

Structural decomposition of EEG signatures of melodic processing

Rebecca S. Schaefer^{a,*}, Peter Desain^a, Patrick Suppes^b

^a Donders Institute for Brain, Cognition and Behavior: Centre for Cognition, Montessorilaan 3, 6525 HE Radboud University Nijmegen, The Netherlands

^b Center for the Study of Language and Information, Ventura Hall, Stanford University, Stanford, CA 94305-4101, USA

ARTICLE INFO

Article history:

Received 30 October 2008

Accepted 11 August 2009

Available online 19 August 2009

Keywords:

Music perception

Electroencephalography

Event-related multiple regression

Imagery

ABSTRACT

In the current study we investigate the EEG response to listening and imagining melodies and explore the possibility of decomposing this response according to musical features, such as rhythm and pitch patterns. A structural model was created based on musical aspects and multiple regression was used to calculate profiles of the contribution of each aspect, in contrast to traditional ERP components. By decomposing the response, we aimed to uncover pronounced ERP contributions for aspects of the encoding of musical structure, assuming a simple additive combination of these. When using a model built up of metric levels and contour direction, 81% of the variance is explained for perceived, and 57% for imagined melodies. The maximum correlation between the parameters found for the same melodic aspect in perception vs. imagery was 0.88, indicating similar processing between tasks. The decomposition method is shown to be a novel analysis method of complex ERP patterns, which allows subcomponents to be investigated within a continuous context.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

In processing a musical stimulus, the information is thought to be grouped and categorized on multiple levels (see [Lerdahl and Jackendoff, 1983](#)). In probing the characteristics of these processes, it can be very challenging to isolate the different musical aspects and separate their related responses. The current study proposes a method to decompose the EEG response to simple melodies, both heard and imagined, and to isolate the brain response common to separate musical aspects of the stimulus.

The investigation into how the melodic structure is stored (as a sequence of pitches, as interval jumps, as a tonal contour, as pitch functions within the key, or as a combination of these) has been very informative in behavioral studies. Differences in reaction times have revealed hierarchical processing of tonality and chord structure ([Bharucha and Krumhansl, 1983](#)), the time course of recognition memory has indicated that contour (the relative up-down pattern of pitches in a melody) may be processed separately from the absolute pitch pattern ([Dowling et al., 1995, 2001](#)). Recognition rates have shown varying levels of attention over different melodic and metric accents ([Jones and Boltz, 1989](#)) and sophisticated scrambling methods have shown interactions of global and local perception (i.e. [Tillmann and Bigand, 2001](#)), among others.

For a number of these musical aspects the EEG response has also been investigated, albeit mostly in the context of ERP components that have been described earlier, such as the P300 oddball response, the mismatch negativity (MMN) and language syntax related components such as the N400. Effects on the P300-complex were found for a number of oddball stimuli, such as minor vs. major mode for musicians ([Halpern et al., 2007](#)) and unexpected pitches in scales ([Krohn et al., 2007](#)). MMNs, occurring when infrequently occurring stimuli are perceived, were also seen for rare chord modulations ([Koelsch et al., 2003](#)) (where the authors refer to an early right anterior negativity as a music-syntactic MMN), out-of-key pitches ([Brattico et al., 2006](#)), and contour violations ([Trainor et al., 2002](#)). ERPs in response to musical rhythms have also been investigated, for instance for different metric levels such as the note, beat and bar level ([Jongsma et al., 2004](#)), as well as subjective accents ([Brochard et al., 2003](#)). On a more global level, parallel to phrases in language, musical phrase endings often induce a so-called closure-positive shift ([Neuhaus et al., 2006](#)).

In a single note, multiple characteristics or functions are combined, such as its absolute and relative pitch, the interval jump it has just made, its absolute and relative duration, hierarchical metric level, role in the harmony and more. How these aspects combine to form a representation is not clear. [Jones and Ralston \(1991\)](#) describe different aspects of the stimulus (i.e. the rhythmic pattern, the pitch pattern) as each having their own accent structures, and assume that the different accent structures combine in some way to form our representation. To investigate these aspects individually, they need to be isolated from each other. This is practically impossible for melodies that inherently

* Corresponding author. Tel.: +31 24 361 5458; fax: +31 24 361 6066.
E-mail address: r.schaefer@donders.ru.nl (R.S. Schaefer).

contain these different layers, and can only be achieved by keeping constant as many of the other aspects as possible. The studies described above tend to use factorial designs, contrasting different conditions to investigate their different effects on the event-related brain response. Although there is generally an assumption that different ERP components combine additively, very few studies have tried to separate the effects of processing different aspects of a stimulus. A relatively new method of separating such intercorrelated variables is using multiple linear regression analysis (e.g. Schaefer and Desain, 2006; Hauk et al., 2006, 2009; Dambacher et al., 2006) to test the degree to which a variable (or in the current study musical aspect) predicts data across all trials, in a continuous context. For fMRI data, the use of continuous regressors imitating the ongoing modulations is more common (see Cohen, 1997) and has already been proposed in the music domain by Janata et al. (2002).

