Title: What makes medical students better listeners?

Authors: De Meo R, Matusz PJ, Knebel JF, Murray MM, Thompson WR, Clarke S

Journal: Current biology : CB

Year: 2016 Jul 11

Issue: 26

Volume: 13

Pages: R519-20

DOI: 10.1016/j.cub.2016.05.024

Current Biology

What makes medical doctors better listeners?

--Manuscript Draft--

Manuscript Number:	
Full Title:	What makes medical doctors better listeners?
Article Type:	Correspondence
Keywords:	Auditory recognition; EEG; human
Corresponding Author:	Stephanie Clarke CHUV Lausanne, SWITZERLAND
First Author:	Rosanna De Meo
Order of Authors:	Rosanna De Meo
	Pawel J Matusz
	Jean-François Knebel
	Micah M Murray
	W. Reid Thompson
	Stephanie Clarke, MD
Suggested Reviewers:	Simon Thorpe, Prof. Centre de Recherche Cerveau et Cognition, Toulouse, France thorpe@cerco.ups-tlse.fr Expert on auditory and visual recognition
	Charles E. Schroeder, Prof. Columbia University, USA cs2388@columbia.edu Expert on auditory cognition and cortical processing of auditory information
	Jan Schnupp, Prof. Oxford University, UK jan.schnupp@dpag.ox.ac.uk Expert on auditory processing
	Olivier Collignon , Prof. University of Trento, Italy oli.collignon@gmail.com Expert on auditory and multisensory cognition
	Lee Miller, Prof. Center for Brain and Mind, UC Davis, USA leemiller@ucdavis.edu Expert on auditory and multisensory processing
	Riita Hari, Prof. Brain Research Unit, Aalto University, Finland riita.hari@aalto.fi Expert on human brain recordings and on auditory processing
	Asif Ghazanfar, Prof. Developmental Neuromechanics and Communication lab, Princeton University, USA asifg@princeton.edu Expert on auditory and multisensory processing in primates
	Chris Petkov, Prof. Institute of Neuroscience, Newcastle University Medical School, UK chris.petkov@ncl.ac.uk Expert on auditory processing

Opposed Reviewers:	

What makes medical doctors better listeners?

Rosanna De Meo[1], Pawel J. Matusz[2-3], Jean-François Knebel[1,2,4], Micah M. Murray[1,2,4-6], W. Reid Thompson[7] and Stephanie Clarke[1,†].

[1] Service of Neuropsychology and Neurorehabilitation, Department of Clinical Neuroscience, University Hospital Center and University of Lausanne, Switzerland

[2] The Laboratory for Investigative Neurophysiology (The LINE), Department of Radiology and Department of Clinical Neurosciences, University Hospital Center and University of Lausanne, Switzerland

[3] Attention, Brain & Cognitive Development Group, Department of Experimental Psychology, University of Oxford, United Kingdom

[4] EEG Brain Mapping Core, Center for Biomedical Imaging (CIBM), Lausanne, Switzerland

[5] Department of Hearing and Speech Sciences, Vanderbilt University Medical Center, Nashville, TN, USA

[6] Department of Ophthalmology, Jules-Gonin Eye Hospital, Lausanne, Switzerland

[7] Division of Pediatric Cardiology, Department of Pediatrics, Johns Hopkins University School of Medicine, Baltimore, MD, USA

†Address correspondence to SC (Stephanie.Clarke@chuv.ch)

Diagnosing heart conditions by auscultation with a stethoscope is an important clinical skill, commonly learnt by medical students. However, clinical proficiency for this skill is in decline [1], and new teaching methods are needed. Successful discrimination of heartbeat sounds is believed to benefit mainly from acoustical training [2]. It is currently unknown whether the neural representation of the meaning contributes to successful discrimination. If so, teaching of auscultation skills should emphasise the link between the heartbeat sound and its meaning. Beyond cardiac auscultation, this issue is of interest for all fields where subtle, but complex perceptual differences identify items in a well-known semantic context.

Previous electrophysiological studies in humans highlighted mechanisms whereby a brief training improved difficult auditory discrimination. Learning to recognize bird species by their songs sharpened the activation of neural representations of the learned species within regions known to be involved in semantic processing [3]. Similarly, auditory spatial discrimination in near-threshold conditions critically involved early stages of cortical processing within auditory areas by increasing the encoding of spatial positions in successful discrimination runs [4].

Thus, successful discrimination of heartbeat sounds may be associated with distinct encoding at later stages of cortical processing with the sharpening of a semantic representation. To address this issue we recruited medical students, who had solid knowledge of cardiology at the time of testing, and analyzed their auditory evoked potentials (AEPs) to sounds of four different types of heartbeats. One cardiac cycle

comprises two consecutive sounds attributed to the closing of the mitral and aortic valves, respectively. The four conditions used in our study differed in the nature of the first or second sound, which could be simple or “split” and indicated the state of the heart (Supplemental Information). For each of the four conditions, auditory recordings from several patients were used, representing the natural variation present in clinical practice. The task was to identify the condition portrayed by each sound, based on recordings of four heart cycles. Prior to EEG recording, subjects practiced recognizing the four types of heartbeats until reaching 70% accuracy. During the EEG recording subjects listened to a new set of stimuli (involving different exemplars of the same four types of heartbeats) and assigned them to the corresponding category (Fig. 1A; Fig. S1). Mean accuracy in the categorization task was below 70% but statistically above the chance level and remained so for the whole recording session; response times were faster for correct than incorrect responses (Supplemental Information).

