



**HAL**  
open science

# Neural sensitivity to statistical regularities as a fundamental biological process that underlies auditory learning: The role of musical practice

Clément François, Daniele Schön

## ► To cite this version:

Clément François, Daniele Schön. Neural sensitivity to statistical regularities as a fundamental biological process that underlies auditory learning: The role of musical practice. *Hearing Research*, 2014, Music: A window into the hearing brain, 308, pp.122-128. 10.1016/j.heares.2013.08.018 . hal-02062416

**HAL Id: hal-02062416**

**<https://hal.science/hal-02062416v1>**

Submitted on 8 Mar 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

NEURAL SENSITIVITY TO STATISTICAL REGULARITIES AS A  
FUNDAMENTAL BIOLOGICAL PROCESS THAT UNDERLIES  
AUDITORY LEARNING: THE ROLE OF MUSICAL PRACTICE.

Clément François<sup>(1, 2)</sup> & Daniele Schön<sup>(3, 4)</sup>

(1) Cognition and Brain Plasticity Unit, Dept. of Basic Psychology (Campus de Bellvitge) & IDIBELL, University of Barcelona, Feixa Llarga s/n, 08907, L'Hospitalet (Barcelona), Spain

(2) Department of Basic Psychology, Faculty of Psychology, University of Barcelona, 08035, Barcelona, Spain

(3) Aix-Marseille Université, INS, Marseille, France

(4) INSERM, U1106, Marseille, France

Corresponding Author : Daniele Schön

[daniele.schon@univ-amu.fr](mailto:daniele.schon@univ-amu.fr)

INSERM U 1106 /Institut de neurosciences des systèmes

Centre MEG, Service de Neurophysiologie Clinique

CHU Timone - Rue Saint Pierre-13385 Marseille Cedex 5

Tel: +33 (0) 4 91 38 55 77 - Fax : +33(0) 4 91 78 99 14

## Abstract:

There is increasing evidence showing that humans and other nonhuman mammals are sensitive to the statistical structure of auditory input. Indeed, neural sensitivity to statistical regularities seems to be a fundamental biological property underlying auditory learning. In the specific case of speech, statistical regularities have been shown to play a crucial role in the acquisition of several linguistic features, from phonotactic to more complex rules such as morphosyntactic rules. Interestingly, a similar sensitivity has been shown with non-speech streams: sequences of sounds changing in frequency or timbre can be segmented on the sole basis of conditional probabilities between adjacent sounds. We recently ran a set of cross-sectional and longitudinal experiments showing that, on the one side, merging music and speech information in song facilitates stream segmentation and, on the other side, musical practice directly affects sensitivity to statistical regularities in speech both at the neural and behavioral levels. Based on recent findings showing the involvement of a fronto-temporal network in speech segmentation, we will defend the idea that enhanced auditory learning observed in musicians can originate via at least three distinct pathways: enhanced low-level auditory processing, enhanced phono-articulatory mapping via the left Inferior Frontal Gyrus and Pre-Motor cortex as well as increased functional connectivity within the whole audio-motor network. Finally we will discuss how these data predict a beneficial use of music in optimizing speech acquisition trajectories both in normal and impaired populations.

Predicting future events is crucial for survival in an ever-changing auditory world. The detection of auditory changes in the environment and the prediction of future auditory events at different time scales are particularly interesting for the study of music and speech auditory processing. Music and speech are two communicative activities that involve the combination of an elementary set of sounds ordered in time according to rules allowing the perception and the production of complex and unlimited utterances or musical phrases. Both activities require several perceptual and cognitive processes ranging from sound identification and categorization to memory storage and retrieval. Most importantly for our present aim, because language and music evolve in time, elementary elements become part of more complex structures following specific (emerging) rules and making specific (emerging) predictions. The study of how the human brain builds these expectations, one of the core mechanisms of language and music acquisition, has a specific interest to inform basic principles of both auditory neuroscience and psychophysiology. In the first part of this article, we will bring evidence showing that neural sensitivities to simple acoustic changes or more complex statistical rules found in auditory sequences are one of the fundamental biological processes that underlies auditory learning for both music and speech. In a second part, we will discuss the role of music, musical practice and training in revealing refined sensitivity to these statistical properties and will finish by proposing testable hypotheses accounting for such an enhanced sensitivity in musicians.

## 1. Cortical correlates of change detection in animal and human models

### 1.1. Change detection in animals

The cerebral correlates of auditory change detection have been widely studied via neurophysiological measures both in humans and animals. Neural sensitivity to low probability auditory events has been shown to take place within the primary auditory cortices of anesthetized cats by using invasive intra-cortical single-unit recordings (Ulanovsky et al., 2003). Indeed,

Ulanovsky and collaborators showed that neurons in this cortical area responded more to rarely occurring pure tones embedded in a repetitive sequence of standard ones. This finding suggested that Stimulus Specific Adaptation (SSA), i.e. the decrease in the firing rate of neurons to repetitive adapted standard tones, might be the correlate of change detection at the level of single neurons, and a possible simple mechanism for auditory memory. In a further study, these authors showed that SSA in primary auditory cortices could occur at different time scales ranging from a few hundred milliseconds to tens of seconds (Ulanovsky et al., 2004). More recently, Yaron and collaborators (2012) have revealed the presence of neurons in primary auditory cortices of anesthetized rats that are sensitive to more detailed information. The authors analyzed intra- and extracellular recordings from the auditory cortex of anesthetized rats in response to two types of oddball sequences wherein the position of rare tones could be either a random or periodic. They found smaller response to tones in periodic sequences than to the same tones in random order. Most interestingly, the activity of primary auditory cortex neurons was clearly shown to be modulated by the order of sound sequences at the time scale of minutes. The authors proposed that such a neuronal sensitivity could be the correlate of statistical regularities processing involved in extracting rather natural syntactic properties of language and music.