To investigate the cognitive involvement in processing melodies separately from the auditory response, the investigated constructs also need to be isolated from the processing of sound, distinguishing the low-level perceptual mechanisms from the higher level cognitive processes. As we have a rich capacity for imagining music, imagery offers a means to get around auditory information processing while still evoking the representation of the melody. By investigating the relationship between perceived and imagined modalities, the level of similarity between the different tasks can also be addressed. This issue becomes particularly important, as the two have been shown to be entwined, both in behavioral and EEG studies.

In previous work that has looked at imagery of music, specific attention has been given to the spatial location of music imagery in the brain, as measured with PET or fMRI. Brain structures implicated in musical imagery were shown to be very similar to those recruited by actual perception, although involvement of primary auditory areas is not always found (c.f. Halpern et al., 2004; Kraemer et al., 2005). However, looking at imagery for melodies in EEG, Janata (2001) found that of a series of imagined notes, only the first elicited an N1 component as is generally seen for actually perceived notes. The subsequent imagined notes did not elicit such a component, most likely indicating the response to the first note to signify a state or task change. Other investigations of music imagery in EEG have focused specifically on musicians, for instance on the motor component induced by imagining the sound of an instrument participants played themselves (Kristeva et al., 2003), or the MMN elicited when a note imagined from notation deviates from a note actually presented (Yamamoto et al., 2005). When specifically investigating imagined rhythmic patterns, Desain and Honing (2003) showed that classification of internally rehearsed rhythms from the EEG signature is possible well above chance level, and preliminary results also show detectability of imagined natural overlearned music from the EEG (Schaefer et al., 2008). More basic investigations include work by Meyer et al. (2007), who found a reduced N1 and almost absent P2 response (see also Scherg et al., 1989) for imagined piano triads. However, these investigations do not address the multiple layers of a musical stimulus, such as metric levels or pitch structure.

In the current study, the full ERP trace of listening and imagining melodies is investigated, and the possibilities of decomposing these traces according to musical phenomena, here referred to as musical aspects, are explored. These musical phenomena or aspects add multiple levels of structure to a musical stimulus, and together form the melody. Examples are the absolute pitch sequence, the relative pitch pattern, the sequence of implied chords, and the rhythmic structure, which is multi-leveled itself. By taking this structure as the basis for the decomposition, we thus use a priori knowledge about the stimulus to decompose the response, similarly to Windsor et al. (2006) and Desain et al. (2008). A structural symbolic model can be

built to represent these layers of structure, defined by simply labeling notes according to different musical aspects. Each aspect can have a number of levels (such as for instance metric depth), but is reduced into a set of components on separate levels (note, beat, bar, etc.) that are either absent or present. This structure can be expressed in a matrix of ones and zeros, shown more elaborately in the method description below. By regressing EEG data according to this structure, the contributions of different musical aspects to melodic processing can be calculated.

If we assume that these structural levels combine linearly in the ERP response, we can find the parameters that are associated with these musical aspects by using simple least-squares linear regression. By regressing the data with only one, or a few components that together correspond to the levels of a musical aspect, the prediction for only that component (such as beat) or aspect (combined levels of meter) can be formed. By calculating how much variance is explained by each aspect or component, the size and significance of that part of the model can be assessed. This procedure is very similar to the deconvolution algorithms used in the analysis of fMRI data (Glover, 1999). In this way, we have different predictor variables for different time segments, depending on the note within the sequence. Furthermore, as we fit the data separately per EEG channel, the topology of the component responses can be investigated, yielding insight into the different cognitive modules involved in the processing of the various aspects of musical structure. In the model we created, we used a three-layered rhythmic structure ('First', or the start of a phrase; 'Beat', or the pulse of the rhythm, and 'Note', the lowest rhythmic level using every event) and contour direction, based on the up-down pattern of the pitch sequence. These particular aspects were selected based on previous report of a distinguishable response in the EEG signal, such as Janata (2001) and Brochard et al. (2003) for rhythm and Trainor et al. (2002) for contour direction changes.

2. Method

2.1. Participants

Eighteen healthy volunteers with normal hearing took part in the study, from the undergraduate student body at Stanford in 2004. Although musical expertise is relevant to musical information processing, we did not make this influence the subject of the current investigation, and musical background was not taken into account. The data for two participants were not used because of extensive artifacts yielding poor signal quality, as judged by visual inspection.

2.2. Stimuli and procedure

This study was part of a longer experiment, in which three experiments were interleaved, separated in blocks. These are not analyzed here. The other tasks were listening to phonemes and sentences while making emotional and word association judgments. Four original melodies were constructed by Elizabeth Margulis, and consisted of three parts, a perception part, an imagery part and a probe tone part. Here, we only analyze the first two sections and will for further purposes ignore the probe tone section. Each melody consisted of 7 notes of which the first six were of equal duration, and the last note was twice this duration. This time-structure, combined with the chosen pitch sequence, induces the impression of a binary beat structure, in which rhythmic groups of two or four events emerge. The notes were presented using a midi-piano sound, followed by the same rhythmic pattern played by a midi high-hat, so as to induce a time lock for the imagery task. For an overview in music notation (in 4/4 measure) of the stimulus melodies as they were presented, see Fig. 1. The stimuli are also available to listen to on <http://www.nici.ru.nl/mmm> under 'Demos and Stimuli'. The participants were instructed to repeat the melody in their head in concurrence with the high-hat sound. Notes were presented with an inter onset interval of 1/3 of a second (180 beats per minute). The key-off time used for each note was 50 ms before the next onset, although due to the natural piano decay after key-release there were no real silent gaps between notes, resulting in legato-style. Each stimulus was presented 36 times.

