signals' were constituted by five sine-wave tones (200 ms length; inter-tone interval 50 ms). In the Bohlen-Pierce equation, we defined $k = 440 \text{ Hz}$ and $n = 0, 4, 6, 7$ or 10 to obtain the maximum number of tone combinations with a low-integer frequency ratio. The 'states' were five 'basic' (joy, peace or sadness) or 'compound' emotions (peace + joy or peace + sadness), depicted as facial expressions. We recruited four professional actors (two male, two female) for a photo shoot session: they were instructed to express naturally each basic emotion; one photo per emotion per actor was included in the final stimulus set. Facial expressions for compound emotions were generated by dividing up expressions of basic emotions into two regions (Ekman and Friesen, 1975): the 'upper' and 'lower face'. Compound emotions were produced by merging the upper face portion of peace and the lower face portion of joy (peace + joy) or sadness (peace + sadness).

Transmission design. Figure 2 (right panel) shows the structure of signaling games. In Game 1, the participant played as receiver and the confederate played as sender. Before the start of the game, the confederate was trained to deliver one of three sets of stimuli. In each set, signals with different contour complexity were included: two high-entropy (Shannon entropy, Entropy = 1), two low-entropy (Entropy = 0.81), one monotone stimulus (control; Entropy = 0). Each set was randomly assigned to an equal number (7) of participants. We expected the participant to learn the associations between tone sequences and emotions consistently used by the confederate in Game 1. In Game 2, the participant played as sender and the confederate as receiver. The sender was asked to reproduce and transmit the system of tone sequences learned in the previous game as they remembered it.

Procedure. Figure 2 (left panel) shows a single trial in signaling games. The sender is privately shown a facial expression from an actor's set (e.g. joy; 3 s duration). Next, he is expected to generate an isochronous five-tone signal from a pool of five



Fig. 2. Example of a trial from the signaling games (left panel) played by participants on day 2 of the study. The top and bottom rows show what the sender and receiver saw on their screens, respectively. The task for the sender was to compose a five-tone sequence to be used as a signal of the simple or compound emotion expressed by the face presented on the screen at the start of the trial, and for the receiver to respond to that signal by choosing the face that the sender had seen. The sender and the receiver converged over trials on a shared mapping of signals (tone sequences) to meanings (emotions). Hand symbols indicate when the sender and the receiver had to produce a response. Feedback was given to both players simultaneously, displaying the face seen by the sender and the face selected by the receiver in a green frame (matching faces; correct) or in a red frame (mismatching faces; incorrect). Time flows from left to right. A diagram of the signaling games played by the participant is also shown (right panel). The participant played as receiver (R) with a confederate of the experimenters playing as sender (S) in Game 1, and roles switched in Game 2.

tones using a standard keyboard, to inform the receiver on the state the sender has seen. The relation between keyboard keys and tones was randomized over trials, so in each trial the sender had to learn anew the key-to-tone mappings.¹ Unheard by the receiver, the sender could try out each composition of the five-tone signal at will; the relation between keyboard keys and tones was fixed. Visual cues were provided on the screen showing the 'height' of each tone. The signal was sent to the receiver, who listened to it via headphones, and chose, using the numeric keypad, one of the five facial expressions. All five expressions were displayed simultaneously on the screen; their position was reshuffled randomly at each trial. Feedback (3 s duration) was given to both players simultaneously: the left-hand side of the screen showed the face seen by the sender, and the right-hand side showed the face chosen by the receiver. For positive feedback, both faces were displayed in a green frame, while for negative feedback a red frame was used. Participants had to play exactly 70 trials in Game 1 and 50 trials in Game 2. Each state (emotion) occurred equally frequently. There was no reward for speed in a trial or in the game: coordination was the only reward for participants.

Data analysis. Our aim was to determine whether MMN peak latencies or amplitudes predict individual performance and changes in the organisation of tone sequences in signaling games. The simplicity and formal explicitness of signaling games allows us to define a set of intuitive, rigorous measures of similarity between two tone sequences ('edit distance'), learning or recall ('coordination, transmission, innovation'), direction of transmission from sender to receiver or vice versa ('asymmetry'), and structural modification ('entropy' and 'proximity') (Supplementary Figure S2).

Signaling measures. Measures of learning and recall were computed using a measure of similarity, the edit distance (Hamming, 1986). The elements compared were the five tones ('tone distance') or interval directions ('ups' and 'downs') of adjacent tones ('contour distance'). A normalized output to [0 1] was given by S/L : S is the number of shared items and L the length of a sequence. Values close to 1 are associated to similar strings. Using the edit distance, we measured to what extent the signaling system transmitted by the confederate was