## 1.2. Change detection in humans

There is growing evidence suggesting that several aspects of speech and music processing in humans involve similar change detection processes, from the detection of acoustic changes to the detection of changes in the statistical structure of a sequence of sounds. Thanks to excellent time resolution, Electroencephalographic (EEG) recordings offer non-invasively access to cortical brain activity on the order of the millisecond in response to external stimuli such as pure tones, harmonic sounds or speech sounds. For instance, EEG has been used to bring the first evidence of sensitivity to rarely occurring auditory events in the auditory cortex. Using tone pips, neural sensitivity to low

probability auditory events was shown in human adults and then in newborns (Hari et al., 1984; Alho et al., 1991). The well-known Mismatch Negativity (MMN), a change-detection component of the auditory Event-Related brain Potentials (ERPs), is elicited by rarely occurring deviant events embedded within a repetitive sequence of standard stimuli (Näätänen et al., 2005, Grimm & Escera 2011). The MMN is considered a measure of pre-attentive, implicit auditory processing which may reflect the formation of an echoic memory trace within the auditory cortices (Näätänen et al., 2005). This ERP component is found for changes in frequency (Sams et al., 1985), intensity (Paavilainen et al., 1993), location (Paavilainen et al., 1989), duration (Schröger and Winkler, 1995) or timbre (Tervaniemi et al., 1997). Importantly, the MMN is also elicited for changes in different acoustic parameters of speech sounds consisting of single vowels or CV syllables (Deguchi et al., 2010, Näätänen 2001). Interestingly for the present review, other studies have demonstrated that a MMN can be elicited by changes in longer, much more complex and/or abstract patterns of musical and speech sounds (Boh et al., 2011; Herholz et al., 2009; Wang et al., 2012). All together, these findings support the idea that detecting changes as well as detecting patterns both require that the auditory system be sensitive to statistical regularities. Furthermore, these results show that auditory change detection processes that take place within the auditory cortices can reflect the processing of formant transition variations sub-serving phoneme and timbre change detection as well as the processing of more complex structures in longer linguistic or musical contexts. We suggest that the detection of a change, strongly linked to the predictive abilities of the auditory system, can be considered as a basic biological process involved both in music and speech sequence learning.

## 2. Sensitivity to statistical structure of auditory input.

### 2.1. Statistical learning of single sounds

In the context of language acquisition, it is now well accepted that sensitivity to speech sounds evolves during infancy (Kuhl 2004). Notably, it has been shown that infants' ability to discriminate

different speech sounds is sensitive to the probabilistic patterns of speech sounds present in the environment. During the first year of their life, while infants between the ages of 0 to 6 months do not differentiate foreign from native phonetic units, after the age of 6 months they show a drastic decrease in the sensitivity to foreign phonological units while becoming attuned to phonetic units of their mother tongue (Kuhl 2004). This property of phonetic learning has been elegantly highlighted in a behavioral experiment with 6 and 8 month-old infants (Maye et al., 2002). One group of infants was familiarized during 2 minutes with a [da] - [ta] continuum presenting a bimodal frequency distribution (tokens from the extremities of the continuum were the most frequent) and another group of infants was familiarized with a similar continuum but presenting a unimodal frequency distribution (tokens from the center of the continuum were the most frequent). Infants were subsequently tested in a phonetic discrimination test. While the group of infants exposed to the bimodal distribution discriminated tokens from the extreme points of the continuum, infants who were previously familiarized with a unimodal distribution did not. These results brought evidence that perceptual learning, as it has been showed for the learning of native phonemes seems to rely on the sensitivity to the statistical properties of the speech sounds present in the environment (Maye et al., 2002). This process is nowadays commonly referred to as statistical learning. Statistical learning has been shown to occur not only at the level of phonemes but also at the word level (Saffran et al., 1996a).

## 2.2. Statistical learning of sequences of sounds

Besides identifying the phonemes present in the input, another crucial step of language acquisition is the ability to extract the words from fluent speech unfolding in time. In regular adult directed speech there are no systematic acoustic cues such as silences or stresses at words boundaries. By contrast, infant directed speech presents exaggerated pitch contour changes and slower speech rate that may largely contribute in solving the speech segmentation problem (Fernald 1992; Thiessen et

al., 2005). Nonetheless, it has been shown that speech can be segmented in an implicit manner based on only one source of information: the statistical structure of the language (Saffran et al., 1996a). Indeed, “syllables that are part of the same word tend to follow one another predictably, whereas syllables that span word boundaries do not” (Saffran et al. 2001). The role of transitional probabilities in speech segmentation (i.e., the probability of syllable Y given syllable X:  $p(Y/X) = \text{frequency of } XY / \text{frequency of } X$ ) has been elegantly shown by Saffran and collaborators in both infants and adults (Saffran et al. 1996a, 1996b; Aslin et al. 1998). Subsequently, studies using EEG and Near Infrared Spectroscopy have also revealed this ability in neonates (Gervain et al. 2008; Teinonen et al. 2009). In these kinds of experiments, the paradigm generally consists of a familiarization (learning) phase followed by a behavioral test. During the exposure phase, participants are exposed to several minutes (depending on the complexity of the stream and on the population of interest) of a statistically structured, continuous flow of artificial syllables lacking any acoustic cues at words boundaries. The test phase for adult participants is typically a two-alternative forced choice (AFC) procedure during which participants have to choose, in each trial, between a familiar pseudo-word that was part of the language and a pseudo-word built with similar syllables, but that was not part of the language (see figure 1 for an illustration of the experimental procedure used in François & Schön 2011). Above chance performance is interpreted as participants’ ability to segment the speech stream on the basis of statistical properties of the input. This ability to extract units thanks to statistical properties of the input has also been shown for non-linguistic stimuli such as sounds with different pitches (Saffran et al. 1999) and timbres (Tillmann and McAdams 2004), as well as with environmental sounds (Sanders et al., 2009). These results suggest that statistical learning is a domain general learning process. Additionally this statistical sensitivity has been revealed in the cotton top tamarin and in rats as well (Hauser et al. 2000, 2002; Ramus et al. 2000; Toro and Trobalon 2005) suggesting that sensitivity to acoustic regularities is a fundamental auditory process that undergirds normal auditory processing, survival in a complex acoustic



environment and even, in many cases, successful expressive and receptive vocal communication. In the domain of music, statistical learning has been shown to be at play for the acquisition of new unfamiliar musical scales as well as for the learning of 12-tone serial musical sequences (Loui et al., 2010; Dienes & Longuet-Higgins 2004). Loui and collaborators familiarized adult participants with melodies originating from a finite-state musical grammar during 30 minutes. The results of this study showed evidence of learning as revealed by recognition, generalization and preference for repeated melodies from the grammar. The authors suggested that sensitivity to statistical regularities is a basic principle accounting for music acquisition and preference. To summarize, these studies bring converging evidence suggesting that sensitivity to statistical regularities is one of the fundamental perceptual-cognitive processes underlying auditory learning and that it may serve as a basic principle for several aspects of language and music learning.

## 2.2. Brain correlates of auditory statistical learning.

A growing number of studies have been conducted in human adults and children to unravel the cortical network involved during such statistical learning of speech and tone sequences. Using fMRI, it has been shown that compared to a random stream (thus containing no predictable patterns of syllables), listening to a statistically structured stream of syllables activated more of the so-called fronto-temporal dorsal pathway including the Superior Temporal Gyrus, the Middle Temporal Gyrus, the Pre-Motor Cortex and the Inferior Frontal Gyrus (Hickok & Poeppel 2007; Cunillera et al., 2009; Rodriguez-Fornells et al., 2009; McNealy et al., 2006; 2010). Using EEG recordings, other studies have shown larger learning related modulations on the N100 and N400 components for structured compared to random sequences of syllables (Cunillera et al., 2009, De Diego Balaguer et al., 2007). Moreover, in a study using sequences of tones, Abia and colleagues reported similar modulations of the N100 and N400 components during the exposure phase suggesting that similar processes might be at play for speech and tone segmentation (Abia et al., 2008).

