ANNA DORA MANCA, FRANCESCO DI RUSSO, MIRKO GRIMALDI

Orderly organization of vowels in the auditory brain: the neuronal correlates of the Italian vowels

The main novelty of the study consists in using the EEG technique to elucidate the functioning of the auditory cortex in speech sound information processing. In searching for the cortical responses of two Italian vowels [a] and [i] we demonstrate that (i) as for the magnetic fields, also the modulations of the early electric N1 component is useful in exploring the mechanisms of phonetic feature extraction. The N1 amplitude and the N1 cortical generators seem to be affected by the vowel features as mapped in F2/F1 ratio so that the vowel [a] that shows a reduced inter-formant distance evokes weaker and anterior activity than [i]. The study thus, confirms that the cortical activity reflects the main spectro-acoustic dissimilarities supporting tonotopy as one of the most prominent organizing rules of the auditory cortices.

1. *Introduction*

The structures of the auditory pathway are specialized in categorizing different classes of sounds. Recent anatomical models of speech have demonstrated the existence of circuits functionally organized for the sound processing. (Hickok, Poeppel, 2007; Hickok et al., 2011; Scott, Johnsrude, 2003; Rauschecker, Tian, 2000). Cumulative findings supported the critical role of the anterior (Scott et al., 2000; Scott, Johnsrude, 2003; Obleser et al., 2006; 2010; Zatorre, Belin, 2001) and the posterior parts (Okada et al., 2010; Hickok, Poeppel, 2004; Wise et al., 2001) of the superior temporal gyrus (STG) in the phonetic and phonological encoding (Rauschecker, Tian, 2000). In general, the common assumption is that some areas along the auditory cortices show an extraordinary capability of extracting and classifying species-specific vocalizations, including the speech sounds (Alain et al., 2001; Scott et al., 2000; Obleser et al., 2010). It has been suggested that speech processing, consists of a cascade of computations that allows decoding streams of continuously varying sounds onto discrete representations. How the mapping procedures occur, it is still poorly understood. At a physical level, speech sounds are the output of the acoustic effects of the movements of specific articulator organs (Kent, 2002; Stevens, 2002). For example, vowels are described by the first three peaks of the spectral envelope, the so-called formant frequencies, i.e. F1, F2 and F3 (Delattre et al., 1952; Albano Leoni, Maturi, 2003) which are correlated with specific articulator configurations. F1 inversely correlates with the tongue height, F2 with the tongue place of articulation, and F2 with F3 are usually used for describing the lip configurations during the vowel production (Stevens, 2002). Intriguingly, the first two formants are crucial-

ly important for vowel categorization (Kent, 2002; Diesch, Luce, 2000). Yet, it seems that spectral relation between formants play a crucial role in determining the principles of functional organization of auditory cortex (Ohl, Scheich, 1997; Diesch, Luce, 1997; 2000).