actually learned by the participant (receiver) in Game 1 ('coordination'), and to what extent the participant (sender) then reproduced ('transmission') or modified ('innovation') that system in Game 2. Formally, 'coordination' is the extent to which sender and receiver during Game 1 use the same mapping of states to signals (i.e. the extent to which they share a signaling system), and is defined as the mean similarity between the signal used by the sender for a given emotion and the set of signals mapped by the participant to the same emotion during the second half of Game 1, where player responses are more consistent and show fewer errors. It ranges from 1 (shared signaling system; the participant uses the same mapping of states to signals as the confederate, and has therefore learned that mapping) to 0 (different signals for a given emotion). 'Transmission' was measured as the contour similarity between pairs of tone sequences for a particular emotion produced with greater frequency by the senders in Game 1 (confederate) and Game 2 (participant) (between-player measure). Faithful transmission of signaling systems learned in Game 1 yields values close to 1. 'Innovation' is the edit contour distance between the signals acquired during Game 1 and transmitted by the participant during Game 2 denoting a particular emotion (within-player measure). Values close to 1 indicate that a significant number of changes were introduced in the signaling system by the participant. If coordination in Game 1 is imperfect (the signaling system transmitted by the confederate and the signaling system learned by the participant do not exactly correspond), transmission and innovation do not mirror each other. A measure of 'asymmetry' during coordination was given by subtracting the number of changes introduced in the signaling system by the sender from the number of changes made by the receiver, divided by the total number of changes (Moreno and Baggio, 2015); it ranges from 1 (the sender is modifying the signaling system more frequently) to -1 (viceversa).

Entropy. The Shannon Entropy (Shannon, 1948) of tone sequences (and of the relative contour transforms) was used as measure of the complexity of melodic surface structures:

$$H(X) = -\sum P(x_i) \log_2 P(x_i)$$

X is the string of 'ups' and 'downs' ('contour smoothness') or tones ('sequence complexity') and $P(x_i)$ is the probability to find item x_i in that string position. A single mean entropy value was

1 A partial exception here was the confederate input in Game 1, see Supplementary Materials section 1.2.1.

computed for the entire set of sequences, and separately for the relative contour transforms. The absolute difference in mean entropy values was computed between the two games and used in the correlation analysis with MMN latencies/amplitudes.

Proximity. Measures of perceptual proximity were applied to the sequences received and reproduced by the participant ('interval distribution', 'interval compression ratio'). The distribution of unison (0 macrotones), small (1 or 2 macrotones) and large intervals (3, 4, 6, 7 or 10 macrotones) was calculated for each set of melodies. The absolute difference in the frequency of unison, small or large intervals was computed between the two games and was used in the correlation analysis with MMN latencies or amplitudes. The interval compression ratio (Tierney et al., 2011) represents a bias in tone sequences toward smaller intervals. The interval compression ratio is defined as the mean absolute interval size (in macrotones) of a tone sequence randomly shuffled ($N = 100$), divided by the mean absolute interval size of the original sequence. Because a scrambled version of a proximal melody typically has larger intervals than the original (Von Hippel, 2000), larger values for this ratio are associated to compressed, small-interval melodies.

Results

Mismatch negativity

The time-course and topographical distribution of the effect elicited by deviant stimuli are typical of the MMN (Figures 3 and 4; Supplementary Figures S2–S4), with a frontal maximum between 120 and 240 ms from deviant onset. Statistical tests confirmed an MMN effect in all three conditions (Table 1). Interaction effects support a stronger MMN at 120–240 ms in the frontal quadrant.

Asymmetry and coordination

Asymmetry in Game 1 was -1 by design (the confederate never changed the current signaling system), and in Game 2 it was negative and different from 0 (median = -0.12 ; $N = 18$; $Z = 3.48$; $P < 0.001$): the sender (participant) changed his mapping of states to signals (learned in Game 1, possibly with modifications) less frequently than the receiver (confederate). Coordination was different from 0 (median = 0.77 ; $N = 90$, $Z = 0.86$, $P < 0.001$): the tone systems were effectively learned during the game. These data show a net flow of information from the sender to the receiver in both games.

Transmission and innovation

Transmission of tone sequences was faithful and significantly different from 0 (contour distance, median = 0.75 ; $N = 90$, $Z = 7.61$, $P < 0.001$), showing scarce innovation (median = 0.25 ; $N = 90$, $Z = 6.74$, $P < 0.001$). Using tone distance, instead, we observed lower levels of transmission (median = 0.2) and higher innovation (median = 0.8), reflecting ease of recall and production of the contours of sequences. This justifies our use of contour distance as a measure of transmission and innovation in subsequent analyses.