Interestingly for the present review, a recent study has provided a precise description on the organization of the STG with respect to statistical sensitivity for speech and non-speech sounds (Tremblay et al., 2012). In this study, participants were presented with random and highly structured sequences of speech and bird songs while performing a visual task. The results revealed several important findings. While some sub-regions of the STG showed a relative preference to speech (i.e. being significantly activated for both speech and bird songs), others showed an absolute preference for speech (i.e. being significantly activated for speech only). Most interestingly, some sub-regions in the Posterior Supra Temporal plane including the medial Transverse Temporal Gyrus were sensitive to the statistical structure of the input independently of the sound category. The authors proposed that the supratemporal plane is functionally organized according to an anteroposterior gradient (figure 2). More anterior regions might be operating specifically for speech at the level of single phonemes while more posterior regions might be involved in processing statistically structured sequences of both speech and non speech sounds. This last study is particularly interesting because it supports the idea that the processing of different categories of structured sequences (speech and non speech sounds) may rely on shared neural processes in the posterior STG of healthy participants. Moreover, it supports Patel's OPERA hypothesis (Patel, 2011), which is presented in an expanded version in the same issue of *Hearing Research* (Patel, 2013).

### 2.3. Sensitivity to statistical structure as a domain general skill

The body of literature presented above suggests that sequencing abilities in music and speech may partly rely on general shared learning mechanisms and that these abilities may be sustained by shared and overlapping cortical resources (Patel 2003; Kraus & Chandrasekaran, 2010). However, studies investigating statistical learning processes have focused on one dimension or the other but not on both at the same time. This renders the comparison of results issued from different tasks, participants, and analyses difficult and not straightforward. In this respect, the use of song is

particularly well suited to the study of the relation between language and music, the advantage being that both linguistic and musical information are merged into one single acoustic signal containing two salient dimensions, thus allowing for a direct comparison within the same experiment. We recently ran a set of behavioral experiments using sung languages with different statistical structures during the exposure phase. In a first experiment (Schön et al., 2008), we compared segmentation performance of a spoken language versus segmentation performance of a sung language. Two groups of participants were familiarized with an artificial language that was either spoken or sung (each syllable being sung on a specific pitch). While 7 min of familiarization were not enough to segment a spoken language, they were sufficient to learn the sung language. These behavioral results suggest that the learning process is facilitated when the statistical structure of the musical and linguistic dimensions match. Because one may claim that song is simply more arousing than speech, we ran a further experiment using another sung language wherein linguistic and musical statistical structures mismatched. The participants' level of performance was exactly in between the spoken and the matching sung version obtained in the two previous experiments. These results confirmed a beneficial effect of both distributional and motivational properties of music in the very first steps of language acquisition (Thiessen et al., 2005; Schön et al., 2008).

### 3. Musical practice facilitates the extraction of statistical regularities

In two recently published articles we studied separately the effect of musical expertise in adults and the effect of active musical training in 8 to 10 year old children on speech segmentation abilities (François & Schön 2011, François et al., 2012, see also Shook et al., 2013 for additional behavioural evidence of enhanced statistical learning of an auditory Morse code language in musicians). The underlying hypothesis was that if some auditory processes involved in both linguistic and musical learning are shared, then, musical training, by enhancing the functioning of these common processes, should also affect language processing. We used both behavioral and

electrophysiological measures to assess speech segmentation skills in participants exposed to several minutes of a sung stream. While in the adult experiment we tested both the recall of linguistic and musical dimensions contained in the sung language, we tested only the recall of the linguistic dimension in children. In the first study with adults, we showed that musicians barely outperformed nonmusicians at the behavioral level: both groups performed above chance level in the linguistic test while staying at chance in the musical test. Interestingly, ERP data were more sensitive and showed a main effect of expertise on the N1 component with larger amplitude for musicians than for nonmusicians in both dimensions. Additionally, we found that the amplitude of a later N400-like fronto-central negative component was more sensitive to the transitional probabilities in musicians compared to nonmusicians, again in both dimensions. These findings suggest that musicians may have more “robust” representations of both musical and linguistic structures shaped during the exposure phase than nonmusicians. Interestingly, in the musical test we found a similar interaction of expertise and sensitivity to the statistical structure on two early ERP components. Musicians showed both a larger P2 and MMN to low transitional probability melodies than to high transitional probability melodies while nonmusicians did not show these effects. The Transitional Probability effect on the P2 and MMN components suggested that musicians learned the musical structure (both tonal and interval structures) better than nonmusicians. In the second study with children, we compared a group of children enrolled in a music-training program during two school years to a matched control group of children enrolled in an equally motivating painting training program. Results were impressive in showing that while no pre-training between-group differences were observed neither at the behavioral nor at the electrophysiological level, only musically trained children showed a large increase in their speech segmentation abilities as revealed by a linear increase in performance after one and after two years of training. Interestingly, after 2 years, this behavioral benefit was accompanied, as in adults, by a greater sensitivity of a fronto-central N400-like component to transitional probabilities in musically trained children than in

painting trained children. Our findings provided causal evidence that music training can enhance the sensitivity to the statistical regularities in speech, a fundamental biological process underlying language learning. Additionally, this result argues in favor of the intuitive and common thinking that music can play a very important role in children's language development notably by facilitating speech segmentation, a building block of language acquisition. In the following sections we will suggest three non-exclusive explanations that may account for the observed benefit of musicians over nonmusicians in speech segmentation.