In the last decade, a series of magnetoencephalographic (MEG) studies has provided new insights on the earliest steps of information processing showing tonotopy as one of the most prominent organizing rules of the human auditory cortex (Pantev et al., 1995; Eulitz et al., 1995). Specifically, it has been shown that distinctive pools of neurons, spatially located at different sites along the auditory pathway, revealed a different sensibility to the properties of the spectral structure of the linguistic stimuli (Diesch, Luce, 1997; 2000). As for vowels, for instance, the phonetic and phonological attributes seems to affect the activations of different sites of the auditory regions, even at the earliest steps of auditory information encoding (Obleser et al., 2003). The behavior of the auditory-evoked responses represent the main tool of the analysis. The N1 component and the magnetic counterpart N1m, is a broad negativity peak occurring between 80 and 150 ms after the stimulus onset, it is generally distributed over the fronto-central scalp area (Wolpaw, Perry, 1975; Woods, 1995). Its properties in terms of amplitude, latency and location of the cortical generators seem to be the cues crucial to describe the mechanisms of the sound speech encoding (Näätänen, Picton, 1987; Eulitz et al., 1995). By means of N1m, MEG data suggested a phonemotopic organization of human auditory cortices (Obleser et al., 2003; 2004; Mäkelä et al., 2003; Shestakova et al., 2004). For example, Obleser and colleagues (2003) revealed weak N1m amplitude values for /a/ because of its spectral envelope in comparing the German vowels [i] with [a]. As for the N1 latency, different studies have stressed that it is affected by acoustic correlates of speech segments peaking early for the first formant frequencies around 1 kHz (Roberts et al., 2000). For example, vowel [a] peaked earlier than vowels [i] or [u] in different studies (Roberts et al., 2000; Diesch, Luce, 1997; Poeppel et al., 1997). Again, the interaction between F1and F2 formants of the vowels, appear to guide vowel cortical mapping (Ohl, Scheich, 1997). Finally, this fact has also demonstrated that, the spatially localization of the N1 generators reflected the acoustic and the related phonological representation of the perceived sounds (Diesch, Luce, 1997; Obleser et al., 2003; 2004; Shestakova et al., 2004; Mäkelä et al., 2003).

In the present work, the cortical responses to two natural Italian vowels [a] and [i] were analyzed by exploring the auditory N1 response of the event-related response (ERP) signals. The aim of the study was twofold: on the one hand, we want to test the tonotopic organization of the auditory cortex by the N1 ERP component and on the other hand, we want to check whether, at different N1 modulations, separable cortical vowel representations can be associated or not. To do this, we analysed amplitude, latency and the cortical generators of the N1 component. It is worthy to note that, to the best of our knowledge, no study has used the electroencephalography (EEG) technique for exploring the vowel cortical mapping procedures, the magnetoencephalography (MEG) technique instead, has

usually been preferred for its high spatial accuracy in the measurements. However, some recent studies have shown that MEG offers no significant advantage over EEG in localizing a focal source (Virtanen et al., 1998; Cohen et al., 1990); moreover, both techniques seem to reveal similar differences in terms of accuracy in the spatial resolution of the cortical sources (Eulitz et al., 1997). Finally, advance in the high-density electrode montage and the EEG source analysis software provide much accurate brain localization measurements (Gevins et al., 1994; Cohen, Halghren, 2003).

Therefore, we expect that the acoustic phonetic features of the vowels shape their neuronal representation and that the electric N1 modulations show the tonotopic structure of the auditory cortex.

2. *Materials and methods*

2.1 Participants

Twelve students (7 males) with a mean age of 25 ± 3 years took part in the experiment. They provided a written informed consent. They were consistently right-handed according to Handedness Edinburgh Questionnaire (Oldfield, 1971). None of them had any known neurological disorder or other significant health problem. The experimental procedure was approved by the local ethics committee.

2.2 Stimuli

The stimuli consisted of the two SI vowels [a] and [i]. Three native Italian male speakers (mean age 33 ± 7 years) realized ten repetitions of each vowel in isolation, at a normal rate. The speech signal was recorded in a soundproof room with CSL 4500 and a Shure SM58-LCE microphone with a sampling rate of 44.1 kHz and an amplitude resolution of 16 bits. The stimuli were normalized for intensity (70 dB/SPL) and for duration (300ms) by using the speech analysis software Praat 5.2 (Boersma, Weenink, 2011). For every speaker (S1, S2, and S3) we selected the tokens with similar contours pitch within every vowel category. Finally, the first (F1), the second (F2) and the third (F3) formant values were measured in the vowel steady tract centered at the midpoint and the acoustic distances (F2-F1) were extracted (Table 1).