Structural changes

Tone sequences were restructured in the course of the signaling games. The entropy of tone sequences decreased from Game 1 to 2 ($Z = -2.673$, $P = 0.008$), more than contour entropy ($Z = -1.835$, $P = 0.06$). These changes reflect the emergence of

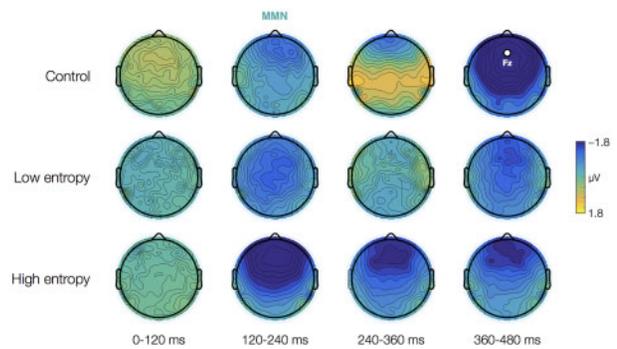


Fig. 3. Topographic isovoltage maps of grand-average ($N = 18$) difference waves between ERPs to deviants and standards in control, low- and high-entropy sequences. The MMN corresponds to the topographic maps shown at 120–240 ms; 0 ms is the onset of the deviant tone (fourth or fifth) in a sequence.

smoother sequences: tones were repeated or varied in small steps. Median 'interval compression ratio' values ($Z = -2.577$, $P = 0.01$) indicate changes toward smaller intervals.

To examine whether ERPs predict participant behavior in signaling games, we correlated MMN latencies or amplitudes with behavioral performance or structural changes to tone sequences during Game 2. Below we focus on the latency (120–240 ms) typical of the MMN (Supplementary Table S1–S8).

Neural predictors of learning and transmission

Figure 5 (upper panel) shows behavioral measures ('coordination', 'transmission', 'innovation accuracy') against MMN peak latencies in the anterior quadrant. MMN data from this quadrant only, in the high-entropy condition, were predictive of individual behavior (Supplementary Table S1). Peak latency was negatively correlated with 'coordination' ($P < 0.004$) and 'transmission' ($P < 0.01$): individuals with shorter MMN peak latencies learned more efficiently the signaling systems in Game 1, and transmitted them more faithfully in Game 2. Shorter MMN latencies also predicted fewer changes to the signaling system across games ('innovation', $P < 0.01$). Moreover, peak latencies from the anterior quadrant predicted 'accuracy' in Game 1 ($P < 0.004$). In Game 2, we observed a positive correlation with the number of changes introduced in the signaling system before coordination occurred, which suggests a greater effort to coordinate for participants with longer MMN peak latencies (Supplementary Table S3). Together, these results point to a relation between longer MMN peak latencies and less accurate or less faithful acquisition and transmission of tonal material.

No correlations in the remaining conditions (and quadrant) were significant (Supplementary Table S1 and S3). Using mean MMN amplitudes, correlations were weaker and right lateralized ($P < 0.05$; Supplementary Tables S2 and S4). No correlations were found in the 240–360 ms post-MMN window (Supplementary Tables S5–S8).

Neural predictors of structural change

Figure 5 (lower panel) shows absolute changes in tonal ('contour smoothness', 'sequence complexity') and interval structure (percentage of 'small' and 'large intervals') against peak latencies in the frontal scalp quadrant. Positive correlations were found with contour smoothness ($P < 0.05$), sequence complexity ($P < 0.001$) and changes in the percentage of smaller ($P = 0.01$) and larger intervals ($P = 0.004$). These data

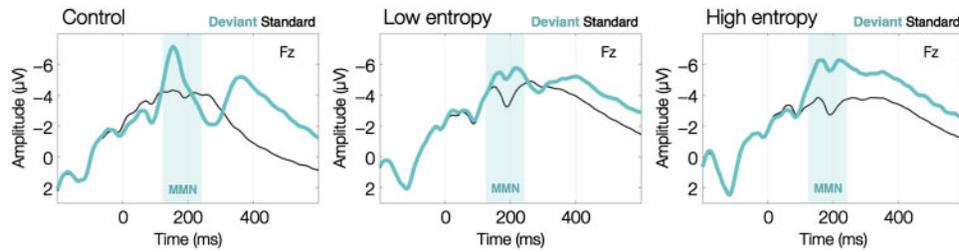


Fig. 4. Grand-average ERPs ($N = 18$) from the frontal midline electrode (Fz) in response to the fourth and fifth tones in standard and deviant stimuli in control, low and high entropy sequences. The waveforms shown here were low-pass filtered at 30 Hz.

Table 1. Repeated-measures analysis of variance for grand-average data

Main effects	Condition	F	df	P
Stimulus-type	HE	28.1	(1,17)	≤ 0.001
	LE	18.4	(1,17)	≤ 0.001
	CTRL	1.24	(1,17)	0.28
	HE	29.5	(1.81, 30.9)	≤ 0.001
	LE	33.8	(1.49, 25.3)	≤ 0.001
	CTRL	28.9	(2.68, 45.6)	≤ 0.001
	HE	104	(1.41, 23.9)	≤ 0.001
	LE	101	(1.54, 26.3)	≤ 0.001
	CTRL	98.6	(1.58, 26.9)	≤ 0.001
Two-way interaction Stimulus-type \times Temporal window	HE	28.8	(2.14, 36.5)	≤ 0.001
	LE	13.6	(2.37, 40.4)	≤ 0.001
	CTRL	36	(2.86, 48.7)	≤ 0.001
	HE	17.1	(1.61, 27.5)	≤ 0.001
	LE	2.04	(1.48, 25.2)	0.15
	CTRL	5.56	(1.45, 24.8)	≤ 0.01
Three-way interaction Stimulus-type \times Temporal window \times Quadrant	HE	14.9	(3.21, 54.7)	≤ 0.001
	LE	4.17	(4.14, 70.4)	≤ 0.01
	CTRL	15.9	(3.72, 63.3)	≤ 0.001

*P-values are Grenhouse-Geisser corrected.