2.3. Equipment

EEG recordings were made at the Suppes Brain Lab at Stanford University, using 21 Model-12 Grass amplifiers and Neuroscans Scan 4 software. Sensors were attached to the scalp of a subject according to the standard 10–20 EEG system (Jasper, 1958), using channel locations Fp1, Fp2, F3, F4, F7, F8, Cz, C3, C4, Pz, P3, P4,



Fig. 1. The four stimuli, each consisting of a melody (E. Margulis), followed by a repetition of the rhythmic pattern with a high-hat sound, to ensure the time lock of the imagined events.

	Melody 1							Melody 2							Melody 3							Melody 4						
first	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0
beat	1	0	1	0	1	0	1	1	0	1	0	1	0	1	1	0	1	0	1	0	1	1	0	1	0	1	0	1
note	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
contour up	0	1	1	0	1	1	0	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0	1	0	0	1	0	1
stimulus																												

Fig. 2. The structural matrix of melodic aspects for each of the four stimulus melodies. In this way, each note has a specific profile made up of its characteristics.

T3, T4, T5, T6, O1, O2 as well as vertical EOG, horizontal EOG, throat EMG and an earlobe reference. The recording bandwidth was from 0.3 to 100 Hz with a sampling rate of 1000 Hz. A computer was used to present auditory stimuli (digitized at 22 kHz) to subjects via small loudspeakers. All analyses were carried out in Matlab (The Mathworks, Natick, MA).

2.4. Analyses

2.4.1. Decomposition

A structural model was created based on the presence or absence of several musical aspects. Although numerous models can be constructed, we here demonstrate the method with a simple model that uses only basic aspects and illustrates the method. The matrix for the a priori model we used is shown in Fig. 2.

This model creates separate parameters for the first note of a phrase, the beat-based metric level (grouping subjectively accented events in even positions), the note level (partialing out the effects that are common for each note), and one pitch-based aspect, contour direction. Note that one needs to take care to define structural aspects and levels sparsely, otherwise the number of parameters (rows in Fig. 2) becomes larger than the number of data points (columns), and the resulting degenerate matrix can no longer be used to fit a unique solution. In our example this means, for instance, that while we have two contour directions in our melodies, we define only one level of contour (up) in the table. The other one (down) is considered as part of the component for every note, as the two are mutually exclusive for these particular melodies. This is only the case because tone repetitions do not occur in these melodies, which would otherwise remove the complementary nature of the two directions. The aspects chosen here each yield a unique parameter, although some necessarily overlap. An example is that ‘First’ and ‘Beat’ always co-occur, defining the two parameters to both contribute to the response to the first note in a melody, but ‘Beat’ also occurs independently of ‘First’. By adding the component ‘Note’, we capture all activity that is common over all notes, such as the sounds of the metronome in the case of imagery, or the common spectral properties of the sound during perception. This then leaves the residual activity, which is specific to each note, to be used to form the other parameters. In a similar vein, the aspect ‘Beat’, which is not explicitly present in the stimulus, but expected to show some EEG response based on the literature in subjective accenting, may overlap with the others, but as it is a distinct grouping of trials, the parameter will only be fitted to activity that is unique to this grouping. To the extent that this cannot be defined, activity will be shared. Of course an alternative model can also be constructed, based on different characteristics of each note, in order to compare their power in explaining the data.

To deconvolve the ERP data, we expand the structure of Fig. 2 in time into a table that defines the structure for each measurement sample, replacing every 1 by a diagonal unity matrix with the length of that segment in timepoints, as shown in Fig. 3. Assuming that different musical aspects each have their own (unknown) response component in the ERP, and that these different responses combine linearly, the structure matrix can be multiplied by the as-yet unknown parameters

(a concatenation of the component segments shown on the right) to predict the data vector (a concatenation of the ERP response to the used stimulus melodies for either perception or imagery).