Table 1 - *Spectral compositions of the vowel stimuli (measured in Hertz) used in the experimental protocol*

Vowel	Speaker	F0	F1	F2	F2-F1
	S1	136	856	1344	488
/a/	S2	122	838	1432	594
	S3	133	816	1373	557
	S1	150	321	2189	1868
/i/	S2	129	313	2470	2157
	S3	134	281	2308	2027

2.3 Experimental Procedure

A schematic representation of the experimental procedure is shown in Figure 1. In a dark and shielded room, subjects sat in front of a computer monitor and were instructed to passively listening to the Italian vowels [a] and [i] randomly presented (Perception task (P)) and then to perform a specific articulatory task when a white screen appeared. In three different sessions, subjects had to (i) produce aloud the perceived vowel i.e., Acoustic Production task (AP), (ii) produce without emitting sound the vowel perceived, i.e. Silent Production task (SP) and (iii) mentally produce the perceived vowel i.e., Articulation Imagery task (AI).

Each trial began with a random exposure to a black screen (200-500ms) followed by an audible stimulus [a] and [i] which was randomly presented for 300 ms. After a randomized time interval (200-500ms) a white fixation cross (500ms) appeared at the centre of the black screen in order to suggest subjects to concentrate and prepare for the task. Another randomized time interval (200-500ms) preceded a white screen (2sec), which triggered the onset of the production and imagination tasks. When the black screen appeared again, the next trial started. In each experimental session, vowels were presented 160 times, 80 per vowel category. Since the perception condition systematically preceded the three articulatory tasks, perceptual data were finally computed on 240 trials per vowel (80 vowels per three Tasks). The auditory interstimulus interval (ISI) ranged from 3100 to 4300ms.

Figure 1 - *Timing of the experimental conditions. The auditory N1 responses (P) preceding the articulatory tasks (AP, SP, AI) were off-line summed and computed on 240 trials per vowel category*

In the present report, only the results concerning the Perception task are considered; the whole study is reported in Manca and Grimaldi 2013.

2.4 EEG Recordings

Continuous EEG was recorded by using a 64-ch actiCAP (BrainProducts GmbH, Germany; 10-20 System) and BrainVision Recorder 1.20 (BrainProducts GmbH, Germany) with a sampling rate of 250 Hz and a band pass filter of 0.1-70 Hz. Eye movements were monitored by several electrodes: for the vertical electrooculogram (VEOG) Fp1- Fp2 and an additional electrode (O2) was attached below the right eye; for the horizontal electrooculogram (HEOG), the electrodes FT9 and FT10 were used. The online reference was at FCz and impedance was kept under 5 Ω by electrogel conductant.

3. *Data analysis*

3.1 EEG Processing

Off-line signal processing was carried out with the software package BrainVision Analyzer 2.0.1 (BrainProducts GmbH, Germany). 1200 ms ERP epochs (including 200 ms of baseline) were extracted according to the onset of the vowel stimulus and digitally filtered by a 1–30 Hz bandpass filter (48db), a notch filter (50 Hz) and re-referenced to the right and left mastoids (TP9-T910). Ocular artifacts were removed by applying the Independent Component Analysis (ICA) algorithm and, additionally, rejection criteria were set (maximum allowed absolute difference =120 µV in 1000ms). Artifact-free segments were separately averaged for each vowel and a baseline correction of 200 ms was applied. Finally, the grand averages were computed across the subjects and for each vowel category. N1 was identified as the most prominent peak between 80 and 150 ms after the stimulus onset on the central medial electrode Cz. The N1 mean peak amplitude was calculated considering an interval of 60 ms cantered at the maximum peak (Table 2).