Results of ANOVA statistics on mean MMN amplitude values in the high-entropy (HE), low-entropy (LE) and control conditions (CTRL).

reflect a reorganization of the melodic surface and interval distribution of sequences, mainly introduced by individuals with longer MMN latencies. In contrast, individuals with shorter MMN latencies were more conservative: specifically, they tended to maintain the melodic structures acquired in the first game.

Working memory

Working memory span scores were not predictive of behavior or structural changes in signaling games (all P -values > 0.05).

Discussion

The main finding of the present study is that inter-individual variability in pre-attentive processing of tone sequences, as revealed by MMN latencies, predicts subsequent learning, transmission and structural reorganization of signaling systems by participants. Below we discuss the possible origins and implications of this variability in the context of experimental research on cultural transmission.

Individual differences in MMN latencies

The auditory MMN reflects pre-attentive sensitivity to regularities and how accurately these are encoded as auditory traces (Alain et al., 1999a,b). Can the MMN be used to track individual differences in auditory processing capacity? Previous studies have shown correlations between the MMN characteristics and individual discrimination capability with complex spectrotemporal patterns (Näätänen et al., 1993; Tervaniemi et al., 2001). Näätänen et al. (1993) showed that poor auditory discrimination is related to a smaller MMN: a monotonic increase in MMN amplitude was observed for participants improving their performance over time. Converging results were obtained with hit-rates and reaction times (Tiitinen et al., 1994; Kraus et al., 1996; Tremblay et al., 1998; Menning et al., 2000) and fMRI (Zatorre, 2013): e.g. auditory cortex responses to microtonal intervals (smaller than the Western semitone) predicted individual discrimination of microtonal patterns (Zatorre et al., 2012). Overall, these data point to neural predispositions (not necessarily innate) in some individuals for processing complex auditory objects (Zatorre, 2013).

One possibility is that ERP latencies reflect connectivity strengths in the relevant networks, while amplitudes are linked

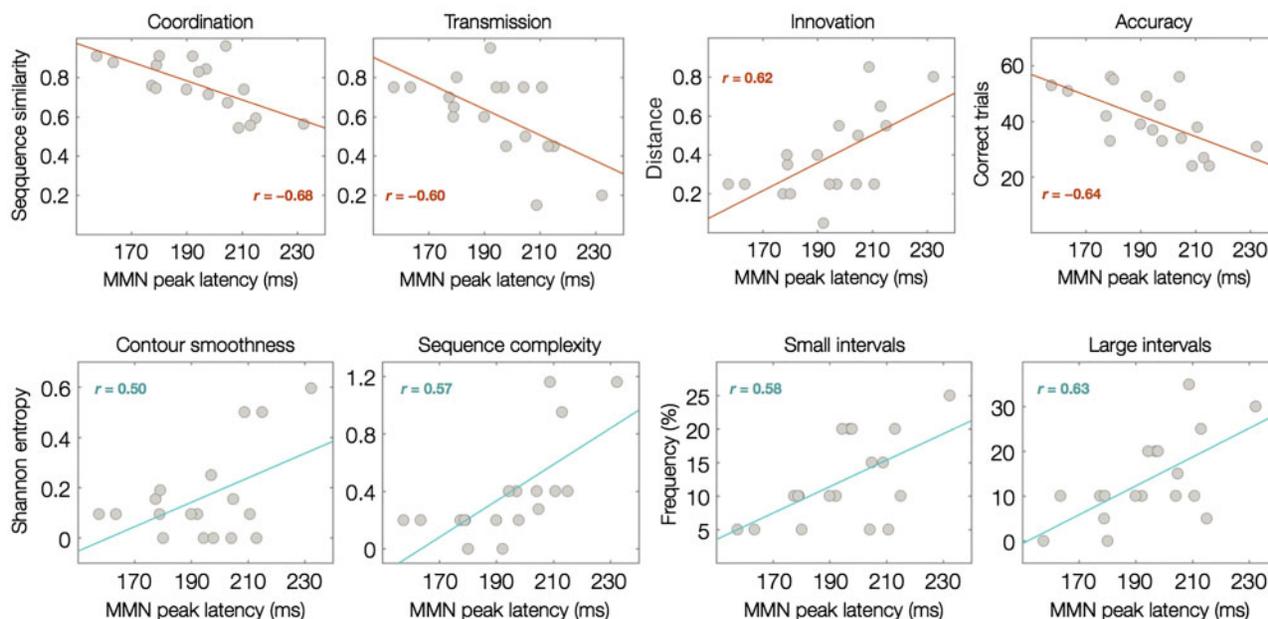


Fig. 5. Pearson product-moment correlations (r) between MMN peak latencies relative to the onset of the deviant tone (0 ms) in high-entropy sequences recorded from the frontal scalp quadrant on day 1 (x-axis), and behavioral (top, orange) and structural (bottom, blue) measures from signaling games on day 2 (y-axis). All MMN peaks fall within the 120–240 ms window. See ‘Methods’ section and Supplementary Figure 1 for further details on the measures. Each point on a scatterplot is one participant ($N = 18$).