The method can formally be described as follows. The input stimulus is a binary sequence which comes from the structural model shown in Figs. 2 and 3. We postulate that the brain responds differently to the presence or absence (1 or 0) of a musical aspect. Finally, we assume that each musical aspect contributes a time-limited waveform which combine linearly to give the total stimulus response. Fig. 5 schematically illustrates this model. The same decomposition model was used for timing signals in Windsor et al. (2006). In algebraic terms this model can be written as:

$$x(t) = \sum_{\tau=1}^L \sum_{a=1}^n I_a(t) R_a(t - \tau) \tag{1}$$

where $x(\tau)$ is the total response at time t , L is the duration of the response, R_a are the temporal responses of the brain to the presence (or absence) of the n musical aspects a in the stimulus, and $I_a(t)$ is an indicator function which has the value 1 if there is an aspect a present at time t , and 0 otherwise. This model can more compactly be expressed in matrix notation using a structure matrix M to encode the indicator functions I_a as,

$$x = \begin{bmatrix} \vdots & \vdots & \vdots \\ I_1(i : i + L) & \cdots & I_n(i : i + L) \\ \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} R_1 \\ \vdots \\ R_n \end{bmatrix} = Mp \tag{2}$$

where x is the column vector of modeled response for each time, the rows of M signify sample times with each row being the previous row shifted 1 element to the right, and p is the concatenation of the different types of response function. Eq. (2) is linear in the temporal responses R_a and p , so these parameters can be found using a least-squares regression with the average measured response.

To test the significance of the contribution of the individual aspects, an F -test was used on the increase of the explained variance, compared to the variance explained by the model without this aspect. The regression analysis was carried out over the mean ERP response of all subjects, as there were not enough data to do this separately and subject the regression coefficients to group statistics. We used all the electrodes, and thus have a percentage of variance the model explains for each electrode position.

2.4.2. ERP latencies

As we are interested in the cognitive processing of these notes, it is likely that at least part of the response we are looking for does not occur within the 333 ms until the next note is presented. To deal with the latency of the response and prevent the data segment we are looking at from still containing much of the response to the previous note, the data segments were shifted forward to follow the stimulus note at a time lag. By shifting the position (not the duration) of the parameter in time, the

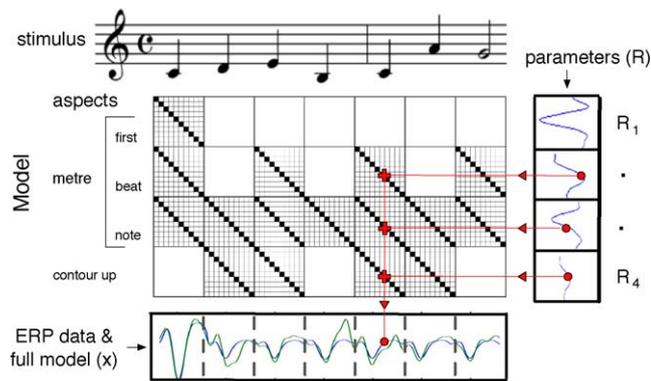


Fig. 3. A schematic overview of the method, using an example of a stimulus melody to create the structural model, and predicting the ERP data from parameter segments that correspond to the model. The parameters are optimized using linear regression.

time lag of the response was accounted for. We optimized the lag of the model behind the stimulus and found that shifting the segments 0.2 s maximizes the predictive value of the model. Interestingly, this probably excludes some of the more low-level perceptual responses, which appear not to contribute to the processing of the chosen aspects. The fact that the model explains more variance at this time lag further argues for the notion that we are looking at cognitively driven ERP responses, as more time is needed for these higher cortical processes.

2.4.3. Inter-task correlation

To assess the commonalities of the decomposition between perception and imagery, correlations between tasks were calculated for both the data and the parameters found. By calculating this for each channel a distribution of task similarity emerges for every aspect taken up into the model.

3. Results

3.1. ERPs

The full ERP trace at C4 (right auditory area) for each melody over all participants is shown in Fig. 4. The ERPs of the perceived melodies show a similar pattern per note, showing a N1-P2 complex for every note (c.f. Scherg et al., 1989). For the imagined melodies, only the first imagined note shows this response, and the ERP flattens after that, much like the responses to imagined notes described in Janata (2001). Although the absence of a clear auditory N1 response is interesting, given the fact that there is still a high-hat sound present as a timekeeper, the absent P2 response for imagined sound events is consistent with Meyer et al. (2007) as well as the finding that this component is modulated by the spectral complexity of the sounds (Shahin et al., 2005), when comparing the richer pitch structure for the perceived melodies than that of just the high-hat sound.

3.2. Regression results

For the model that was tested, the explained variance of the full model at Cz, as well as for each isolated aspect, is shown in Table 1.

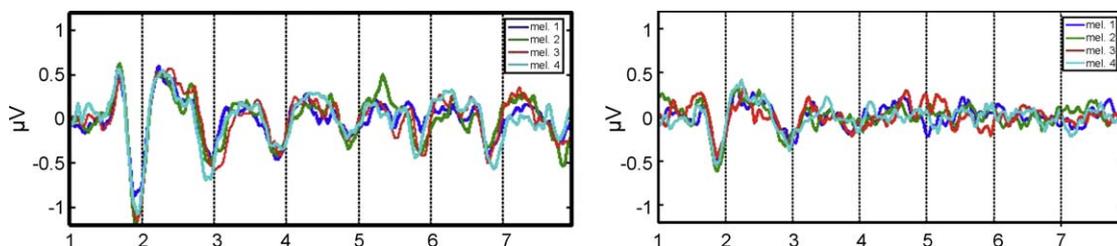


Fig. 4. ERP signatures at electrode C4 (over right auditory area) for the whole melodies, perception (left) and imagery (right). The ticks on the x-axis mark the notes in the melody.