Table 2 - *N1 latency (ms) and amplitude (µV) values of the vowels for each subject and mean values (with standard deviation)*

N1	Latency		Amplitude	
	[a]	[i]	[a]	[i]
Sub 1	108	109	−1.98	−2.36
Sub 2	111	104	−.89	−1.11
Sub 3	104	96	−.44	−.54
Sub 4	111	116	−3.64	−4.49
Sub 5	107	107	−1.15	−2.04
Sub 6	107	111	−2.56	−3.31
Sub 7	111	109	−1.77	−1.66
Sub 8	108	120	−2.33	−1.76
Sub 9	95	111	−.74	−1.05
Sub 10	111	111	−1.01	−1.48
Sub 11	124	120	−2.24	−3.06
Sub 12	111	116	−1.60	−1.38
mean	108 (6.6)	110 (6.8)	−1.69 (0.9)	−2.01 (1.1)

3.3 Source Analysis

Estimation of the N1 intracranial sources was carried out using the BESA 2000. We used the spatiotemporal source analysis of BESA that estimates location, orientation and time course of equivalent dipolar sources by calculating the scalp distribution obtained for a given model (forward solution). This distribution was then compared to that of the actual auditory evoked potentials (AEP). Interactive changes in source location and orientation lead to minimization of residual variance between the model and the observed spatiotemporal AEP distribution. The optimal set of parameters was found in an iterative manner by searching for a minimum in

the compound cost function. Latency ranges for fitting N1 was between 100 and 120 ms. The source generator to each vowel was calculated on the grand average N1 waveform and was modelled as a symmetric pair of equivalent current dipoles (ECD) one on each hemispere. The accuracy of the source model was evaluated by measuring its residual variance as a percentage of the signal variance, as described by the model, and by applying residual orthogonality tests (ROT; e.g., Bocker et al., 1994). The resulting individual time series for the dipole moments (the source waves) were subjected to an orthogonality test, referred to as a source wave orthogonality test (SOT; Bocker et al., 1994).

3.2 Statistical Analysis

A t-test was run to compare the N1 amplitude and latency values at the Cz electrode to the vowels [a] and [i]. Furthermore, a 2x4 repeated measure analysis of variance ANOVA having as within factors Vowel type [a, i] and Electrode Type, was run on N1 amplitudes to test hemispheric asymmetries. Analysis was carried out at the electrode sites where the components were maximal in amplitude: at C1, CP1 for the left hemisphere, and C2 and CP2 for the right hemisphere.

   Differences in source localization within the left hemisphere (the right ECD was specular) were evaluated by means of one-way ANOVAs to test distance between the vowels on each spatial plane (anterior-posterior; x, lateral medial; y and superior-inferior; z), and the absolute Euclidean distance. All statistical results were corrected with the Greenhouse–Geisser method, whereas Bonferroni post-hoc method was applied to adjust the univariate output of repeated measurements of variances for violations of the compound symmetry assumption. In source analysis, Turkey (HSD) post-hoc method was applied. All t-statistics were evaluated for significance at the 5% level.

4. *Results*

4.1 Vowel comparison: Analysis on the N1 responses

Figure 2A shows the N1 waveforms at Cz electrode, where the activity was prominent for each vowel, in black line for [a] and in dashed line for [i]. From statistically analysis emerged that the vowels differ significantly in terms of the N1 peak amplitude at Cz electrode, $t(11)=2.40$, $p<0.05$. The vowel [i] recorded a higher value than [a]; N1 amplitude to [a] was of -1.69 μV whereas N1 to [i] had a peak of -2.01 μV. In contrast, the N1 peak latency did not statistically differ, $t(11)=0.98$, $p = 0.371$. The response to [a] had a latency of 108 ms, the response to [i] of 110 ms.