to the extent of active tissue (Cardenas et al., 2005; Baggio and Hagoort, 2011; Baggio, 2012). This might explain why in our experiment the best predictors were MMN latencies and not amplitudes. Variation in MMN latencies may then be related to learning or maturational states of the relevant cortical networks (Moore and Guan, 2001; Trainor et al., 2003; Choudhury and Benasich, 2011). Inter-individual differences in temporal integration, moreover, may be a source of variation in MMN latencies: temporal integration is used by the auditory system to track input regularities, create a unified perceptual model (Bregman and Ahad, 1995) and predict upcoming events. Violations of expected regularities would then explain the generation of mismatch-related brain potentials (Winkler, 2007; Winkler et al., 2009 for reviews).

Temporal integration is key to create and maintain auditory representations of the environment (Sussman et al., 1999). In people with poor integration capacity, rapidly presented information, as in our experiment, may limit the accuracy of the perceptual model, and delay detection of the deviant sound. Sensory memory might be rapidly overwritten by novel inputs unless the stimuli are not rapidly processed (Christiansen and Chater, 2016). Although impairments in rapid temporal integration mechanisms have been reported in special or clinical populations (Bishop, 2007; Williamson and Stewart, 2010), it is likely that temporal integration abilities occur in a continuum. Individual differences in MMN characteristics are indeed more pronounced between- than within-individuals (Pekkonen et al., 1995). Collecting these evidences, we suggest that interindividual differences in MMN latencies observed here reveal individual constraints in auditory processing, rather than contextual or global, non-functional factors. In addition, if these low-level mechanisms govern perceptual information processing (Näätänen and Winkler 1999), they may account for performance at the behavioral level (Novak et al., 1992; Tiitinen et al., 1994; Tervaniemi et al., 1997).

Sequence complexity and behavioral prediction

Demands on temporal integration change with the complexity of the rules violated (Todd et al., 2010) as well as with individual differences (Näätänen and Winkler, 1999). From low- to high-entropy sequences, rules become more complex and less predictable, increasingly taxing the auditory system. Thus, control and low-entropy melodies (Fujioka et al., 2004; van Zuijen et al., 2004) draw on primitive grouping mechanisms and ‘top-down schemata’ largely shared across individuals of the same musical culture (Snyder, 2009). Here, differences in temporal integration, reflected by MMNs, are minimal (van Zuijen et al., 2004). Instead, greater individual variation in ERPs has been shown for complex patterns (Näätänen et al., 1993; Boh et al., 2011). In this condition, different integration capacities may be reflected at the behavioral level (Näätänen and Winkler 1999). Consistent with this notion, only MMN latencies recorded with high-entropy stimuli in our study were predictive of behavioral performance.

Learning, transmission and modification of symbolic systems

One key finding here is that learning and transmission are predicted by the ‘timing’ of brain processes: shorter MMN latencies were observed in faithful transmitters; longer latencies were found in individuals who innovate more the artificial tone system. This agrees with previous findings (Schröger et al., 1992; Näätänen et al., 1993; Tervaniemi et al., 1997). The result that latencies correlate with learning (‘coordination’) and successive recall (‘transmission’) provide additional support to the memory interpretation of the MMN, up to now addressed using auditory discrimination tasks (Näätänen et al., 2007).

Perhaps most interestingly, MMN latencies predict changes in the ‘structure’ of signals: individuals with longer latencies reorganised the melodic contour and interval distributions of sequences more often; smoother and proximal melodic forms

became more frequent from Game 1 to Game 2. Importantly, these changes follow the Gestalt principles of auditory perceptual proximity and good continuation: strings with such properties are easier to memorize and recall (Deutsch, 2013). Previous experimental studies support the view that cultural systems tend toward compressibility and simplicity (Tamariz and Kirby, 2015; Verhoef et al., 2014) to fit the capability and limits of human memory (Chater and Vitányi, 2003). Our work shows that these pressures can appear after a single round of learning and transmission, and suggests that the extent of the adaptation of cultural systems toward simpler forms may relate to individual constraints on information processing.