Table 1

Explained variance and significance results for the full combined model as well as per aspect, for electrode Cz.

Model	Task			
	Perception		Imagery	
	r^2 (adj)	p	r^2 (adj)	p
Full model	83(81)%	< 0.0001	63(57)%	< 0.01
First	48%	< 0.0001	47%	< 0.04
Beat	0%	n.s.	6%	< 0.01
Note	55%	< 0.0001	17%	< 0.01
Contour direction	11%	< 0.065	10%	n.s.
Residue	17%		37%	

At this electrode, the full model explains the maximal amount of variance. The parameters found through the decomposition are shown in Fig. 5, showing the ERP in the top row, followed by the parameters found for each aspect. The bottom two rows show the full model, built up from these parameters, and the residue of the data when the model is subtracted. This figure also shows the scalp maps of the r^2 explained by a single aspect within the model, as well as the power distribution (s^2) of the signal that is responsible for explaining this variance. Most of the variance in the data is explained by perhaps the least musical aspects in the model, namely the note level and the start of the phrase, which by themselves explain up to around 55% and 48% of the variance, for perception ($p < 0.001$, $p < 0.001$) and imagery ($p < 0.04$, $p < 0.01$), respectively. Contour direction shows a trend to significance in perception, however in imagery it does not reach significance. Alternately, the beat-level subjective accenting does not reach significance during perception, but explains 6% of the data for imagery with $p < 0.01$.

3.3. Correlations between tasks

When calculating the correlation between perception and imagery tasks for each of the estimated parameters over the scalp, the decomposed parameters correlate much more highly than the ERP data do between tasks. To compare the parameters from the decomposition with the full ERP signature of the non-decomposed data, we used the complete length of each melody. We must keep in mind that for the parameter correlations, this causes parts of the melody trace to become zero (due to the structural model), which increases the correlations. The topographic maps of the correlations are shown in Fig. 6. They show the degree of correlation for the full data signature to be smoothly distributed over the scalp, with a maximum of about 0.6, whereas the separate musical aspects show much higher correlations (up to 0.88) and more pronounced patterns on the topographic maps. This shows that the similarities between tasks have different distributions per aspect in the model. These distributions are most likely produced by the activity of combined networks of brain regions. For the note level, it is interesting to note that even though what each note has in common is a midi-piano sound for perception and a high-hat

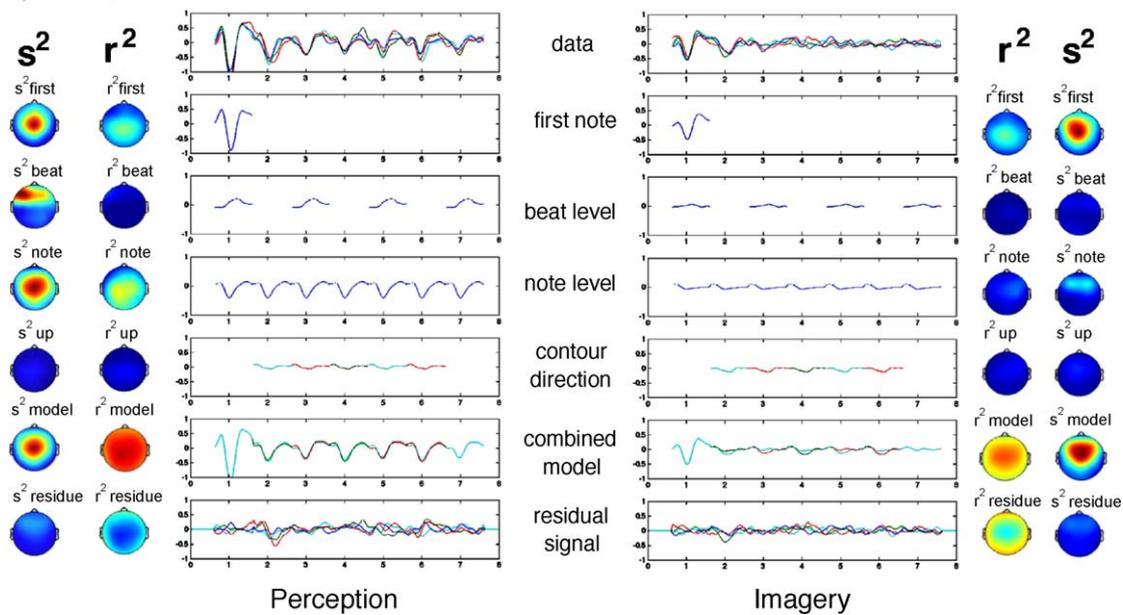


Fig. 5. Additive effect of melodic aspects for perception (left) and imagery (right). Each color represents one of the stimulus melodies. The top plots show the ERP data, followed by graphs of the parameters derived from the decomposition for first, beat, note and contour direction. Only the contour parameter differs per stimulus, which makes for a different contribution to each combined or full model. The two bottom plots show the full models for each melody and the residues left after subtraction of the model from the ERP. The channel shown here is Cz, where the combined model explained the maximum amount of variance. The top plots show the distribution of power of the signal of this parameter (s^2) and the explained variance in the column closest to the waveforms (r^2) and the distribution of power of the signal of this parameter (s^2).