Figure 2 - *(A) Grand-averaged N1 responses (n=12 subjects) elicited by the vowel [a] (black line) and by [i] (dashed line). (B) shows the topographic maps of the N1 activities*
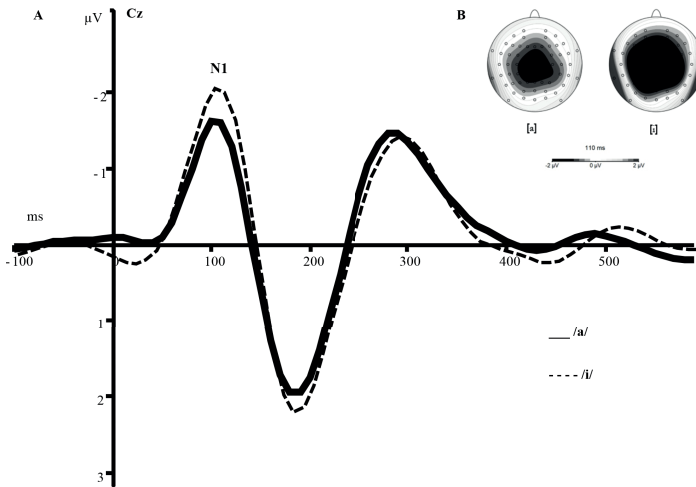
As showed in the topographical maps in Figure 2B, the N1 responses to [a] and [i] showed a different scalp distribution, which is more medial for [a] and more laterally distributed for [i]. Laterality ANOVA revealed a significant difference for the main effect of Vowel Type, $F(1,11)=4.44$, $p<0.05$ showing stronger negative activity in CP1 electrode for [i] relative to [a]; yet, hemisphere affected N1 modulation, $F(3,33)=3.11$, $p<0.05$, showing a larger activity in the left hemisphere for both vowels.

4.2 Source Analysis

The intracranial localization of the N1s sources for the two vowels is shown in Figure 3. The N1 response to each vowel was localized within the bilateral primary auditory cortex in Brodmann area (BA) 41, in the bilateral STG, approximately.

Figure 3 - *Intracranial source localization of N1s. In red the ECD to the vowel [i] in black ECD to the vowel [a]*

Averaging the Talairach coordinates across vowels the mean locations resulted ±52, 14, 10 (mm on x, y, z planes). Table 3 reports the Talairach coordinates within the left hemisphere.

Table 3 - *Mean (and standard deviation) Talairach coordinates of the source location of the two vowels for the N1 response in the left hemisphere*

N1	X	Y	Z
[a]	−49 (7)	−9 (2)	11 (4)
[i]	−54 (8)	−18 (4)	9 (3)
Mean	−52	−14	10

Statistical analysis showed that the vowel type influenced the N1 source location along the anterior-posterior y-axis (F(1,11)=9.87, p<0.01), the N1 generators were separated by a distance of 9 mm revealing a more anterior location for the vowel [a] relative to [i]. However, along the medial-lateral axis (x), even though the distance between vowels did not reach the statistically significance (F(1,11)=4.21, p=0.06), the N1 cortical generator to [a] was more lateral than the N1 source to [i], (5 mm). Along the superior-inferior z-axis, the difference (1 mm) was not significant (F(1,11)<1, ns). The Euclidean distance between the two vowels (10 mm) was significant (F(1,11)=11.05, p<0.01).

5. *Discussion*

The results of the present study suggest that the electric N1 response correlates to the functioning of the auditory cortex such that it mirrors its tonotopic structure. Therefore, N1 reveals as a suitable tool for observing the earliest steps of the brain mechanisms of extraction of linguistically relevant information. Specifically, the N1 amplitude pattern seems to index the cortical responsiveness of the auditory system to the spectro-temporal features of the natural Italian vowels [a] and [i]. The vowel [a] with close formant peaks elicited a reduction of the auditory activity at the scalp relative to vowel [i], which is characterized by a large F2-F1 distance of about 2000 Hz (cf. Table 1). A weaker auditory activity for [a] than for [i] is explained by referring to the so-called inhibitory formant frequency principle whereby the spectral relation rather than the extraction of separate formant peaks allows the vowel distinction (Ohl, Scheich, 1997; Diesch, Luce, 2000). These results are in line with previous MEG studies that have shown the effects of the acoustic structure of German (Obleser et al., 2003; 2004), Russian (Shestakova et al., 2004), and French vowels (Mäkelä et al., 2003) on the auditory mechanisms underlying speech processes.