Individual processing constraints and cultural evolution

One leading hypothesis in cultural transmission research is that symbolic systems are shaped and constrained by properties of the human mind/brain (Christiansen and Chater, 2008). Humans transform information they receive from others, introducing features which fit general constraints on learning and memory (Boyd and Richerson, 1988). These constraints may be amplified during transmission, making the cultural system easier to acquire, use and transmit (Kirby, 2001; Griffiths and Kalish, 2007; Kirby et al., 2007, 2008; Smith and Kirby, 2008; Verhoef et al., 2014). Their existence has so far been inferred largely from observations at the population level (Griffiths et al., 2008; Kirby et al., 2008). Such constraints are often assumed to be too small to be observed and measured in single individuals and, as a result, no evidence has supported this view by direct neural investigations.

Here, we have shown that such constraints can be manifest at the individual level after only one instance of transmission, leaving distinct behavioral and neural traces. We have shown that the extent of change varies significantly across individuals. Our study provides initial neural evidence for the role of learning and memory constraints in the cultural transmission of symbolic systems (Christiansen and Chater, 2008), and musical systems in particular (Kleeman, 1985; Sawa, 1989; Huron, 2001; Trainor, 2015).

Supplementary data

Supplementary data are available at SCAN online.

Funding

No funding was received, in the form of grants or otherwise, than ordinary internal support for experimental research at SISSA.

Acknowledgements

We thank the anonymous reviewers for the helpful comments and their insights on paper organization.

Conflict of interest. None declared.

References

Alain, C., Achim, A., Woods, D.L. (1999a). Separate memory-related processing for auditory frequency and patterns. *Psychophysiology*, **36**, 737–44.

- Alain, C., Cortese, F., Picton, T.W. (1999b). Event-related brain activity associated with auditory pattern processing. *Neuroreport*, **9**, 3537–41.
- Alho, K., Sainio, K., Sajaniemi, N., Reinikainen, K., Näätänen, R. (1990). Event-related brain potential of human newborns to pitch change of an acoustic stimulus. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, **77**, 151–5.
- Baggio, G., Hagoort, P. (2011). The balance between memory and unification in semantics: A dynamic account of the N400. *Language and Cognitive Processes*, **26**, 1338–67.
- Baggio, G. (2012). Selective alignment of brain responses by task demands during semantic processing. *Neuropsychologia*, **50**, 655–65. 2012.
- Bishop, D.V.M. (2007). Using mismatch negativity to study central auditory processing in developmental language and literacy impairments: where are we, and where should we be going? *Psychological Bulletin*, **133**, 651.
- Boh, B., Herholz, S.C., Lappe, C., Pantev, C. (2011). Processing of complex auditory patterns in musicians and nonmusicians. *PLoS One*, **6**, e21458.
- Boyd, R., Richerson, P.J. (1988). *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.
- Bregman, A.S., Ahad, P.A. (1995). *Demonstrations of Auditory Scene Analysis: The Perceptual Organization of Sound*. Auditory Perception Laboratory, Psychology Department, Montreal: McGill University.
- Cardenas, V.A., Chao, L.L., Blumenfeld, R., et al. (2005). Using automated morphometry to detect associations between ERP latency and structural brain MRI in normal adults. *Human Brain Mapping*, **25**, 317–27.
- Chater, N., Vitányi, P. (2003). Simplicity: a unifying principle in cognitive science? *Trends in Cognitive Sciences*, **7**, 19–22.
- Choudhury, N., Benasich, A.A. (2011). Maturation of auditory evoked potentials from 6 to 48 months: prediction to 3 and 4 year language and cognitive abilities. *Clinical Neurophysiology*, **122**, 320–38.
- Christiansen, M.H., Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, **31**, 489–509.
- Christiansen, M.H., Chater, N. (2016). The now-or-never bottleneck: a fundamental constraint on language. *Behavioral and Brain Sciences*, **39**, e62.
- Cross, I. (2005). Music and meaning, ambiguity and evolution. *Musical Communication*, 27–43, Oxford: Oxford University Press.
- Dehaene, S., Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron*, **56**, 384–98.
- Dehaene, S., Cohen, L., Morais, J., Kolinsky, R. (2015). Illiterate to literate: behavioral and cerebral changes induced by reading acquisition. *Nature Reviews Neuroscience*, **16**, 234–44.
- 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.
- Deutsch, D. (2013). Grouping mechanisms in music. In Deutsch D., editors. *The Psychology of Music*, pp. 183–248. San Diego: Elsevier.
- Ekman, P., Friesen, W.V. (2003). *Unmasking the Face: A Guide to Recognizing Emotions from Facial Clues*. Los Altos: Malor Books.
- Foster, J.L., Shipstead, Z., Harrison, T.L., Hicks, K.L., Redick, T.S., Engle, R.W. (2015). Shortened complex span tasks can reliably measure working memory capacity. *Memory and Cognition*, **43**, 226–36.
- Fujioka, T., Trainor, L.J., Ross, B., Kakigi, R., Pantev, C. (2004). Musical training enhances automatic encoding of melodic