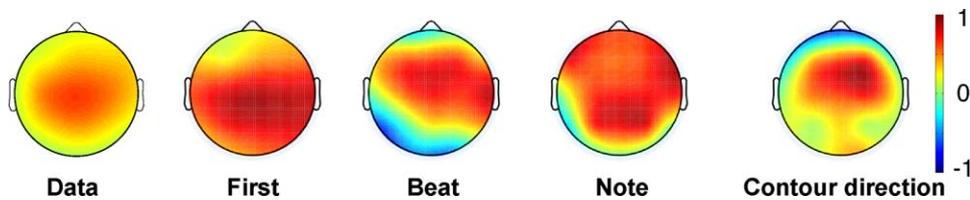


Fig. 6. The distribution of correlations per channel between perception and imagery, for the original (left) and decomposed (others) data, with a different distribution for the state-change or ‘first’, the beat-level, the component common over all notes (‘note’) and contour direction changes.

Table 2

Correlations between perception and imagery, for each parameter found through the decomposition as well as for the full ERP trace. Channels are shown that represent the maximum correlation for each parameter, marked with *.

Channel	Decomposition parameter				ERP data
	First	Beat	Note	Contour dir.	
Cz	0.88*	0.74	0.69	0.66	0.62*
C4	0.84	0.67	0.72	0.73*	0.52
T4	0.84	0.88*	0.74	0.02	0.30
F8	0.46	0.26	0.88*	-0.26	0.30

sound for imagery, the correlations are still high. This supports the idea that the correlations are not likely to be due to early perceptual responses, considering the difference in timbre between the two, as well as the time lag used in the analyses. Table 2 shows the correlations for the channels with maximum correlations for the different parameters as well as for the original data.

4. Discussion

A method has been described to decompose ERP data, illustrated with measurements of perception and imagery of simple melodies. As opposed to analysis methods that contrast multiple conditions, we here isolate responses to specific

subcomponents of the processes we are interested in, exploiting a priori knowledge of the stimulus.

When testing a model based on a multi-levelled rhythmic aspect combined with the melodic contour direction, the parameters found through the decomposition explain a substantial amount of the variance, slightly less so for imagery than for perception but highly significant in both. Rhythmic aspects were found to be more easily isolated than melodic or pitch-driven aspects. Even though the beat-level component did not reach significance during perception, the explained variance in imagery shows subjective rhythmicization of the identical sounds presented to provide a time lock. We suspect that at least some of the residual signal is also related to aspects that were not taken up into the model. A much larger number of trials would be needed to investigate this further. The most variance was explained by the lowest rhythmic level, partialing out what is common over all notes in the melody. For perception this relates to the processing of a midi-piano note, for imagery it relates to the high-hat sound used to timelock the task. This means that, aside from the low-level rhythmic processing, the auditory response to this sound is also captured in this parameter, resulting in a component that cannot necessarily be considered musical. Although other means of synchronization could have been used (i.e. a flashing led), the low-level rhythmic processing would always overlap with the timekeeper.

The parameters found through the decomposition do not only model the contribution of a musical aspect to the ERP signature,

but also allow comparisons across tasks (i.e. perception and imagery). A high correlation between the parameters found for different tasks also supports the idea that the decomposition isolates aspects of the brain response that are related to the musical aspect under consideration. As with statistically separated components or calculated dipoles, the parameters found through decomposition are not physically isolable. As such their physical shapes are hard to compare with ERP responses previously found for the melodic aspects included in our model. In order to fully investigate the effects of these aspects, stimuli would need to be constructed that include comparable numbers of trials of each aspect or level of aspect, likely to produce melodies that sound less natural than the stimuli used here. That said, the decomposition does offer a look at EEG signatures that could not otherwise be revealed, and the result showing the 'Beat'-level to explain variance in the Imagery condition corresponds well with previous results on subjective rhythmization (Brochard et al., 2003). The contour component, which only approaches significance for perception, but does not at all for imagery, shows a very interesting localization of the correlation between tasks, corresponding well with areas which also distinguish between different perceived and imagined music stimuli (Schaefer et al., 2008).

The decomposition method offers a new possibility of investigating complex ERP patterns and allows processing components to be visualized within a continuous context. This decomposition method assumes that the EEG traces of the processing of each aspect of the stimulus combine additively to form the whole EEG signature of a note. This is a reasonable hypothesis if the subcomponents originate from different source locations, all measured at each electrode. However, the method also assumes independent contributions from each subcomponent (but see Boltz (1989) for interactions in behavioral responses). The assumption of modularity of music processing in the brain is not new, and has been elaborately described by Peretz and Coltheart (2003). The concept of additive responses is already present in general ERP research as well, when interpreting the averaged waveforms as containing responses related to different subcomponents of processing. In few cases, Independent Component Analysis (c.f. Makeig et al., 1997) or Principal Component Analysis (c.f. Dien et al., 2003) has been used to separate the different brain responses. However, these methods are blind to the stimulus and do not allow structuring according to what is known about the stimulus or task, needing post hoc interpretation of the resulting components. By creating a model based on what we are interested in, the predictive value of that particular process or stimulus component on the data can be assessed.