#### 4 Why musical practice may facilitates speech segmentation

##### 4.1 Enhanced low-level auditory processing for speech and musical sounds in musicians.

The study of professional musicians is of great interest because it can reveal how the long term coupling of auditory and motor systems can fine-tune the brain. Studying professional musicians allows determining the functional and structural modifications fostered by musical practice (at least when revealed by longitudinal studies). Indeed, musicians' brain is considered as a good model of neuroplasticity both at the structural and functional levels (Münte et al., 2002). Compared to non-musicians, musicians show an enhanced cortical attentive and pre-attentive processing of linguistic (for example see Chobert et al., 2011 for Voice Onset Time, frequency and duration deviances) and musical features (see for example Vuust et al., 2011). This enhancement is reflected by larger amplitude and/or shorter latencies of many ERP components such as the N1, P2, MMN and P300 (Pantev et al., 1998; Koelsch et al., 1999; Shahin et al., 2003; Van Zuijen et al., 2005). For instance, adult musicians show larger MMN than nonmusicians to deviance in contour and interval structure inserted in repetitive 5 notes melodies (Fujioka et al., 2004). Even at the sub-cortical level, adult, child and preschooler musicians show more robust encoding of linguistic and musical features as reflected by earlier and larger brainstem responses compared to non-musicians (Kraus and Chandrasekaran, 2010; Musacchia et al., 2007; Strait et al., 2012, 2013; Parbery-Clark et al., 2011,

Strait & Kraus, 2013). Most of these studies used cross-sectional approaches and causation can only be inferred in the case of well-controlled longitudinal studies with pseudo-random assignment of the participants. For instance, Lappe and collaborators (2008) compared MMNs elicited by deviant 3 piano tones sequences embedded in standard sequences before and after 2 weeks of musical training in one group with sensory-motor training and in another with auditory training only. The results showed a greater enhancement of MMN amplitude in the sensory-motor trained group compared to the auditory training group. More recently, Chobert and collaborators revealed that 6 months and 12 months of active musical training in 8 to 10 year old children were sufficient to induce an enhanced MMN in response to changes in VOT and duration of CV syllables (Chobert et al., 2012). At the structural level, Hyde and collaborators (2009) showed that 15 months of musical training in 6-year-old children enhanced auditory discrimination performance in melodic and rhythmic tasks as well as in a finger motor sequencing task. Moreno and colleagues (2009) trained 8-year-old nonmusician children with music or painting for 6 months and revealed better pitch discrimination abilities for both melodies and sentence prosody in the musically trained group only. These last studies demonstrate that active musical training positively impacts multiple functional aspects of the brain. Most importantly, these studies show that music training facilitates both pre-attentive and attentive processing of acoustic deviance in linguistic and musical inputs, one the basic biological process involved both in music and speech sequence learning. Recently, Paraskevopoulos and colleagues (2011) used MEG recordings in a MMN paradigm to study the effect of musical practice on the neural and behavioral sensitivity to statistical structure of tone sequences. The stimuli consisted of repetitive three tone sequences as standard sequences, which were interleaved with deviant sequences in which the final tone was changed compared to standard ones. Musical experts and nonmusicians showed similar level of performance and similar MMN amplitude to deviance. Nonetheless, compared to nonmusicians musical experts showed a larger difference between standards and deviants on an early ERP component, the P50. These results

suggested that enhanced P50 to deviant tones might index a better encoding of transitional probabilities in musicians than in nonmusicians (Paraskevopoulos et al., 2011). Even more recently, Skoe and colleagues conducted an experiment demonstrating sensitivity to statistical in evoked responses understood to be generated by the auditory midbrain/brainstem (Skoe et al., 2013). They used brainstem recordings during the presentation of random and statistically structured sequences of doublets of tones while participants watched a silent photo-slideshow. Results showed that the best learners were participants having the most enhanced response to structured sequences. Interestingly, there is evidence showing that musicians have larger planum temporale than nonmusicians (Keenan et al, 2001; Luders et al, 2004) and that white matter integrity in the right hemisphere predicts the success in a pitch-related grammar-learning task (Loui et al., 2011). Additionally, trans-magnetic brain stimulation applied to Heschl's gyrus modulates the level of performance in a pitch discrimination task (Mathys et al., 2010). Taken together, these results show that an enhanced neural sensitivity of the statistical properties in auditory input within the temporal lobes and in the Inferior Colliculus of musicians might probably contribute to enhanced auditory learning skills.

#### 4.2. Enhanced IFG/PMC functioning

In addition to the involvement of the auditory cortex, prefrontal cortices (left IFG and PMC) are clearly activated during the exposure phase of an artificial language learning task (Cunillera et al., 2009; McNealy et al., 2006, 2010). IFG/PMC regions have been shown to be involved in language perception and production and in music processing as well (Vigneau et al., 2006; Koelsch et al., 2005; Brown et al., 2006). Grey matter density in the IFG has been found to correlate with the level of language proficiency (Stein et al., 2012) and in the number of years of phonetic training in expert phoneticians (Golestani et al., 2011). Interestingly, there is evidence showing that white matter integrity in the left-IFG predicts the level of performance in an artificial grammar learning

task (Flöel et al., 2009) and that musicians show increased gray matter density and volume in this very same brain region (Slumming et al., 2002). Thus a greater recruitment of associative frontal areas in musicians compared to nonmusicians could also contribute to enhanced neural sensitivity by acting as a top-down modulatory process onto low-level auditory cortices or as phono-articulatory matching process (Rodriguez-Fornells et al., 2009).

#### 4.3. Enhanced connectivity from pSTG to IFG/PMC

Finally, it is worth to consider feedforward and feedback neuronal projections within the speech segmentation network. The fibers of white matter forming the well-known Arcuate Fasciculus or Superior Longitudinal Fasciculus (AF/SLF) directly connect the posterior STG to the IFG/PMC. Lesions of the AF/SLF induce impairment not only of phonological and word repetition but also in verbal short-term memory tasks (Benson et al., 1973; Damasio and Damasio, 1980; Anderson et al., 1999). Moreover, studies have found that the AF/SLF is notably decisive in mapping speech sound sequences to articulatory sequences of gestures (Catani et al., 2005; Schmahmann et al., 2007) and in word learning (Lopez-Barroso et al., 2013). Interestingly, both adult musicians and in 8-year old children who followed 2 years of musical training show a more developed AF/SLF than nonmusicians (Oechslin et al., 2010; Wan & Schlaug, 2010, Halwani et al., 2011). Taken together these results suggest that, an increased connectivity between auditory regions within the temporal lobe and motor regions within the frontal lobe might induce enhanced segmentation skills in musicians compared to nonmusicians.

#### Conclusion

Further research will be needed to decipher the relative contribution of these three non exclusive possibilities suggested above and notably thanks to the use of innovative data processing techniques involving both structural and functional measures applied to specific well designed experiments



(François et al., 2012a). For instance, the analyses of ERPs during the exposure phase are of great interest because implicit electrophysiological measures (EEG/MEG) can reveal subtle differences between groups in language acquisition that might be underestimated by behavioral estimates (McLaughlin et al., 2004; Abla et al., 2008; Cunillera et al., 2009; De Diego Balaguer et al., 2007). Additionally, implicit electrophysiological measures can be very useful and informative when studying specific populations such as very young infants (Teinonen et al., 2009; Kudo et al., 2011) who are unable to deliver an overt behavioral response while being at the same time the most interesting population for disentangling the contribution of nature vs. nurture in the acquisition of language and music. The use of implicit measures could be also useful in the case of patients with communicative disorders such as patients with aphasia or children with autistic disorders or in the case of patients with motor disorders such as in Parkinson's disease. It is important to keep in mind that the present article focused exclusively on first-order statistical regularities. Besides these, natural languages make use of higher order statistics, for instance in hierarchical syntactical structures. Interestingly, recent work done by Brod and Opitz (2013) brings evidence of enhanced sensitivity to long distance dependencies in the learning of an artificial language for musicians compared to nonmusicians. This last finding suggests that musical practice might also benefit the learning of more complex syntactical rules in more natural language. The effect of singing in facilitating the speech segmentation as well as the effect of musical practice in enhancing brain functioning at the different levels of processing involved in auditory sequencing skills strongly promote the use of music as a tool for optimizing speech acquisition trajectories both in normal and impaired populations.