- contour and interval structure. *Journal of Cognitive Neuroscience*, **16**, 1010–21.
- Gottselig, J.M., Brandeis, D., Hofer-Tinguely, G., Borbély, A.A., Achermann, P. (2004). Human central auditory plasticity associated with tone sequence learning. *Learning and Memory*, **11**, 162–71.
- Griffiths, T.L., Kalish, M.L. (2007). Language evolution by iterated learning with Bayesian agents. *Cognitive Science*, **31**, 441–80.
- Griffiths, T.L., Kalish, M.L., Lewandowsky, S. (2008). Theoretical and empirical evidence for the impact of inductive biases on cultural evolution. *Philosophical Transactions of the Royal Society of London B*, **363**, 3503–14.
- Hamming, R.W. (1986). *Coding and Information Theory*. Englewood Cliffs, NJ: Prentice-Hall, Inc.
- Hannagan, T., Amedi, A., Cohen, L., Dehaene-Lambertz, G., Dehaene, S. (2015). Origins of the specialization for letters and numbers in ventral occipitotemporal cortex. *Trends in Cognitive Sciences*, **19**, 374–82.
- Hsieh, S., Hornberger, M., Pigué, O., Hodges, J.R. (2012). Brain correlates of musical and facial emotion recognition: evidence from the dementias. *Neuropsychologia*, **50**, 1814–22.
- Huron, D. (2001). Tone and voice: A derivation of the rules of voice-leading from perceptual principles. *Music Perception*, **19**, 1–64.
- Juslin, P.N., Sloboda, J.A. (2001). *Music and Emotion: Theory and Research*. Oxford: Oxford University Press.
- Kirby, S. (2001). Spontaneous evolution of linguistic structure: an iterated learning model of the emergence of regularity and irregularity. *IEEE Transactions on Evolutionary Computation*, **5**, 102–10.
- Kirby, S., Cornish, H., Smith, K. (2008). Cumulative cultural evolution in the laboratory: an experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*, **105**, 10681–6.
- Kirby, S., Dowman, M., Griffiths, T.L. (2007). Innateness and culture in the evolution of language. *Proceedings of the National Academy of Sciences*, **104**, 5241–5.
- Kleeman, J.E. (1985). The parameters of musical transmission. *The Journal of Musicology*, **4**, 1–22.
- Kraus, N., McGee, T.J., Carrell, T.D., Zecker, S.G., Nicol, T.G., Koch, D.B. (1996). Auditory neurophysiologic responses and discrimination deficits in children with learning problems. *Science*, **273**, 971–3.
- Lense, M.D., Gordon, R.L., Key, A.P., Dykens, E.M. (2014). Neural correlates of cross-modal affective priming by music in Williams syndrome. *Social Cognitive and Affective Neuroscience*, **9**(4), 529–37.
- Lewis, D. (1969). *Convention. A Philosophical Study*. Cambridge, Mass: Harvard University Press.
- Lopez-Calderon, J., Luck, S.J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, **8**, 1–14.
- Mathews, M.V., Pierce, J.R., Reeves, A., Roberts, L.A. (1988). Theoretical and experimental explorations of the Bohlén-Pierce scale. *Journal of the Acoustical Society of America*, **84**, 1214–22.
- Menning, H., Roberts, L.E., Pantev, C. (2000). Plastic changes in the auditory cortex induced by intensive frequency discrimination training. *Neuroreport*, **11**, 817–22.
- Moore, J.K., Guan, Y.L. (2001). Cytoarchitectural and axonal maturation in human auditory cortex. *Journal of the Association for Research in Otolaryngology*, **2**, 297–311.
- Moreno, M., Baggio, G. (2015). Role asymmetry and code transmission in Signaling Games: An experimental and computational investigation. *Cognitive Science*, **39**, 918–43.
- Näätänen, R., Gaillard, A.W., Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, **42**, 313–29.
- Näätänen, R., Paavilainen, P., Rinne, T., Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology*, **118**, 2544–90.
- Näätänen, R., Schröger, E., Karakas, S., Tervaniemi, M., Paavilainen, P. (1993). Development of a memory trace for a complex sound in the human brain. *NeuroReport*, **4**, 503–6.
- Näätänen, R., Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin*, **125**, 826.
- Novak, G., Ritter, W., Vaughan, H.G. (1992). The chronometry of attention-modulated processing and automatic mismatch detection. *Psychophysiology*, **29**, 412–30.
- Nowak, I., Baggio, G. (2016). The emergence of word order and morphology in compositional languages via multi-generational signaling games. *Journal of Language Evolution*, lzw007.
- Orsini, A., Grossi, D., Capitani, E., Laiacona, M., Papagno, C., Vallar, G. (1987). Verbal and spatial immediate memory span: normative data from 1355 adults and 1112 children. *Italian Journal of Neurological Sciences*, **8**, 537–48.
- Paavilainen, P. (2013). The mismatch-negativity (MMN) component of the auditory event-related potential to violations of abstract regularities: A review. *International Journal of Psychophysiology*, **88**, 109–23.
- Palmer, S.E., Schloss, K.B., Xu, Z., Prado-León, L.R. (2013). Music-color associations are mediated by emotion. *Proceedings of the National Academy of Sciences*, **110**, 8836–41.
- Pekkonen, E., Rinne, T., Näätänen, R. (1995). Variability and replicability of the mismatch negativity. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, **96**, 546–54.
- Sawa, G.D. (1989). Oral transmission in arabic music, past and present. *Oral Tradition*, **4**, 254–65.
- Schröger, E., Näätänen, R., Paavilainen, P. (1992). Event-related potentials reveal how non-attended complex sound patterns are represented by the human brain. *Neuroscience Letters*, **146**, 183–6.
- Seppänen, M., Brattico, E., Tervaniemi, M. (2007). Practice strategies of musicians modulate neural processing and the learning of sound-patterns. *Neurobiology of Learning and Memory*, **87**, 236–47.
- Shannon, C.E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, **27**, 623–56. 379–423,
- Skyrms, B. (2010). *Signals: Evolution, Learning, and Information*. Oxford: Oxford University Press.
- Smith, K., Kirby, S. (2008). Cultural evolution: implications for understanding the human language faculty and its evolution. *Philosophical Transactions of the Royal Society B*, **363**, 3591–603.
- Snyder, B. (2009). Memory for music. In: Hallam, S., Cross, I., Thaut, M., editors. *The Oxford Handbook of Music Psychology*, pp. 171–83. Oxford: Oxford University Press.
- Sussman, E., Ritter, W., Vaughan, H.G. (1999). An investigation of the auditory streaming effect using event-related brain potentials. *Psychophysiology*, **36**, 22–34.
- Tamariz, M., Kirby, S. (2015). Culture: Copying, compression, and conventionality. *Cognitive Science*, **39**, 171–83.
- Todd, J., Myers, R., Pirillo, R., Drysdale, K. (2010). Neuropsychological correlates of auditory perceptual inference: a mismatch negativity (MMN) study. *Brain Research*, **1310**, 113–23.
- Tervaniemi, M., Ilvonen, T., Karma, K., Alho, K., Näätänen, R. (1997). The musical brain: Brain waves reveal the neurophysiological basis of musicality in human subjects. *Neuroscience Letters*, **226**(1), 1–4.