Furthermore, the topographic differences between the vowels, showed a lateral position of the ECD source elicited by the vowels [i] relative to the ECD of the vowel [a] (Figure 2B) suggesting us the presence of different cortical generators (Näätänen, Picton, 1987). By source analysis, we found that the cortical generators of the activities evoked by the vowels were in primary auditory cortex (A1), mapped in Brodmann area BA 41, bilaterally. Imaging data has widely demonstrated the existence of a sophisticated dual-stream architecture of the auditory pathway where the core regions (BA41), work selectively during the speech information process-

ing (for a review see Hickok, Poeppel, 2004; Scott et al., 2003). From our data, we can assume that the N1 responses index the start of the sensory-based perceptual processes from the bilateral A1, (Hickok, Poeppel, 2007) thus triggering the speech recognition mechanisms (Scott et al., 2003; Formisano et al., 2008; Hickok, Poeppel, 2007; De Witt, Rauschecker, 2012). The displacement of the N1 generators along the auditory cortices gives further confirmation in that perspective. ECD to [a] results to be anterior to ECD source to [i] supporting the assumption that the depth of the source is stimulus frequency-dependent; in other word, higher frequency, deeper the location of the ECD source accounting for the N1 waves (Woods, 1995; Eulitz et al., 1995). In our study, the deep location of the ECD to [i] along the anterior-posterior axis, allows us to assume that different cortical loci are given by a different vowel acoustic structure. Actually, the spectral envelope of the vowels, as mapped in the F2-F1 acoustic space, affects the neural sensitivity of the auditory tissues (Diesch et al., 1996; Mesgarani et al., 2008; 2014). The study thus, confirms that the cortical activity reflects the main spectro-acoustic dissimilarities of the vowels, and it supports tonotopy as one of the most prominent organizing rules of the auditory cortices.

6. *Conclusion*

Exploiting for the first time the ERP N1 component for speech cortical representation, we provided further confirmation of the existence of tonotopy maps along the auditory system. Even though we are aware of some important limits of the present study such as the presentation of only two types of stimuli, it is worth noting that this is the first investigation on the Italian vowel system. The vowels [a] and [i] represent the two maximally different vowels within numerous vocalic systems therefore, they were suitable tools for the initial phases of our tonotopic inquiry. We expected that vowels differing in acoustic and articulator patterns could determine discrete cortical maps. Actually, we are working on the entire Salento Italian vowel system to establish specifically, how information on phonemes is implemented in the human brain with the aim to investigate to what extent the acoustic properties or rather the phonological features influence the cortical representation of speech sounds.

References

ALAIN, C., ARNOTT, S.R., HEVENOR, S., GRAHAM, S. & GRADY, C.L. (2001). "What" and "where" in the human auditory system. In *Proceedings of the national academy of sciences*, 98 (21), 12301-12306.

ALBANO LEONI, F., MATURI, P. (Eds.) (1995). *Manuale di fonetica*. Roma: Carocci.

BOCKER, K.B., BRUNIA, C.H. & CLUITMANS, P.J. (1994). A spatio-temporal dipole model of the readiness potential in humans finger movement. In *Electroenchephalogy and Clinical Neurophysiology*, 91, 275-285.

COHEN, D., CUFFIN, B.N., YUNOKUCHI, K., MANIEWSKI, R., PURCELL, C., COSGROVE, G.R., IVES, J. , KENNEDY, J.G. & SCHOMER, D.L. (1990). MEG versus EEG localization test using implanted sources in the human brain. In *Annals of Neurology*, 28, 811-817.

COHEN, D., HALGREN E. (2003). Magnetoencephalography (Neuromagnetism). In *Encyclopedia of Neuroscience*. Elsevier, 615-622.