Here, we have presented evidence for enhanced sensitivity to statistical regularities in musicians than in nonmusicians. We have proposed that such a benefit might rely on a complex interaction between enhanced low-level auditory processing in both cortical and subcortical regions, enhanced functioning of associative frontal areas, and enhanced connectivity between those areas. One of the

possible mechanisms for such an enhanced sensitivity to statistical regularities has been proposed to involve corticofugal projections that shapes low-level auditory regions in a top-down manner (Skoe et al., 2013). However, future research is needed to clearly define the respective roles of bottom-up and top-down mechanisms for engendering neural sensitivity to acoustic regularities and how they are strengthened in musicians.

## References

- Abla, D., Katahira, K., Okanoya, K. 2008. On-line Assessment of Statistical Learning by Event related Potentials. *Journal of Cognitive Neuroscience* 20 (6), 952-964.
- Alho, K., Sainio, K., Sajaniemi, N., Reinikainen, K., Näätänen, R. 1991. Event-related brain potential of human newborns to pitch change of an acoustic stimulus. *Electroencephalography and clinical neurophysiology* 77 (2), 151-155.
- Anderson, J., Gilmore, R., Roper, S., Crosson, B., Bauer, R., Nadeau, S., Beversdof, D., Cibula, J., Rogish, M., Kortencamp, S., Hughes, J.D., Gonzalez Rothi, L.J., Heilman, K.M. 1999. Conduction aphasia and the arcuate fasciculus: a reexamination of the Wernicke-Geschwind model. *Brain and Language* 70, 1-12.
- Aslin, R.N., Saffran, J.R., Newport, E.L. 1998. Computation of conditional probability statistics by 8-month-old infants. *Psychological Science* 9, 321-324.
- Benson, D., Sheremata, W., Bouchard, R., Segarra, S., Prince, D., and Geschwind, N. 1973. Conduction aphasia: a clinicopathological study. *Archives of Neurology* 28, 339 -346.
- Boh, B., Herholz, S.C., Lappe, C., Pantev, C. 2010. Processing of complex auditory patterns in musicians and nonmusicians. *PloS One* 6 (7): e21458.
- Brod, B., Opitz, G. 2013. Does it really matter? Separating the effects of musical training on syntax acquisition. *Frontiers in Psychology* 3:543. doi: 10.3389/fpsyg.2012.00543.
- Catani, M., Jones, D.K., Ffytche, D.H. 2005. Perisylvian language networks of the human brain. *Annals of Neurology* 57 (1):8-16.
- Chobert, J., Marie, C., François, C., Schön, D., Besson, M. 2011. Enhanced passive and active processing of syllables in musician children. *Journal of Cognitive Neuroscience* 23(12), 3874-3887.
- Chobert, J., François, C., Velay, J.L., Besson, M. 2012. Twelve Months of Active Musical Training in 8- to 10-Year-Old Children Enhances the Preattentive Processing of Syllabic Duration and Voice Onset Time. *Cerebral Cortex* doi: 10.1093/cercor/bhs377

Cunillera, T., Càmarà, E., Toro, J. M., Marco-Pallares, J., Sebastián-Galles, N., Ortiz, H., Pujol, J., Rodríguez-Fornells, A. 2009. Time course and functional neuroanatomy of speech segmentation in adults. *Neuroimage* 48 (3), 541-553.

Damasio, H., and Damasio, A. 1980. The anatomical basis of conduction aphasia. *Brain* 103, 337-350.

De Diego Balaguer, R., Toro, J.M., Rodriguez-Fornells, A., Bachoud-Lévi, A.C. 2007. Different Neurophysiological Mechanisms Underlying Word and Rule Extraction from Speech. *PLoS One* 2 (11), e1175.

Deguchi, C., Chobert, J., Brunellière, A., Nguyen, N., Colombo, L., Besson, M. 2010. Pre-attentive and attentive processing of French vowels. *Brain Research* 1366, 149-161.

Dienes, Z., Longuet-Higgins, C. 2004. Can musical transformations be implicitly learned? *Cognitive Science* 28, 531-558.

Fernald, A. 1992. Human maternal vocalizations to infants as biologically relevant signals,” in *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, eds J. Barkow, L. Cosmides, and J. Tooby (Oxford: Oxford University Press), 391–428.

Flöel, A., de Vries, M.H., Scholz, J., Breitenstein, C., Johansen-Berg, H. 2009. White matter integrity in the vicinity of Broca's area predicts grammar learning success. *Neuroimage* 47 (4), 1974-1981.

François, C., Schön, D. 2011. Musical expertise boosts implicit learning of both musical and linguistic structures. *Cerebral Cortex* 21 (10), 2357-2365.

François, C., Chobert, J., Besson, M., Schön, D. 2012. Music Training for the Development of Speech Segmentation. *Cerebral Cortex* doi: 10.1093/cercor/bhs180.

François, C., Tillmann, B., Schön, D. 2012a. Cognitive and methodological considerations on the effects of musical expertise on speech segmentation. *Annals of the New York Academy of Science* 1252, 108-115.

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, 1010-1021.

Gervain, J., Macagno, F., Cogoi, S., Peña, M., Mehler, J. 2008. The neonate brain detects speech structure. *Proceedings of the National Academy of Sciences USA* 105, 14222-14227.

Golestani, N., Price, C.J., Scott, S.K. 2011. Born with an ear for dialects? Structural plasticity in the expert phonetician brain. *Journal of Neuroscience* 31 (11), 4213-4220.

Grimm, S., Escera, C. 2011. Auditory deviance detection revisited: evidence for a hierarchical novelty system. *International Journal of Psychophysiology* 85 (1), 88-92.

Halwani, G.F., Loui, P., Rüber, T., Schlaug, G. 2011. Effects of practice and experience on the arcuate fasciculus: comparing singers, instrumentalists, and non-musicians. *Frontiers in*

Hari, R., Hämäläinen, M., Ilmoniemi, R., Kaukoranta, E., Reinikainen, K., Salminen, J., Alho, K., Näätänen, R., Sams, M. 1984. Responses of the primary auditory cortex to pitch changes in a sequence of tone pips: neuromagnetic recordings in man. *Neuroscience letters* 50 (1-3),127-32.