Electrical neuroimaging analyses of the AEPs analyzed brain activity that was elicited by the second heart cycle by comparing trials with correct (Hits) vs. incorrect identifications (Misses), regardless of the specific type of heartbeat (Fig. 1A-B; Supplemental Information). Analysis of AEPs across the electrode montage revealed a significant difference over the time period of 280-310ms after the onset of the second heart cycle. Estimations of brain sources of differences observed during this period revealed greater activity for Misses than Hits in the left middle frontal gyrus (MFG) and in the right prefrontal cortex. Of note, no such differences were present when the same analysis was performed on AEPs evoked by the first heart cycle of the excerpt, i.e., before the subjects heard sufficient acoustic information to label each heartbeat sound (Supplemental Information).

Our results indicate that successful fine-grained discrimination, as those of heartbeat sounds, relies critically on labelling, which is provided by semantic, rather than acoustic representations (Fig. 1C). Several lines of evidence support this interpretation. First, successful heartbeat discrimination modulated the same MFG region as previous studies on semantic categorization in humans [5-6] and non-human primates [7]. Similarly, intracranial recordings revealed differential activity during correct vs. incorrect auditory categorization within frontal, but not supratemporal sites; correct responses being associated with lower high-gamma activity [8]. The same frontal region was associated with expertise in birdsong recognition, again lesser activity being associated with correct recognition [3]. Taken together, these studies highlight the critical contribution of higher-order processing and particularly, of the sharpening of neural representations within the semantic “expert” network as a necessary step for correct recognition. When a percept activates a sharply tuned semantic representation, it is likely to be correctly recognized whereas a broader activation tends to be associated with incorrect recognition.

Second, differences in neural activity between correct and incorrect heartbeat categorization occurred during the same period that characterized semantic processing in previous studies. Semantic processing starts as early as 70 ms post-stimulus onset with segregation of broad semantic categories, whereas more narrow categories are discriminated at subsequent stages [9]. Within-category discriminations [3] are known to occur at later time-periods, which partially overlap with the critical time-period revealed here. As highlighted by the temporal hierarchy model of auditory object recognition [9], this critical period is well embedded in the semantic processing sequence.

Third, perceptual processing within early-stage auditory areas was shown to play a role in auditory spatial discrimination [4] and in learning to categorize novel sounds [10], in two instances without a semantic link. In our task, differences in neural processing at the level of early-stage auditory areas were not critical to successful discrimination. Control analyses confirmed that our methods were sufficiently sensitive to detect acoustic differences across heartbeat types (Supplemental Information), since AEPs to acoustically different heartbeat sounds (simple vs. split), analyzed independently of whether they were correctly recognized or not, yielded differential activity at early latencies (45-55 ms) within left auditory cortices.

In conclusion, successful discrimination of heartbeat sounds depends critically on the access to the neural representations of their meaning. To our knowledge, this is the first demonstration that fine-grained discrimination of sound objects, with only minor differences in auditory features, can be enhanced by labelling provided by the semantic link. Beyond the conceptual importance, our findings are of practical relevance to auscultation training, where emphasis should be put on linking the sound and its meaning. The same approach is likely to be beneficial in other instances, such as visual recognition in radiological analysis or in relearning to recognize environmental sounds after cochlear implants.

Acknowledgments

We thank Mr. Gilles Weber for high quality photographs. Financial support was provided by the Swiss National Science Foundation (grants 320030_159707 to SC and 320030-149982 to MMM as well as the National Center of Competence in Research project “SYNAPSY, The Synaptic Bases of Mental Disease” [project 51AU40_125759]).

References

[1] Mangione, S. (2001). Cardiac auscultatory skills of physicians-in-training: a comparison of three English-speaking countries. The American journal of medicine, 110(3), 210-216.

[2] Butter, J., McGaghie, W. C., Cohen, E. R., Kaye, M., & Wayne, D. B. (2010). Simulation-based mastery learning improves cardiac auscultation skills in medical students. Journal of general internal medicine, 25(8), 780-785.

[3] De Meo, R., Bourquin, N. M. P., Knebel, J. F., Murray, M. M., & Clarke, S. (2015). From bird to sparrow: Learning-induced modulations in fine-grained semantic discrimination. NeuroImage, 118, 163-173.

[4] Spierer, L., Murray, M. M., Tardif, E., & Clarke, S. (2008). The path to success in auditory spatial discrimination: electrical neuroimaging responses within the supratemporal plane predict performance outcome. Neuroimage, 41(2), 493-503.

[5] Seger, C. A., & Miller, E. K. (2010). Category learning in the brain. Annual review of neuroscience, 33, 203.

[6] Liebenthal, E., Desai, R., Ellingson, M. M., Ramachandran, B., Desai, A., & Binder, J. R. (2010). Specialization along the left superior temporal sulcus for auditory categorization. Cerebral Cortex, 20(12), 2958-2970.

[7] Cohen, Y. E., Hauser, M. D., & Russ, B. E. (2006). Spontaneous processing of abstract categorical information in the ventrolateral prefrontal cortex. Biology letters, 2(2), 261-265.

[8] Steinschneider, M., Nourski, K. V., Rhone, A. E., Kawasaki, H., Oya, H., & Howard III, M. A. (2014). Differential activation of human core, non-core and auditory-related cortex during speech categorization tasks as revealed by intracranial recordings. Frontiers in neuroscience, 8.

[9] De Lucia, M., Clarke, S., & Murray, M. M. (2010). A temporal hierarchy for conspecific vocalization discrimination in humans. The Journal of Neuroscience, 30(33), 11210-11221.

[10] Ley, A., Vroomen, J., Hausfeld, L., Valente, G., De Weerd, P., & Formisano, E. (2012). Learning of new sound categories shapes neural response patterns in human auditory cortex. The Journal of Neuroscience, 32(38), 13273-13280.