A number of interesting questions are still left unanswered here. Notably the effect of musical expertise, which was not added as a variable in the current design, could influence the latency and localization of musical aspects (e.g. for rhythm see Vuust et al., 2005). Additionally, the effects of memory processes are unclear in the current paradigm. Working memory must have a big effect initially, when internally rehearsing unknown melodies, whereas after a few repetitions, this is probably less of an issue. However, we now did not have enough trials to investigate this effect over time as the experiment progressed.

Here, the data segment used to find the contribution of a musical aspect was lagged behind to ensure that a response to a previous note was minimally present. The model could alternatively be extended to deal with overlapping responses, at the cost of an increased danger of redundancy in the model. Other additions might be to extend the decomposition to regress spatially as well as temporally, by recording high-density EEG with more channels. Also, time-frequency analysis may reveal more patterns by extending the features to non-time locked activity as well as time locked. To further investigate the usefulness of this method, different types of stimuli and tasks are needed.

In conclusion, a method was presented that allows decomposition of continuous EEG signals, using multiple linear regression combined with a structural model of the stimulus or task. This method offers a more precise look at EEG signatures of subprocesses that cannot be disentangled by looking at the ERPs. Although further investigation is necessary, this method expands the possibilities of analyzing electrophysiological measurements as well as opening up possible paradigms for investigating neural responses.

Acknowledgements

The authors thank Logan Grosenick and Tim Uy at the Suppes Brain Lab for collecting the data, Barbara Tillmann and Olaf Hauk for very useful comments on earlier versions of the manuscript, Elizabeth Margulis for the stimuli and Jason Farquhar for help with the formal description of the method.

References

- Bharucha, J., Krumhansl, C.L., 1983. The representation of harmonic structure in music—hierarchies of stability as a function of context. *Cognition* 13 (1), 63–102.
- Boltz, M., 1989. Perceiving the end: effects of tonal relationships perceiving the end: effects of tonal relationships on melodic completion. *Journal of Experimental Psychology-Human Perception and Performance* 15 (4), 749–761.
- Brattico, E., Tervaniemi, M., Näätänen, R., Peretz, I., 2006. Musical scale properties are automatically processed in the human auditory cortex. *Brain Research* 1117, 162–174.
- Brochard, R., Abecasis, D., Potter, D., Ragot, R., Drake, C., 2003. The "ticktock" of our internal clock: direct brain evidence of subjective accents in isochronous sequences. *Psychological Science* 14 (4), 362–366.
- Cohen, M.S., 1997. Parametric analysis of fMRI data using linear systems methods. *NeuroImage* 6 (2), 93–103.
- Dambacher, M., Kliegl, R., Hofman, M., Jacobs, A.M., 2006. Frequency and predictability effects on event-related potentials during reading. *Brain Research* 1084, 89–103.
- Desain, P., Blankespoor, J., Farquhar, J., Gielen, C.C.A.M., 2008. Detecting spread spectrum pseudo random noise tags in eeg/meg using a structure-based decomposition. In: *Proceedings of the 4th International Brain-Computer Interface Workshop*, Graz, Austria.
- Desain, P., Honing, H., 2003. Single trial erp allows detection of perceived and imagined rhythm. In: *Proceedings of the RENCON Workshop, International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 1–4.
- Dien, J., Frishkoff, G.A., Cerbone, A., Tucker, D.M., 2003. Parametric analysis of event-related potentials in semantic comprehension: evidence for parallel brain mechanisms. *Cognitive Brain Research* 15, 137–153.
- Dowling, W.J., Kwak, S.Y., Andrews, M.W., 1995. The time-course of recognition of novel melodies. *Perception and Psychophysics* 57 (2), 136–149.
- Dowling, W.J., Tillmann, B., Ayers, D.F., 2001. Memory and the experience of hearing music. *Music Perception* 19 (2), 249–276.
- Glover, G.H., 1999. Deconvolution of impulse response in event-related BOLD fMRI. *NeuroImage* 9, 416–429.
- Halpern, A.R., Martin, J.S., Reed, T.D., 2007. An ERP study of major–minor classification in melodies. *Music Perception* 25 (3), 181–191.
- Halpern, A.R., Zatorre, R.J., Bouffard, M., Johnson, J.A., 2004. Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia* 42 (9), 1281–1292.
- Hauk, O., Davis, M.H., Ford, M., Pulvermüller, F., Marslen-Wilson, W.D., 2006. The time course of visual word recognition as revealed by linear regression analysis of ERP data. *NeuroImage* 30, 1383–1400.
- Hauk, O., Pulvermüller, F., Ford, M., Marslen-Wilson, W.D., Davis, M., 2009. Can I have a quick word? Early electrophysiological manifestations of psycholinguistic processes revealed by event-related regression analysis of the EEG. *Biological Psychology* 80 (1), 64–74.
- Janata, P., 2001. Brain electrical activity evoked by mental formation of auditory expectations and images. *Brain Topography* 13 (3), 169–193.
- Janata, P., Birk, J.L., Van Horn, J.D., Tillmann, B., Bharucha, J.J., 2002. Cortical topography of tonal structures underlying western music. *Science* 298, 2167–2170.
- Jasper, H.H., 1958. The ten-twenty electrode system of the international federation. *Electroencephalography and Clinical Neurophysiology* 10, 371–375.
- Jones, M.R., Boltz, M., 1989. Dynamic attending and responses to time. *Psychological Review* 96 (3), 459–491.
- Jones, M.R., Ralston, J.T., 1991. Some influences of accent structure on melody recognition. *Memory and Cognition* 19 (1), 8–20.
- Jongsma, M.L.A., Desain, P., Honing, H., 2004. Rhythmic context influences the auditory evoked potentials of musicians and nonmusicians. *Biological Psychology* 66 (2), 129–152.