Hauser, M.D., Newport, E.L., Aslin, R.N. 2000. Segmentation of the speech stream in a non-human primate: statistical learning in cotton-top tamarins. *Cognition* 78 (3), B53-B64.\$

Herholz, S.C., Lappe, C., Pantev, C. 2009. Looking for a pattern: An MEG study on the abstract mismatch negativity in musicians and nonmusicians. *BMC Neuroscience* 10: 42.

Hickok, G., Poeppel, D. 2007. The cortical organization of speech processing. *Nature Reviews. Neuroscience* 8, 393-402.

Hyde, K.L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A.C., Schlaug, G. 2009. Musical training shapes structural brain development. *Journal of Neuroscience* 29 (10), 3019-3025.

Keenan, J.P., Thangaraj, V., Halpern, A.R., Schlaug, G. 2001. Absolute pitch and planum temporale. *Neuroimage* 14 (6), 1402-1408.

Koelsch, S., Schröger, E., Tervaniemi, M. 1999. Superior pre-attentive auditory processing in musicians. *Neuroreport* 10 (6), 1309-1313.

Koelsch, S., Fritz, T., Schulze, K., Alsop, D., Schlaug, G. 2005. Adults and children processing music: an fMRI study. *Neuroimage* 25(4):1068-1076.

Krampe, R.T., Ericsson, K.A. 1996. Maintaining excellence: deliberate practice and elite performance in young and older pianists. *Journal of Experimental Psychology* 125, 331-359.

Kraus, N., Chandrasekaran, B. 2010. Music training for the development of auditory skills. *Nature Reviews. Neuroscience*. 11, 599-605.

Kudo, N., Nonaka, Y., Mizuno, N., Mizuno, K., Okanoya, K. 2011. On-line statistical segmentation of a non-speech auditory stream in neonates as demonstrated by event-related brain potentials. *Developmental Science* 14 (5), 1100-1106.

Kuhl, P.K. 2004. Early language acquisition: cracking the speech code. *Nature Reviews. Neuroscience* 5 (11), 831-843.

Lappe, C., Herholz, S.C., Trainor, L.J., Pantev, C. 2008. Cortical plasticity induced by short-term unimodal and multimodal musical training. *Journal of Neuroscience* 28 (39), 9632-9639.

Lopez-Barroso, D., Catani, M., Ripolles, P., Dell'acqua, F., Rodriguez-Fornells, A., de Diego-Balaguer, R. 2013. Word learning is mediated by the left arcuate fasciculus. *Proceedings of the National Academy of Sciences USA*. Early Edition. [www.pnas.org/cgi/doi/10.1073/pnas.1301696110](http://www.pnas.org/cgi/doi/10.1073/pnas.1301696110)

Loui, P., Wessel, D.L., Hudson Kam, C.L. 2010. Humans Rapidly Learn Grammatical Structure in a New Musical Scale. *Music Perception* 27 (5), 377-388.

Loui, P., Li, H., Schlaug, G. 2011. White Matter Integrity in Right Hemisphere Predicts Pitch-Related Grammar Learning. *Neuroimage* 55 (2), 500-507.

Lovio, R., Pakarinen, S., Huottilainen, M., Alku, P., Silvennoinen, S., Näätänen, R., Kujala, T. 2009. Auditory discrimination profiles of speech sound changes in 6-year-old children as determined with the multi-feature MMN paradigm. *Clinical Neurophysiology* 120, 916–921.

Luders, E., Gaser, C., Jancke, L., Schlaug, G. 2004. A voxel-based approach to gray matter asymmetries. *Neuroimage* 22 (2), 656-664.

Mathys, C., Loui, P., Zheng, X., Schlaug, G. 2011. Non-invasive brain stimulation applied to Heschl's gyrus modulates pitch discrimination. *Frontiers in Psychology* 1, 193.

May, P.J., Tiitinen, H. 2010. Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology* 47 (1), 66-122.

Maye, J., Werker, J.F., Gerken, L. 2002. Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82 (3), B101-B111.

McLaughlin, J., Osterhout, L., Kim, A. 2004. Neural correlates of second -language word learning: minimal instruction produces rapid change. *Nature Neuroscience*, 7 (7), 703.

McNealy, K., Mazziota, J.C., Dapretto, M. 2006. Cracking the language code: neural mechanisms underlying speech parsing. *Journal of Neuroscience* 26, 7629-7639.

McNealy, K., Mazziota, J.C., Dapretto, M. 2010. The neural basis of speech parsing in children and adults. *Developmental Science* 13 (2), 385-406.

Moreno, S., Marques, C., Santos, A., Santos, M., Castro, S.L., Besson, M. 2009. Musical training influences linguistic abilities in 8-year-old children: more evidence for brain plasticity, *Cerebral Cortex* 19, 712-713.

Münte, T.F., Altenmüller, E., Jäncke, L. 2002. The musician's brain as a model of neuroplasticity. *Nature Reviews. Neuroscience* 3(6), 473-478.

Musacchia, G., Sams, M., Skoe, E., Kraus, N. 2007. Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proceedings of the National Academy of Sciences USA* 104 (40), 15894-15898.

Näätänen, R., Gaillard, A.W.K., Mäntysalo, S. 1978. Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica* 42 (4), 313-329.

Näätänen, R., Jacobsen, T., Winkler, I. 2005. Memory-based or afferent processes in mismatch negativity (MMN): a review of the evidence. *Psychophysiology* 42 (1), 25-32.

Oechslin, M.S., Imfeld, A., Loenneker, T., Meyer, M., Jäncke, L. 2010. The plasticity of the superior longitudinal fasciculus as a function of musical expertise: a diffusion tensor imaging study. *Frontiers in Human Neuroscience* 3, 76.

Paavilainen, P., Karlsson, M.L., Reinikainen, K., Näätänen, R. 1989. Mismatch negativity to change in spatial location of an auditory stimulus. *Electroencephalography and clinical neurophysiology* 73 (2), 129-141.

Paavilainen, P., Tiitinen, H., Alho, K., Näätänen, R. 1993. Mismatch negativity to slight pitch changes outside strong attentional focus. *Biological psychology* 37 (1), 23-41.

Pakarinen, S., Lovio, R., Huotilainen, M., Alku, P., Näätänen, R., Kujala, T. 2009. Fast multi-feature paradigm for recording several mismatch negativities (MMNs) to phonetic and acoustic changes in speech sounds. *Biological psychology* 82 (3), 219-226.

Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L.E., Hoke, M. 1998. Increased auditory cortical representation in musicians. *Nature* 392 (6678), 811-814.