Figure

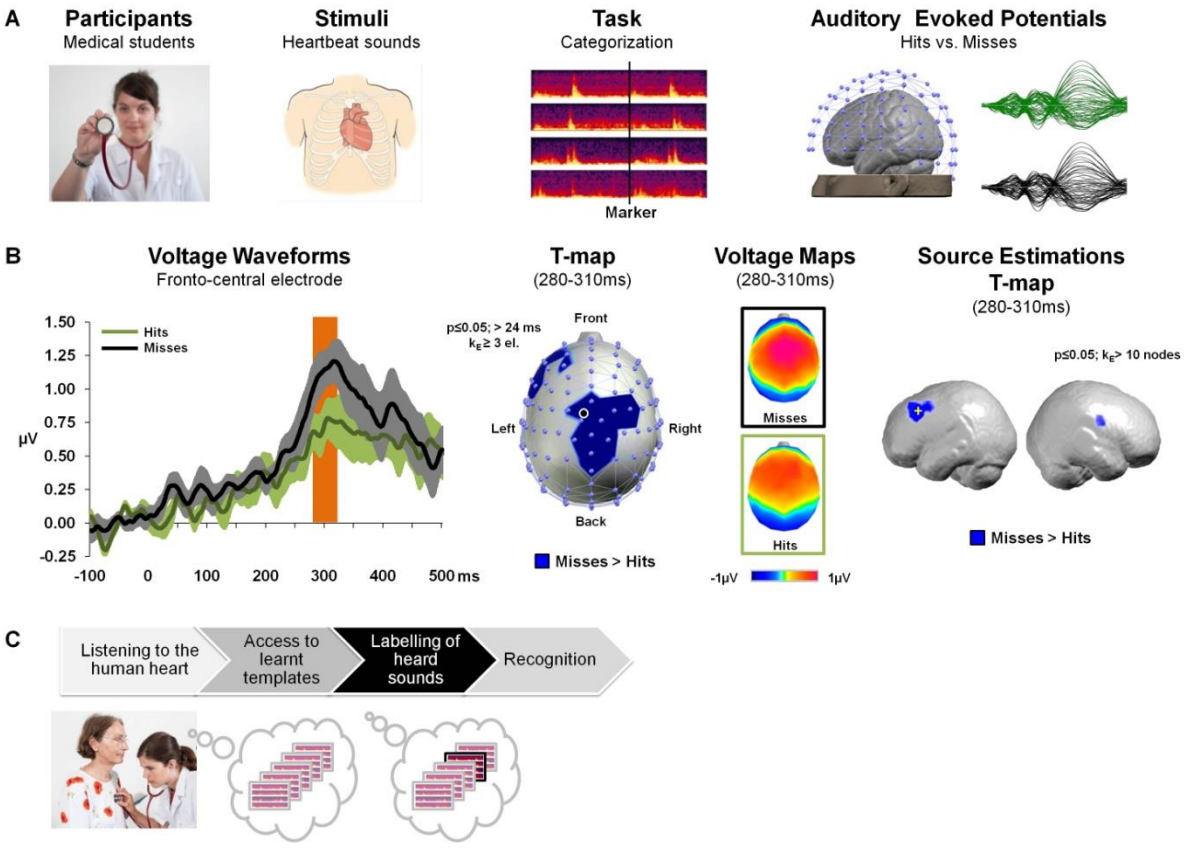
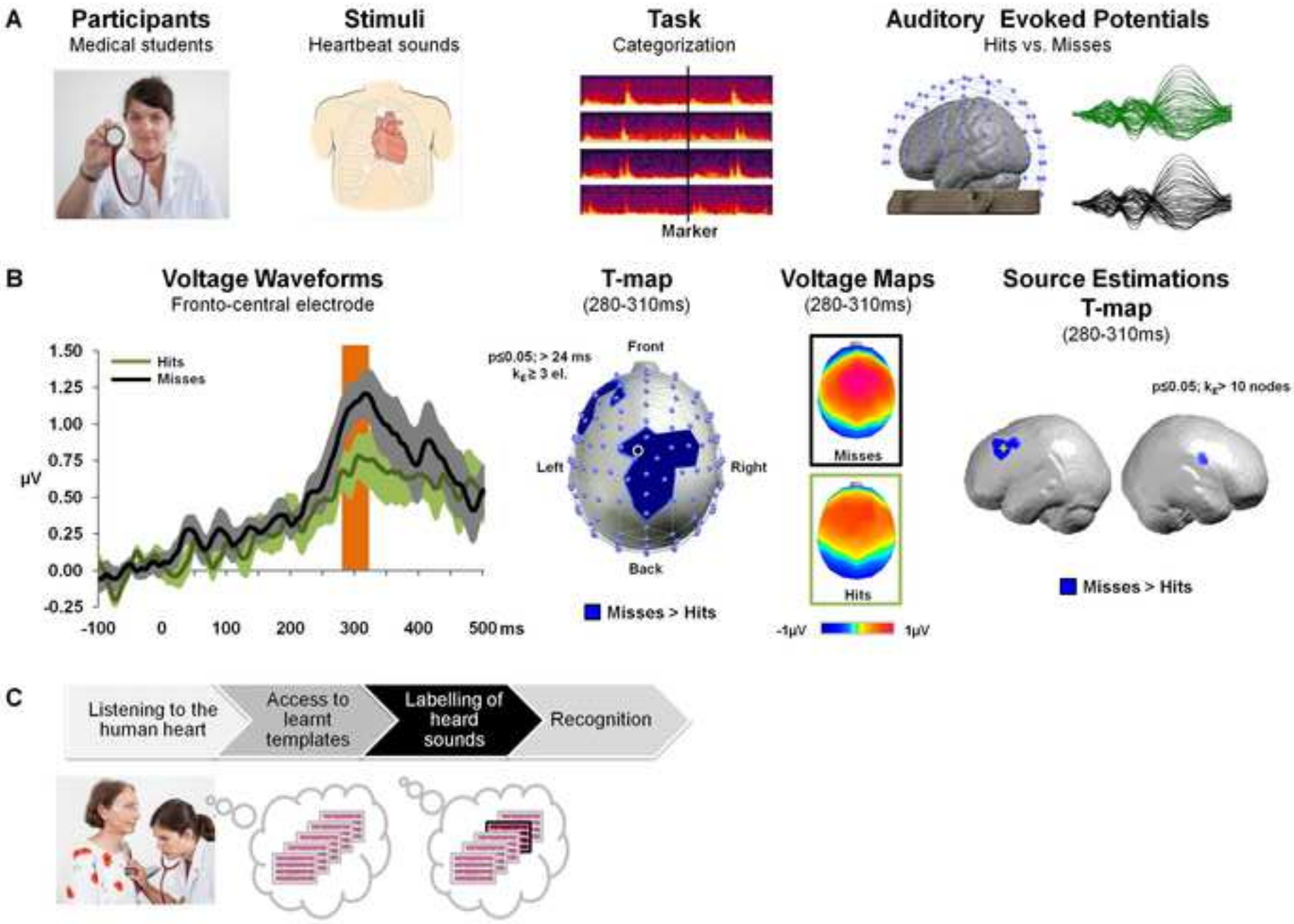