DELATTRE, P., LIBERMAN, A.M., COOPER, F.S. & GERSTMAN, L.J. (1952). An experimental study of the acoustic determinants of vowel color; observations on one-and two-formant vowels synthesized from spectrographic patterns. In *Word,* 8, 195-210.

DEWITT, I., RAUSCHECKER, J.P. (2012). Phoneme and word recognition in the auditory ventral stream. In *Proceedings of the National Academy of Sciences*, 109 (8), E505-E514.

DIESCH, E., EULITZ, C., HAMPSON, S. & ROSS, B. (1996). The neurotopography of vowels as mirrored by evoked magnetic field measurements. In *Brain Language*, 53, 143-168.

DIESCH, E., LUCE, T. (1997). Magnetic fields elicited by tones and vowel formants reveal tonotopy and nonlinear summation of cortical activation. In *Psychophysiology*, 34 (5), 501-510.

DIESCH, E., LUCE, T. (2000). Topographic and temporal indices of vowel spectral envelope extraction in the human auditory cortex. In *Journal of cognitive Neuroscience*, 12 (5), 878-893.

EULITZ, C., DIESCH, E., PANTEV, C., HAMPSON, S. & ELBERT, T. (1995). Magnetic and electric brain activity evoked by the processing of tone and vowel stimuli. In *Journal of Neuroscience*, 15, 2748-2755.

EULITZ, C., EULITZ, H. & ELBERT, T. (1997). Differential outcomes from magneto-and electroencephalography for the analysis of human cognition. In *Neuroscience letters*, 227 (3), 185-188.

FORMISANO, E., DE MARTINO, F., BONTE, M. & GOEBEL, R. (2008). 'Who' is saying "what"? Brain-based decoding of human voice and speech. In *Science*, 322, 970-973.

GEVINS, A., LE, J, MARTIN, N.K., BRICKETT, P., DESMOND, J. & REUTTER, B. (1994). High resolution EEG: 124-channel recording, spatial deblurring and MRI integration methods. In *Electroencephalography and clinical Neurophysiology*, 90 (5), 337-358.

HICKOK, G., POEPPEL, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. In *Cognition*, 92 (1), 67-99.

HICKOK, G., POEPPEL, D. (2007). The cortical organization of speech processing. In *Nature Reviews Neuroscience*, 8, 393-402.

HICKOK, G., HOUDE, J. & RONG F. (2011). Sensorimotor integration in speech processing: computational basis and neural organization. In *Neuron*, 69 (3), 407-422.

KENT, R.D. (Ed.) (2002). *Acoustic analysis of speech*. Singular Publishing Group.

MÄKELÄ, A.M., ALKU P. & TIITINEN, H. (2003). The auditory N1m reveals the left-hemispheric representation of vowel identity in humans. In *Neuroscience Letters*, 353, 111-114.

MANCA, A.D., GRIMALDI, M. (2013). Perception, production, articulation and imagery articulation of Italian vowels: an ERP study. In *Proceedings of the 9th Conference of Multimodalità e Multilingualità: La sfida più avanzata della comunicazione orale.* AISV, (Eds.), Bulzoni, 213-217.

Mesgarani, N., David, S.V., Fritz, J.B. & Shamma, S.A. (2008). Phoneme representation and classification in primary auditory cortex. In *Journal of Acoustical Society of America*, 123, 899-909.

Mesgarani, N., Cheung, C., Johnson, K. & Chang, E.F. (2014). Phonetic feature encoding in human superior temporal gyrus. In *Science*, 343, 1006-1010.

Näätänen, R., Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. In *Psychophysiology*, 24 (4), 375-425.

Obleser, J., Elbert, T., Lahiri, A. & Eulitz, C. (2003). Cortical representation of vowels reflects acoustic dissimilarity determined by formant frequencies. In *Cognitive Brain Research*, 15, 207-213.

Obleser, J., Lahiri, A. & Eulitz, C. (2004). Magnetic Brain response mirrors extraction of phonological features from speakers vowels. In *Journal of Cognitive Neuroscience*, 16, 31-39.