Paraskevopoulos, E., Kuchenbuch, A., Herholz, S.C., Pantev, C. 2011. Statistical learning effects in musicians and non-musicians: An MEG study. *Neuropsychologia* doi:10.1016/j.neuropsychologia.2011.12.007

Parbery-Clark, A., Strait, D.L., Kraus, N. 2011. Context-dependent encoding in the auditory brainstem subserves enhanced speech-in-noise perception in musicians. *Neuropsychologia* 49 (12), 3338-3345.

Patel, A.D. 2003. Language, music, syntax and the brain. *Nature Neuroscience* 6 (7), 674-681.

Patel, A.D. 2011. Why would musical training benefit the neural encoding of speech ? The OPERA hypothesis. *Frontiers in Psychology*. 2, 142. doi: 10.3389/fpsyg.2011.00142. eCollection

Patel, A.D. 2013. Musical training, neural plasticity, and speech processing: what are the principles. *Hearing Research*.

Ramus, F., Hauser, M.D., Miller, C., Morris, D., Mehler, J. 2000. Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science* 288 (5464), 349-351.

Rodríguez-Fornells, A., Cunillera, T., Mestres-Missé, A., De Diego-Balaguer, R. 2009. Neurophysiological mechanisms involved in language learning in adults. *Philosophical Transaction of the Royal Society B: Biological Science* 364, 3711-3735.

Saffran, J.R., Newport, E.L., Aslin, R.N. 1996a. Word segmentation: The role of distributional cues. *Journal of Memory and Language* 35, 606-621.

Saffran, J.R., Aslin, R.N., Newport, E.L. 1996b. Statistical learning by 8-month-old infants. *Science* 274, 1926-1928.

Saffran, J.R., Johnson, E., Aslin, R.N., Newport, E.L. 1999. Statistical learning of tone sequences by human infants and adults. *Cognition* 70 (1), 27-52.

Saffran, J.R., Senghas, A., Trueswell, J.C. 2001. The acquisition of language by children. *Proceedings of the New York Academy of Sciences USA* 98, 12874-12875.

- Sams, M., Paavilainen, P., Alho, K., Näätänen, R. 1985. Auditory frequency discrimination and event-related potentials. *Electroencephalography and clinical neurophysiology* 62 (6), 437-448.
- Sanders, L.D., Ameral, V., Sayles, K. 2009. Event-related potentials index segmentation of nonsense sounds. *Neuropsychologia* 47, 1183-1186.
- Schmahmann, J.D., Pandya, D.N., Wang, R., Dai, G., D'Arceuil, H.E., de Crespigny, A.J., Wedeen, V.J. 2007. Association fibre pathways of the brain: parallel observations from diffusion spectrum imaging and autoradiography. *Brain* 130 (3), 630-653.
- Schön, D., Boyer, M., Moreno, S., Besson, M., Peretz, I., Kolinsky, R. 2008. Song as an aid for language acquisition. *Cognition* 106 (2), 975-983.
- Schröger, E., Winkler, I. 1995. Presentation rate and magnitude of stimulus deviance effects on human pre-attentive change detection. *Neuroscience letters* 193 (3), 185-188.
- Shook, A., Marian, V., Bartolotti, J., Schroeder, S.R. 2013. Musical Experience Influences Statistical Learning of a Novel Language. *The American Journal of Psychology* 126 (1), 95-104.
- Skoe, E., Krizman, J., Spitzer, E., Kraus, N. in press. The auditory brainstem is a barometer of rapid learning. *Neuroscience* doi:10.1016/j.neuroscience.2013.03.009.
- Sluming, V., Barrick, T., Howard, M., Cezayirli, E., Mayes, A., Roberts, N. 2002. Voxel-based morphometry reveals increased gray matter density in Broca's area in male symphony orchestra musicians. *Neuroimage* 17 (3), 1613-1622.
- Shahin, A., Bosnyak, D.J., Trainor, L.J., Roberts, L.E. 2003. Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. *Journal of Neuroscience* 23 (13), 5545-5552.
- Stein, M., Federspiel, A., Koenig, T., Wirth, M., Strik, W., Wiest, R., Brandeis, D., Dierks, T. 2012. Structural plasticity in the language system related to increased second language proficiency. *Cortex* 48 (4), 458-465.
- Strait, D.L., Parbery-Clark, A., Hittner, E., Kraus, N. 2012. Musical training during early childhood enhances the neural encoding of speech in noise. *Brain and Language* 123 (3), 191-201.
- Strait, D.L., Parbery-Clark, A., O'Connell, S., Kraus, N. 2013. Biological impact of preschool music classes on processing speech in noise. *Developmental Cognitive Neuroscience*. 6C, 51-60.
- Strait, D.L., Kraus, N. 2013. Interactions between musical training, speech processing and cognitive function from preschool to adulthood: musicians as a model of auditory learning. *Hearing Research*.
- Teinonen, T., Fellman, V., Näätänen, R., Alku, P., Huotilainen, M. 2009. Statistical language learning in neonates revealed by event-related brain potentials. *BMC Neuroscience* 13, 10-21.
- Tervaniemi, M., Winkler, I., Näätänen, R. 1997. Pre-attentive categorization of sounds by timbre as revealed by event-related potentials. *Neuroreport* 8 (11), 2571-2574.

Thiessen, E.D., Hill, E., Saffran, J.R. 2005. Infant-directed speech facilitates word segmentation. *Infancy* 7, 49-67.

Tillmann, B., McAdams, S. 2004. Implicit learning of musical timbre sequences: statistical regularities confronted with acoustical (dis)-similarities. *Journal of Experimental Psychology: Learning Memory and Cognition* 30, 1131-1142.

Toro, J.M., Trobalón, J.B. 2005. Statistical computations over a speech stream in a rodent. *Perception and Psychophysics* 67 (5), 867-875.

Tremblay, P., Baroni, M., Hasson, U. 2013. Processing of speech and non-speech sounds in the supratemporal plane: Auditory input preference does not predict sensitivity to statistical structure. *Neuroimage* 66, 318-332.

Ulanovsky, N., Las, L., Nelken, I. 2003. Processing of low-probability sounds by cortical neurons. *Nature Neuroscience* 6, 391-398.

Ulanovsky, N., Las, L., Farkas, D., Nelken, I. 2004. Multiple time scales of adaptation in auditory cortex neurons. *Journal of Neuroscience* 24, 10440-10453.

Van Zuijen, T., Sussman, E., Winkler, I., Näätänen, R., Tervaniemi, M. 2005. Auditory organization of sound sequences by a temporal or numerical regularity – a mismatch negativity study sounds comparing musicians and non-musicians. *Cognitive Brain Research* 23, 270-276.