Figure 1: Paradigm and main results. A) General procedure. Medical students were trained to recognize four different heartbeat sounds. Brain activity during the categorization task was recorded with the electroencephalogram (EEG). Auditory evoked potentials (AEPs) analyses focused on brain responses relative to correct (Hits) and incorrect (Misses) recognition and were calculated according to each participant's behavioural responses. New markers were created offline and marked the beginning of the second cardiac cycle. This represents the first point in time when acoustical differences were available to discriminate the four heart sounds. Further details of the paradigm and the electrophysiological analyses are provided in the Supplemental Information. **B) Main results.** AEPs resulting from trials for correct (Hits, in green) and incorrect (Misses, in black) recognition are shown from an exemplar electrode, here Cz (mean ± s.e.m shown). Activity between the two experimental conditions was statistically compared across the electrode montage using a millisecond-by-millisecond paired t-test (with a spatial criterion of at least 3 contiguous electrodes, a temporal criterion of a minimum of 24 ms and a $p \leq 0.05$). Time period of interest were defined when more than 10% of the electrodes, here at least 13 electrodes, showed a significant difference using the above mentioned criteria (indicated by the orange rectangle). The difference was observed over the 280-310 ms post-marker period. Statistical t-maps are displayed on the electrode montage and show the spatial distribution of the electrodes resulting significant over the 280-310 ms post-marker period. All electrodes showed a higher activity for incorrect (Misses) as compared to correct (Hits) recognition. Voltage maps for both experimental conditions are averaged over the 280-310 ms time-window and displayed on the electrode montage. Significant differences in distributed source estimations within the previously defined time period were observed within the left middle frontal gyrus (MFG) and right pre-/postcentral gyrus (PoG/PrG). Statistical t-maps are displayed on the average MNI brain ($p \leq 0.05$, $k_E \geq 10$ nodes). Both clusters showed a higher activity for incorrect (Misses) as compared to correct (Hits) recognition. The maximal t-value (yellow +) was located in the left MFG cluster. **C) Model.** Auditory fine-grained discrimination requires successive steps: listening to the heartbeat, the access to the learnt templates, the access of heard sound and finally the recognition of the heartbeat sound. Our results showed that unsuccessful recognition of heartbeat sounds is due to limitations in higher-order processes such as labelling rather than limitations in low-level processes, i.e. at the perceptual level.

Figure

Click here to download Figure DeMeo et al._Current biology_main figure.jpg



Participants

Thirteen 4th year and above medical students of the University of Lausanne participated in the experiment after providing written informed consent. Two participants were excluded due to poor performance during the training session, leaving eleven participants (7 females; aged 22-31, mean age±SD = 26±0.8 years, one left-handed, all native French-speakers) in the analyses reported here. The procedures of this study were approved by the Ethics Committee of the University of Lausanne. None of the participants had a history of neurological or psychiatric illnesses, and all reported normal hearing. All individuals received a basic training in auscultation during their Bachelor's degree in Medicine. The possible exposures to auscultation during internships varied between participants but did not affect the discrimination of heartbeat sounds (for example, 6th year students did not perform better than 4th year students at baseline). None of the participants had no particular musical training.

When they participated in our study, medical students already had solid knowledge about the heart and its function. They knew the events that produce the two sounds during a healthy heartbeat, the physiological basis of the two normal variants (i.e. the influence of the respiratory cycle on them). They understood also the pathological variants: their anatomical and physiological basis and their treatments, and were familiar with other diagnostic procedures and methodology (MRI, echocardiographic imaging, etc). They have also examined autopsy and surgical specimens and have seen numerous illustrations of those. Thus, they had vast semantic knowledge within which they could and certainly did interpret the acoustical percept.

Stimuli

Stimuli consisted of heartbeat sounds. Acoustically, the normal cardiac cycle is composed of two sounds, S1 and S2, which correspond to the closure of the mitral and aortic valves, respectively. The time lapse between S1 and S2 corresponds to systole, during which blood is ejected from the ventricles. Then, the time lapse between S2 and the next S1 is diastole, during which the blood is filling the ventricles [1]. Acoustic changes in these sounds can indicate physiological variation or be signs of pathological states. Additional sounds may occur during systole, including the *early systolic click* (ESC), due to the opening of an abnormal aortic valve. This variant represents a pathological state. S2 consists of two components (A2 and P2), corresponding to the closure of the aortic and pulmonary valves [2]. In normal hearts, these 2 components are split during inspiration and fused during expiration. In contrast, when both components of S2 remain constantly split (wide split of S2, WS), the variant is pathological and is an important sign of congenital defect in the atrial septum. Likewise, the absence of S2 splitting during the breathing cycle is also pathologic, representing possibly pulmonary hypertension or absence of either the pulmonary or aortic valve [3].