- Tervaniemi, M., Rytkönen, M., Schröger, E., Ilmoniemi, R.J., Näätänen, R. (2001). Superior formation of cortical memory traces for melodic patterns in musicians. *Learning and Memory*, **8**, 295–300.
- Tierney, A.T., Russo, F.A., Patel, A.D. (2011). The motor origins of human and avian song structure. *Proceedings of the National Academy of Sciences*, **108**, 15510–5.
- Tiitinen, H., May, P., Reinikainen, K., Näätänen, R. (1994). Attentive novelty detection in humans is governed by pre-attentive sensory memory. *Nature*, **372**(6501), 90–2.
- Trainor, L.J. (2015). The origins of music in auditory scene analysis and the roles of evolution and culture in musical creation. *Philosophical Transactions of the Royal Society B*, **370**, 20140089.
- Trainor, L., McFadden, M., Hodgson, L., et al., (2003). Changes in auditory cortex and the development of mismatch negativity between 2 and 6 months of age. *International Journal of Psychophysiology*, **51**, 5–15.
- Tremblay, K., Kraus, N., McGee, T. (1998). The time course of auditory perceptual learning: Neurophysiological changes during speech sound training. *Neuroreport*, **9**, 3557–60.
- Van Zuijen, T.L., Sussman, E., Winkler, I., Näätänen, R., Tervaniemi, M. (2004). Grouping of sequential sounds—an event-related potential study comparing musicians and non-musicians. *Journal of Cognitive Neuroscience*, **16**, 331–8.
- Verhoef, T., Kirby, S., de Boer, B. (2014). Emergence of combinatorial structure and economy through iterated learning with continuous acoustic signals. *Journal of Phonetics*, **43**, 57–68.
- Von Hippel, P. (2000). Redefining pitch proximity: Tessitura and mobility as constraints on melodic intervals. *Music Perception: An Interdisciplinary Journal*, **17**, 315–27.
- Williamson, V.J., Stewart, L. (2010). Memory for pitch in congenital amusia: Beyond a fine-grained pitch discrimination problem. *Memory*, **18**, 657–69.
- Winkler, I. (2007). Interpreting the mismatch negativity. *Journal of Psychophysiology*, **21**, 147–63.
- Winkler, I., Denham, S.L., Nelken, I. (2009). Modeling the auditory scene: predictive regularity representations and perceptual objects. *Trends in Cognitive Sciences*, **13**, 532–40.
- Zatorre, R.J. (2013). Predispositions and plasticity in music and speech learning: neural correlates and implications. *Science*, **342**, 585–9.
- Zatorre, R.J., Delhommeau, K., Zarate, J.M. (2012). Modulation of auditory cortex response to pitch variation following training with microtonal melodies. *Frontiers in Psychology*, **2**, 544.