Vigneau, M., Beaucousin, V., Hervé, P.Y., Duffau, H., Crivello, F., Houdé, O., Mazoyer, B., Tzourio-Mazoyer, N. 2006. *Neuroimage* 30 (4),1414-1432.

Vuust, P., Brattico, E., Glerean, E., Seppänen, M., Pakarinen, S., Tervaniemi, M., Näätänen, R. 2011. New fast mismatch negativity paradigm for determining the neural prerequisites for musical ability. *Cortex* 47(9), 1091-1098.

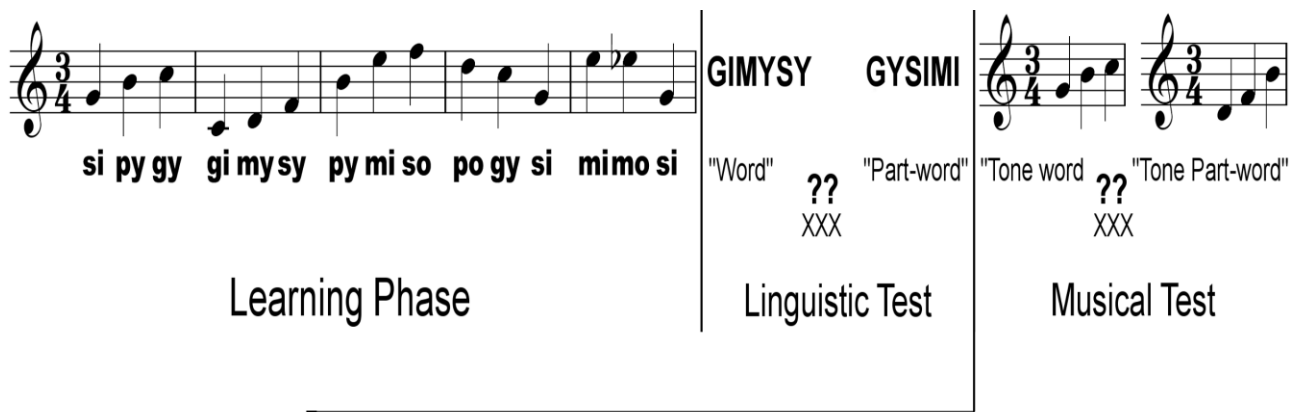
Wan, C.Y., Schlaug, G. 2010. Music making as a tool for promoting brain plasticity across the life span. *The Neuroscientist* 16 (5), 566-577.

Wang, X.D., Gu, F., He, K., Chen, L.H., Chen, L. 2012. Preattentive extraction of abstract auditory rules in speech sound stream: a mismatch negativity study using lexical tones. *PloS One* 7 (1), e30027.

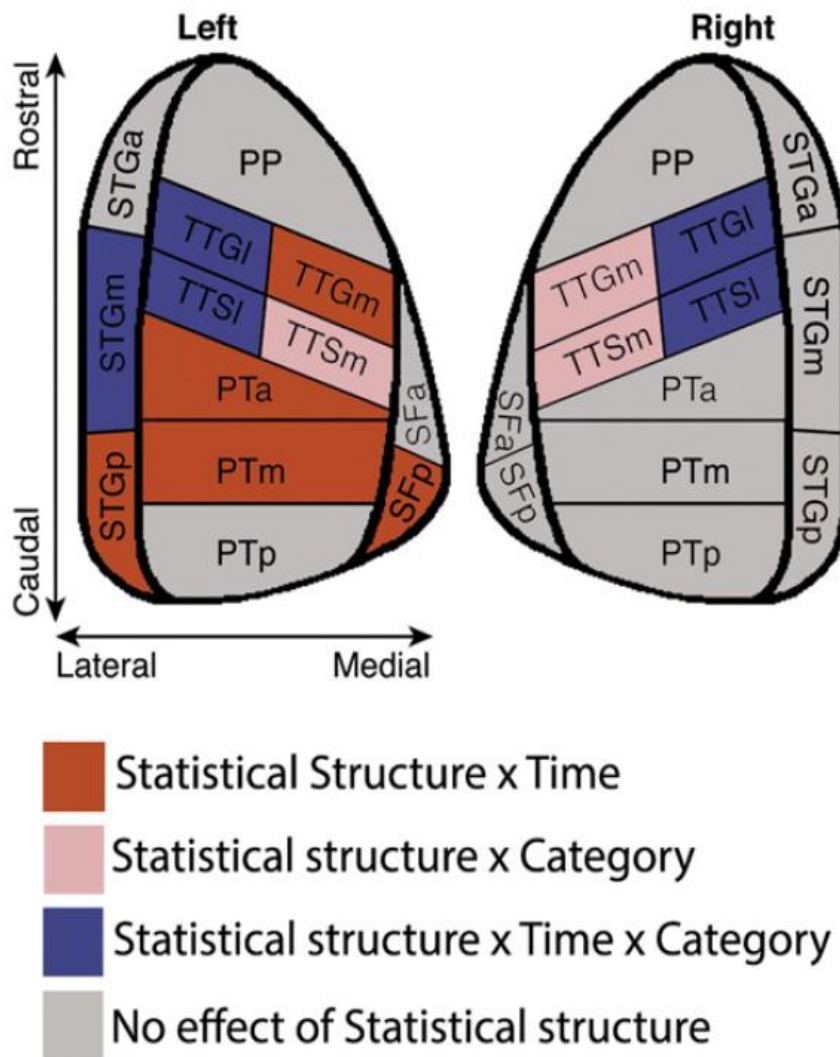
Yaron, A., Hershenhoren, I., Nelken, I. 2012. Sensitivity to complex statistical regularities in rat auditory cortex. *Neuron* 76 (3), 603-615.

## **Figure Legend:**





**Figure 1:** Illustration of the experimental design used in François & Schön 2011. Participants were presented with a statistically structured stream of sung syllables (exposure phase) and then tested with a 2AFC for recognition of both linguistic and musical structures. Stimuli were presented through loudspeakers.



**Figure 2:** From Tremblay et al., 2013, the figure illustrates the patterns of activity to sound category, statistical structure, exposure time and their interactions within the bilateral supratemporal areas of healthy adult participants. The anteroposterior gradient is visible from the color-coded patterns of responses with sensitivity to the statistical structure independently of sound category in more posterior areas and absence of sensitivity in more anterior areas. Regions' legend: PP = planumpolare; TTG = transverse temporal gyrus (m = medial, l = lateral); TTS = transverse temporal sulcus (m = medial, l = lateral); PT = planum temporale (a = anterior, m = middle, p=posterior); SF=caudal sylvian fissure (a=anterior, p=posterior); STG=superior temporal gyrus (a= anterior, m = middle, p = posterior).