The heart sounds were taken from a multimedia database of digital recordings from a large cohort of patients examined at the Johns Hopkins Outpatient Pediatric Cardiology Clinic (http://murmurlab.org). This database contains pathological as well as normal and innocent variants of heartbeat sounds [4]. A total of 40 exemplars representing 2 full cardiac cycles, repeated twice, were selected by one of the authors (WRT) and then processed by Adobe Audition (16 bit mono; 22'050 Hz digitization; 2.2 seconds duration; 50 ms rise and decay time to avoid clicks). The dataset was divided into 4 categories of 10 sounds each: i) normal variable split of S2 (VS); ii) both S2 are single (BS); iii) both S2 are widely split (WS); and iv) early systolic click (ESC). Two sets of sounds were created (Figure S1) with one used during the Training session (set 1, 5 sounds in each category) and the other used during the Categorization session (set 2, 5 sounds in each category). The same sounds were used for all participants during training and the categorization session.

Procedure and Task

Subjects performed a continuous categorization task, which required the labeling of four variants of heart sounds. That is, on each trial a sound was presented twice and required to be categorized as one of the four variants. To ensure comparable performance between participants during the categorization task, all were trained beforehand on a subset of sounds until they reached a minimal accuracy rate of 70%. The procedure used during the training and the categorization task is similar as in [5].

The training session consisted of alternated blocks of learning, where participants passively listen to heart sound variants, and of training assessments, similar to the categorization task. In the latter, sounds were presented twice and participants were asked to match them with the correct category. Participants received feedback on their responses only during the training assessments. A minimum of 3 training assessments was required. The training session ended with the first block where the minimal performance threshold was reached. Participants who did not have 70% of accuracy after the 7th assessment block were excluded from the experiment (here, 2 participants). On average, participants needed 5 assessments to reach the performance threshold (range [3, 7]). Only behavioral data were recorded in this session.

During the categorization task, each sound was presented for 2.2 seconds, followed by a randomized inter-stimulus interval (ISI) ranging from 500 and 800 ms (with steps of 100 ms). Sounds were presented in a randomized order within each block. Categorization blocks consisted of 60 trials each, equally divided between the four cardiac events (i.e. 15 trials each, with 5 sounds repeated three times for each condition). All participants performed a total of 12 blocks. Behavioral and electrophysiological data were recorded during this session.

The experiment took place in a sound-attenuated chamber, where participants were seated in front of a computer monitor. The auditory stimuli were presented over inserted earphones (ER-4P; www.etymotic.com) and the volume was adjusted to a comfortable level (~ 70dB) (sound level meter: CESVA SC160). E-Prime (Psychology Software Tools, Inc.; www.psnet.com) was used to control stimulus delivery and to record behavioral responses.

EEG acquisition and pre-processing

Continuous EEG was acquired from 128 scalp electrodes (sampling rate of 1024 Hz) using a Biosemi ActiveTwo AD-Box. The origin of each epoch corresponded to the beginning of the third beat (i.e. first beat of the second cardiac cycle, see marker on Figure 1A), which corresponds to the first point in time where all cardiac events can potentially be discriminated. Because our interest was mainly in brain responses related to participants' accuracy, for each trial new markers were generated offline according to participant's behavioral responses. Data pre-processing and analyses were performed using CARTOOL ([6]; http://

sites.google.com/site/fbmlab/cartool). Auditory evoked potentials (AEPs) were calculated by averaging epochs spanning from 100 ms pre-marker to 500 ms post-marker onset for the eight experimental conditions (i.e. hits and misses for the four cardiac events) for each subject. A ± 60 μV artifact rejection criterion was applied for each trial and epochs containing eye blinks and other artifacts were removed. The continuous EEG were filtered (non-causal second order Butterworth with -12db/octave roll-off; 0.1Hz high-pass, 40Hz low-pass; 50Hz notch). The filters were computed linearly in both forward and backward directions to eliminate phase shifts. Single trials were baseline corrected using the pre-marker period. Data from artifact electrodes of each subject were interpolated using 3-D splines [7]. Prior to group-averaging, the number of accepted trials within each participant was equalized. The total number of accepted EEG epochs for each condition was equated within each participant using the condition with the fewest number of accepted trials as the anchor (mean (±SE) 260±13, [range: 132, 373]) (similar procedure as in [5]). In the final step, data were re-calculated using an average reference.

Electrical neuroimaging analyses

The AEP analyses were based on the hypothesis that a differential neural response would be found between correct (Hits) and incorrect (Misses) recognition of heartbeat sounds. Furthermore, we hypothesized that this differential neural response occurs after the first point in time when cardiac events are perceptually different. These electrical neuroimaging analyses allowed us to define time periods showing sustained modulations in the activity of small clusters of electrodes, as well as alterations in the configuration of brain generators by visualizing averaged AEP topographies at the scalp. Differences observed between AEPs were used for defining time periods for the analysis of source estimations [8-9]. Source estimations were calculated using the local auto-regressive average distributed linear inverse solution (LAURA, [10-11]. The time period of interest was selected on the basis of sustained differences across the two conditions at clusters of electrodes. Activity for both conditions, Hits and Misses, was statistically compared using a millisecond-by-millisecond paired t-test, with a temporal criterion of ≥ 24 ms and a spatial criterion of ≥ 3 contiguous electrodes and a p-value ≤ 0.05 for significant effects that correct for multiple comparisons [12] using the Statistical Toolbox for Electrical Neuroimaging (STEN; http://unil.ch/line/home/menuinst/about-the-line/software--analysis-tools.html#standard_412). Time periods of interest were defined when more than 10% of the electrodes, here at least 13 electrodes, exhibited a significant difference using the above mentioned criteria. In this experiment, only one time window fulfilled all criteria and was defined as the period of interest for the next steps of the analysis. T-values were calculated at each time-point over the electrode montage. To display the location of electrodes showing a significant effect, their averaged t-values over the defined time period were displayed on a 3-D reconstruction of the electrode montage. Additionally, the averaged voltage maps over the significant time period were displayed for both conditions (Figure 1B).