Obleser, J., Scott, S.K. & Eulitz, C. (2006). Now you hear it, now you don't: Transient traces of consonants and their unintelligible analogues in the human brain. In *Cerebral Cortex*, 16, 1069-1076.

Obleser, J., Leaver, A.M., Vanmeter, J. & Rauschecker, J.P. (2010). Segregation of vowels and consonants in human auditory cortex: evidence for a hierarchical organization. In *Frontiers of Psychology,* 1 (232).

Ohl, F.W., Scheich, H. (1997). Orderly cortical representation of vowels based on formant interaction. In *Proceedings of of the National Academy of Sciences,* 94, 9440-9444.

Okada, K., Rong, F., Venezia, J., Matchin, W., Hsieh, I.H., Saberi, K. & Hickok, G. (2010). Hierarchical organization of human auditory cortex: evidence from acoustic invariance in the response to intelligible speech. In *Cerebral Cortex*, 20 (10), 2486-2495.

Pantev, C., Bertrand, O., Eulitz, C., Verkindt, C., Hampson, S., Schuierer, G. & Elbert, T. (1995). Specific tonotopic organizations of different areas of the human auditory cortex revealed by simultaneous magnetic and electric recordings. In *Electroencephalography and clinical neurophysiology*, 94 (1), 26-40.

Poeppel, D., Phillips, C., Yellin, E., Rowley, H.A., Roberts, T.P. & Marantz, A. (1997). Processing of vowels in supratemporal auditory cortex. In *Neuroscience Letters*, 221, 145-148.

Rauschecker, J.P., Tian, B. (2000). Mechanisms and streams for processing of "what" and "where" in auditory cortex. In *proceedings of the National Academy of Sciences*, 97 (22), 11800-11806.

Roberts, T.P., Ferrari, P., Stufflebeam, S.M. & Poeppel, D. (2000). Latency of the auditory evoked neuromagnetic field components: stimulus dependence and insights toward perception. In *Journal of Clinical Neurophysiology*, 17 (2), 114-129.

Scott, S.K., Blank, C.C., Rosen, S. & Wise, R.J. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. In *Brain*, 123 (12), 2400-2406.

Scott, S.K., Johnsrude, I.S. (2003). The neuroanatomical and functional organization of speech perception. In *Trends in Neurosciences*, 26 (2), 100-107.

Shestakova, A., Brattico, E., Soloviev, A., Klucharev, V. & Huotilainen, M. (2004). Orderly cortical representation of vowel categories presented by multiple exemplars". In *Brain Cognitive Research,* 21, 342-350.

Stevens, K.N. (2002). Toward a model for lexical access based on acoustic landmarks and distinctive features. In *Journal of Acoustical Society of America*, 111 (4), 1872-1891.

Virtanen, J., Ahveninen, J., Ilmoniemi, R.J., Näätänen, R. & Pekkonen, E. (1998). Replicability of MEG and EEG measures of the auditory N1/N1m-response. In *Electroencephalography and clinical Neurophysiology*, 108, 291-298.

Wise, R.J., Scott, S.K., Blank, S.C., Mummery, C.J., Murphy, K. & Warburton, E.A. (2001). Separate neural subsystems withinWernicke's area'. In *Brain*, 124 (1), 83-95.

Wolpaw, J.R., Penry, J.K. (1975). A temporal component of the auditory evoked response. In *Electroencephalography and Clinical Neurophysiology*, 39, 609-620.

Woods, D.L. (1995). The component structure of the N1 wave of the human auditory evoked potential. In *Electroencephalography and Clinical Neurophysiology-Supplements only*, (44), 102-109.

Zatorre, R.J., Belin, P. (2001). Spectral and temporal processing in human auditory cortex. In *Cerebral Cortex*, 11 (10), 946-953.