- Koelsch, S., Gunter, T., Schröger, E., Friederici, A.D., 2003. Processing tonal modulations: an ERP study. *Journal of Cognitive Neuroscience* 15 (8), 1149–1159.
- Kraemer, D.J.M., Macrae, C.N., Green, A.E., Kelley, W.M., 2005. Sound of silence activates auditory cortex. *Nature* 434, 158.
- Kristeva, R., Chakarov, V., Schulte-Mönting, J., Spreer, J., 2003. Activation of cortical areas in music execution and imagining: a high-resolution EEG study. *NeuroImage* 20, 1872–1883.
- Krohn, K.I., Brattico, E., Valimäki, V., Tervaniemi, M., 2007. Neural representations of the hierarchical scale pitch structure. *Music Perception* 24 (3), 281–296.
- Lerdahl, F., Jackendoff, R., 1983. *A Generative Theory of Tonal Music*. MIT Press, Cambridge, Mass.
- Makeig, S., Jung, A., Bell, D., Ghahremani, A.J., Sejnowski, T., 1997. Blind separation of auditory event-related brain responses into independent components. *Proceedings of the National Academy of Sciences of the United States of America* 94, 10979–10984.
- Meyer, M., Elmer, S., Baumann, S., Jancke, L., 2007. Short term plasticity in the auditory system: differential neural responses to perception and imagery of speech and music. *Restorative Neurology and Neuroscience* 25, 411–431.
- Neuhaus, C., Knösche, T.R., Friederici, A.D., 2006. Effects of musical expertise and boundary markers on phrase perception in music. *Journal of Cognitive Neuroscience* 18 (3), 472–493.
- Peretz, I., Coltheart, M., 2003. Modularity of music processing. *Nature Neuroscience* 6 (7), 688–691.
- Schaefer, R.S., Desain, P., 2006. Probing the representations of melody: an ERP study. In: Baroni, M., Addesi, A., Caterina, R., Costa, M. (Eds.), *Proceedings of the 9th International Conference on Music Perception and Cognition (ICMPC9)*, Bologna, Italy.
- Schaefer, R.S., Guimaraes, M.P., Desain, P., Suppes, P., 2008. Detecting imagined music from EEG. In: Miyazaki, K., Hiraga, Y., Adachi, M., Nakajima, Y., Tsuzaki, M. (Eds.), *Proceedings of the 10th International Conference of Music Perception and Cognition*. Causal Productions, Sapporo, Japan.
- Scherg, M., Vajsar, J., Picton, T.W., 1989. A source analysis of the late human auditory evoked potentials. *Journal of Cognitive Neuroscience* 1 (4), 336–355.
- Shahin, A., Roberts, L.E., Pantev, C., Trainor, L.J., Ross, B., 2005. Modulation of P2 auditory-evoked responses by the spectral complexity of sounds. *NeuroReport* 16 (16), 1781–1785.
- Tillmann, B., Bigand, E., 2001. Global context effect in normal and scrambled musical sequences. *Journal of Experimental Psychology-Human Perception and Performance* 27 (5), 1185–1196.
- Trainor, L.J., McDonald, K.L., Alain, C., 2002. Automatic and controlled processing of melodic contour and interval information measured by electrical brain activity. *Journal of Cognitive Neuroscience* 14 (3), 430–442.
- Vuust, P., Pallesen, K.J., Bailey, C., van Zuijen, T., Gjedde, A., Roepstorff, A., Østergaard, L., 2005. To musicians, the message is in the meter: preattentive neuronal responses to incongruent rhythm are left-lateralized in musicians. *NeuroImage* 24, 560–564.
- Windsor, L., Desain, P., Penel, A., Borkent, M., 2006. A structurally guided method for the decomposition of expression in music performance. *JASA* 119, 1182–1193.
- Yamamoto, M., Matsuda, M., Itoh, K., Uno, A., Karino, S., Saitoh, O., Kaneko, Y., Yatomi, Y., Kaga, K., 2005. Auditory imagery mismatch negativity elicited in musicians. *NeuroReport* 16 (11), 1175–1178.