We estimated the sources underlying the effects observed at the scalp using a distributed linear inverse solution together with the LAURA regularization approach [10-11]. LAURA uses a realistic head model, and the solution space included 3005 nodes, selected from a 6x6x6mm grid of 3005 equally distributed within the grey matter of the Montreal Neurological Institute's (MNI) averaged brain. The results of the above mentioned electrode analysis defined a time-window of interest for which intracranial sources were estimated and statistically compared between both conditions (here 280-310ms post-marker onset). The statistical analysis of source estimations was performed first by averaging the AEP data across time generating a single data point for each participant and condition. Then the inverse solution was estimated for each of the 3005 nodes and submitted to a paired t-test. Differences were considered as reliable if they fulfilled a spatial criterion of at least 10 contiguous significant nodes [as in 5]. The results of the source estimations, t-values, are displayed on the MNI averaged brain in Figure 1B.

Behavioral data analysis

The performance in discriminating the different heartbeat conditions was measured through the assessments and the categorization blocks as percentages of correct responses for each experimental condition. Mean accuracy rates for the four experimental conditions were compared during the last training assessment (LTA) and the categorization task ("CAT"). Since the main interest was in the correct ("Hits") and incorrect ("Miss") recognition of sounds, the mean accuracy rate at LTA was compared with the mean accuracy rate throughout the CAT with a paired t-test. The analysis of reaction times compared time responses between Hits ("H") and Miss ("M") during the categorization task. Because participants had to wait until the end of the sound to answer and were not limited in time to answer, extreme reaction time values were removed. For each participant, only values between percentiles 5 and 95 were selected throughout the categorization task. Mean reaction times for Hits were compared to mean reaction times for Miss with a paired t-test.

Behavioral results

During the Training session participants performed on average (mean ± SE) 5 ± 1 training assessments in order to reach the minimal accuracy threshold of 70%. They significantly improved their performance in associating the heartbeat sounds with the corresponding category: from the baseline assessment (A1) at 47 ± 3.8% (mean ± SE) to the last training assessment (LTA) at 74 ± 1.2%. Mean performance during the categorization task was lower than during the last training assessment which was (mean ± SE) 74 ± 1.2% for the last training assessment and 60 ± 2.8% for the categorization session. When compared with a paired t-test, it showed a significant decrease in performance ($t_{(10)}$ = 5.065; p <0.001). These results suggest that even after a general training in discriminating different cardiac events, their recognition remained difficult. Although, the use of a new set of sounds induces a significant decrease in performance, they remain well above chance level (< 25%). Interestingly, accuracy rates during the categorization task remained constant throughout the whole session. Moreover, we inspected the errors committed during the categorization task. While patterns differed between participants, they nonetheless remained consistent throughout the whole session within each participant. This would suggest that participants commit errors due to semantic confusions rather than because of attentional lapses.

Reaction times were calculated for all categories as a function of accuracy starting at the end of each sound until the button press. Mean reaction times (mean ± SE) 840 ± 140ms for Hits and 977 ± 168ms for Misses and were compared with a paired t-test. Participants were faster when they correctly recognized sounds ($t_{(10)}$ = -4; p = 0.003). Furthermore, mean reaction times for both

conditions (Hits and Misses) are long after the AEPs differences time-window. Thus, AEPs differences are not readily caused by executive functions or movement preparations.

Supplemental electrical neuroimaging analysis 1

This supplemental analysis aimed at identifying the spatio-temporal dynamics of Hits vs. Misses trials before subjects heard acoustical information giving access to the meaning.

This analysis was performed on the same participants and electrophysiological dataset as those described above. The epoch's origin was set at the beginning of each sound, i.e. on the first beat of the first cardiac cycle. The same single trials as those used in the previous analysis were selected and were pre-processed using CARTOOL. AEPs were calculated by averaging epochs spanning from 100ms pre- to 500ms post-stimulus onset for the four cardiac events (VS, BS, WS and ESC) and each subject. The artifact rejection parameters and filters applied were identical to those used in the main analysis. In order to ensure similar signal-to-noise ratios for all conditions, we equalized the number of accepted trials for each condition within each participant. Prior to group-averaging, data from artifact electrode of each subject were interpolated using 3-D splines. Data were re-calculated to an average reference and baseline corrected using the pre-stimulus interval.

In these analyses, we expect that differences between correct (Hits) and incorrect (Misses) recognition of heartbeat sounds would not involve the semantic-"expert" network. Differences observed between AEPs were used to define time periods of interest for the analysis of source estimations. The same procedure as in the main analysis was used to define periods of interest and calculate source estimations. In this analysis, only one time window fulfilled all criteria and was defined as the period of interest for the next steps of the analysis and was between 330-360 ms post-stimulus onset. We estimated the sources underlying the effects observed at the scalp using the LAURA regularization approach [10-11]. Intracranial sources were estimated and statistically compared between both conditions on the averaged abovementioned time-window. The statistical analysis of source estimations was performed first by averaging the AEP data across time generating a single data point for each participant and condition. Then the inverse solution was estimated for each of the 3005 nodes and submitted to a paired t-test. Differences were considered as reliable if they fulfilled a spatial criterion of at least 10 contiguous significant nodes [as in 5].

Differential activity between Hits and Misses at the beginning of each excerpt was found within left occipital regions over the 330-360ms post-stimulus onset (Figure S2). The patterns of differential activity as well as its time-window do not overlap with those found in our main analysis, i.e. in frontal and prefrontal regions over the 280-310 ms post-marker time-window. These results support the idea that successful heartbeat sounds discrimination depends critically on the access to neural representations of meaning rather than attentional mechanisms.

Supplemental electrical neuroimaging analysis 2

This supplemental analysis aimed at identifying the spatio-temporal dynamics of acoustic differentiation. We therefore compared early (within the first 100ms after stimulus onset) neural responses evoked by the four cardiac events, irrespectively of the behavioral responses. The source estimation analyses were based on the hypothesis that differential neural responses would be found within auditory areas at early latencies between the four cardiac events if acoustical differences were perceived across the four sounds.

This analysis was performed on the same participants and electrophysiological dataset as those described above. The epoch's origin was set at the beginning of each sound, i.e., on the first beat of the first cardiac cycle. The same single trials as those used in the previous analysis were selected and were pre-processed using CARTOOL. AEPs were calculated by averaging epochs spanning from 100ms pre- to 200ms post-stimulus onset for the four cardiac events (VS, BS, WS and ESC) and each subject. The artifact rejection parameters and filters applied were identical to those used in the main analysis. In order to ensure similar signal-to-noise ratios for all conditions, we equalized the number of accepted trials for each condition within each participant. Prior to group-averaging, data from artifact electrode of each subject were interpolated using 3-D splines. Data were re-calculated to an average reference and baseline corrected using the pre-stimulus interval. Source estimations were compared millisecond-by-millisecond with a one-way ANOVA with four levels (VS, BS, WS and ESC) and based on the LAURA source estimation method. The array of 3005 nodes was used with a temporal criterion of at least continuous 11ms and a cluster size of at least 10 contiguous nodes.

Differential activity between the four cardiac events was found within the left middle temporal gyrus (MTG) over the ~45-55ms post-stimulus onset (Figure S3). This differential activity in early latencies in the vicinity of auditory areas shows that acoustical differences between the cardiac events are processed at the beginning of each sound. The absence of early differences in auditory areas in our main analysis, i.e. at the beginning of the second cardiac cycle, support the idea that correct and incorrect categorization of sounds do not rely on low-level features per se, but rather on their meaning.

The location of the differential activity in the vicinity of auditory areas in early latencies between the different cardiac events demonstrates that acoustical differences are processed but do not represent a critical step for the categorization of heart sounds. Indeed, we did not find differential activity in early latencies in our main analysis (i.e. based on the third beat) but in later latencies (< 200 ms), which supports limitations in higher order stages of processing, such as labeling, rather than perceptual limitations when sounds are not correctly categorized.

Supplemental References

[1] Efimov, I. R., Huang, D. T., Rendt, J. M., & Salama, G. (1994). Optical mapping of repolarization and refractoriness from intact hearts. Circulation,90(3), 1469-1480.

[2] Barma, S., Chen, B. W., Man, K. L., & Wang, J. F. (2014). Quantitative Measurement of Split of the Second Heart Sound (S2). IEEE/ACM Trans Comput Biol Bioinform, 12(4), 851-860.

[3] Debbal, S. M., & Bereksi-Reguig, F. (2006). Analysis and study of the variation of splitting in the second heartbeat sound of wavelet transform. Journal of medical engineering & technology, 30(5), 298-305.

[4] Tuchinda, C., & Thompson, W. R. (2001). Cardiac auscultatory recording database: delivering heart sounds through the Internet. In Proceedings of the AMIA Symposium (p. 716). American Medical Informatics Association.

[5] De Meo, R., Bourquin, N. M. P., Knebel, J. F., Murray, M. M., & Clarke, S. (2015). From bird to sparrow: Learning-induced modulations in fine-grained semantic discrimination. NeuroImage, 118, 163-173.

[6] Brunet, D., Murray, M. M., & Michel, C. M. (2011). Spatiotemporal analysis of multichannel EEG: CARTOOL. Computational intelligence and neuroscience, 2011, 2.

[7] Perrin, F., Pernier, J., Bertnard, O., Giard, M. H., & Echallier, J. F. (1987). Mapping of scalp potentials by surface spline interpolation. Electroencephalography and clinical neurophysiology, 66(1), 75-81.

[8] Murray, M. M., Brunet, D., & Michel, C. M. (2008). Topographic ERP analyses: a step-by-step tutorial review. Brain topography, 20(4), 249-264.

[9] Michel, C. M., & Murray, M. M. (2012). Towards the utilization of EEG as a brain imaging tool. Neuroimage, 61(2), 371-385.

[10] Grave de Peralta Menendez, R. G., Andino, S. G., Lantz, G., Michel, C. M., & Landis, T. (2001). Noninvasive localization of electromagnetic epileptic activity. I. Method descriptions and simulations. *Brain topography*, *14*(2), 131-137.

[11] Grave de Peralta Menendez, R. G., Murray, M. M., Michel, C. M., Martuzzi, R., & Andino, S. L. G. (2004). Electrical neuroimaging based on biophysical constraints. *Neuroimage*, *21*(2), 527-539.

[12] Guthrie, D., & Buchwald, J. S. (1991). Significance testing of difference potentials. *Psychophysiology*, *28*(2), 240-244.

FIGURES

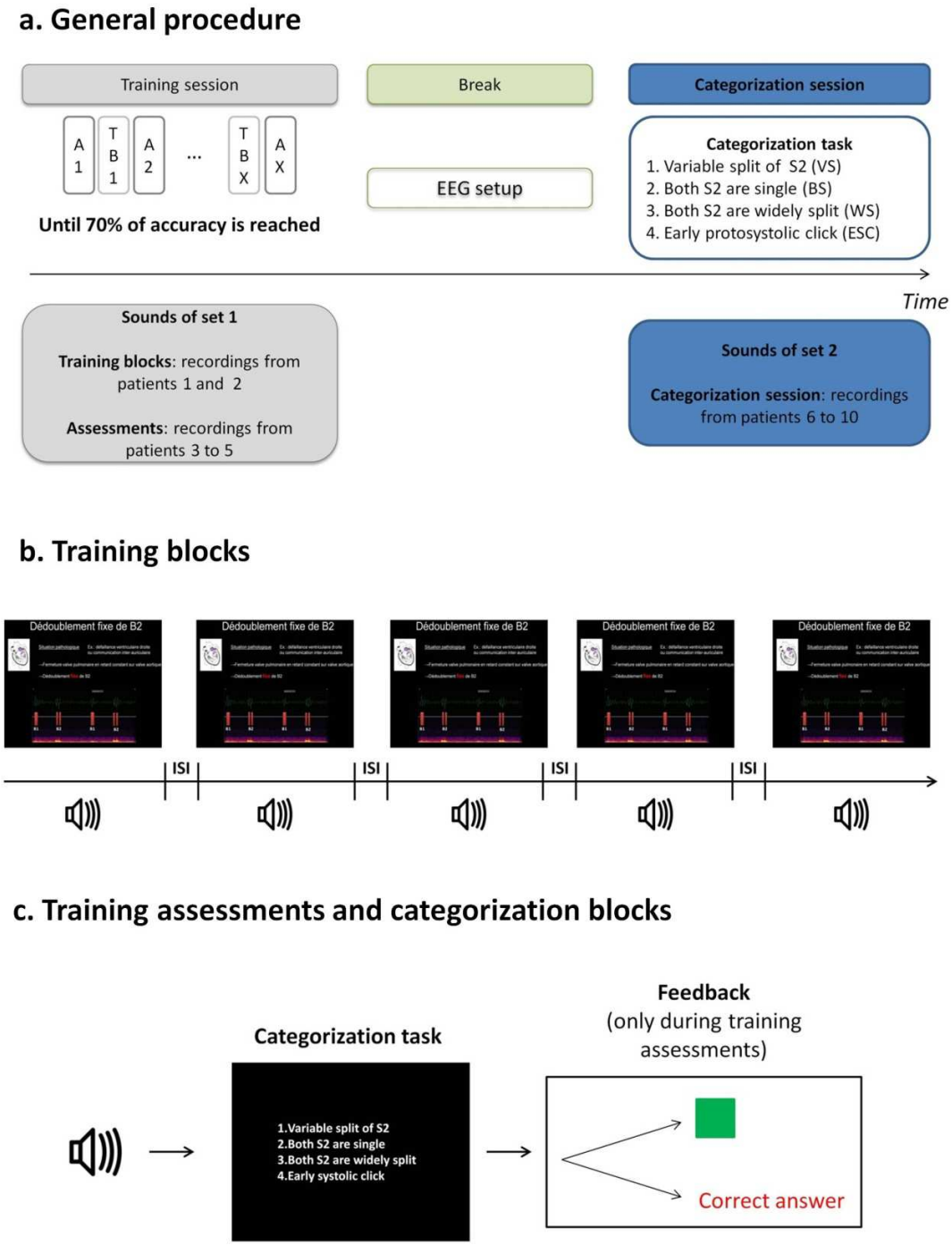


Figure S1: Description of the experimental procedure. a) General procedure. Participants performed two sessions: a *Training session* during which the four types of heart sounds were learned (training blocks, "TB1-TBX") until participants reached a minimum of 70% of accuracy (measured with the assessements, "A1-AX"), and a *Categorization session* requiring the categorization of each heart sound into the corresponding cardiac event. **b) Training blocks.** During the training session, participants passively listen to heart sounds while an detailed explanation of the cardiac event was prensented on the screen. **c) Training assessments and categorization blocks.** Participants' peformance was assessed through the training assessments and the blocks in the categorization session. Each of them consisted of the presentation of the sound (each sound repeated twice) followed by a multiple choice task. Participants received a feedback of their response only during the training assessements (expect for the baseline test).

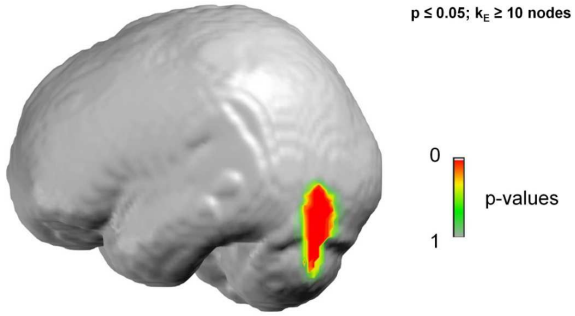


Figure S2: Results supplemental analysis 1, source estimations p-maps (330-360 ms). The results of the paired t-test comparing Hits vs. Misses trials at the beginning of the sound (i.e. at the beginning of the first cardiac cycle) revealed differential activity within the left occipital regions over the 330-360 ms post-stimulus onset period (with the following criteria: cluster size of at least 10 contiguous nodes and a p-value ≤ 0.05). The results of source estimations, p-values, are displayed on the MNI averaged brain.

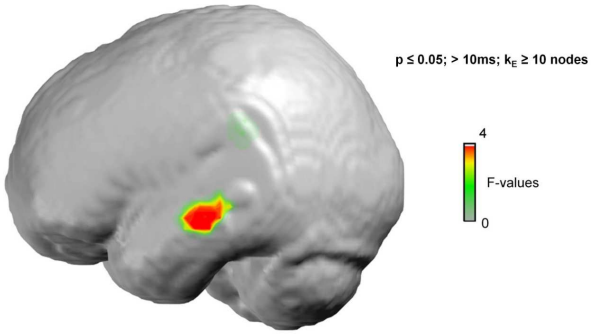


Figure S3: Results supplemental analysis 2, source estimations F-maps (45-55 ms). The results of the one-way ANOVA comparing the four cardiac events at the beginning of the sound revealed differential activity within the left middle temporal gyrus over the 45-55ms post-stimulus onset period (with the following criteria: duration of at least 10 consecutive TF (approximately 11 ms) and a cluster size of at least 10 contiguous nodes). The results of source estimations, p-values, are displayed on the MNI averaged brain.