

SELECTIVE ASSOCIATIVE PHONAGNOSIA AFTER RIGHT ANTERIOR TEMPORAL STROKE

Simona Luzzi¹, Michela Coccia¹, Gabriele Polonara², Carlo Reverberi³, Gabriella Ceravolo¹,
Mauro Silvestrini¹, Fabio Fringuelli⁴, Sara Baldinelli¹, Leandro Provinciali¹ and Guido Gainotti⁵,
6

¹ Department of Clinical and Experimental Medicine, Polytechnic University of Marche, Italy

² Department of Odontostomatological and Specialized Clinical Sciences, Polytechnic University
of Marche, Italy.

³ Department of Psychology, University Miano-Bicocca, NeuroMi Center for Neuroscience, Milan,
Italy

⁴ Department of Nuclear Medicine, Ospedali Riuniti of Ancona, Italy

⁵ Institute of Neurology - Catholic University of the Sacred Heart – Rome – Italy

⁶ IRCCS Fondazione Santa Lucia, Department of Clinical and Behavioral Neurology, Rome

The article is accepted for publication in Neuropsychologia

doi: 10.1016/j.neuropsychologia.2017.05.016

Correspondence to:

Simona Luzzi

Clinica Neurologica

Dipartimento di Medicina Sperimentale e Clinica

Università Politecnica delle Marche – Ospedali Riuniti di Ancona

Via Conca, 1

Torrette di Ancona

Italy

Tel. 071/5964530

Fax. 071/887262

e-mail: simonaluzzi@yahoo.it; s.luzzi@univpm.it

ABSTRACT

We report the case of a 48 year old men who developed a selective impairment in famous voice recognition after ischemic stroke in right subcortical structures (lenticular nucleus and head of the caudate) and right anterior temporal lobe. He underwent fibrinolytic treatment. During the following days he progressively recovered and was discharged without neurological focal sign. Patent foramen ovale was found. When he got back to his house he noticed that he was unable to recognize the voice of his favoured singers and needed to ask who was the singer to his relatives. Neuropsychological examination revealed a selective impairment in famous voice recognition in the absence of alteration of voice perception, face perception and famous face recognition. All other neuropsychological domains were spared. In particular language, memory and executive functions were intact. Neuroimaging carried out by means of PET and MRI revealed two small ischemic lesions in the right subcortical region, involving lenticular and caudate nuclei and in the right temporal pole. To our knowledge, this is the first case described in literature of a patient showing a selective associative phonagnosia after right anterior temporal stroke. The present case helps to clarify the brain circuits underlying famous voice recognition and adds evidence in favour of a right hemisphere involvement in processing knowledge of familiar voices. These findings are discussed in relation to current models of brain organization of person-specific and general semantic knowledge.

Key words: associative phonagnosia; right temporal pole; famous voice recognition

1. INTRODUCTION

Due to the variety and specificity of relations that human beings have with members of the social groups to which they belong, identification of familiar people is a fundamental biological function of the human species. For this reason, a complex multimodal recognition system has evolved in the brain to quickly and efficiently accomplish this difficult process of individual persons identification. Faces and voices are the most important stimuli in the visual and auditory domains through which we recognize familiar people and we interact with other persons, because they convey critical information not only about identity, but also about gender, age and emotional status of known and unknown people. Clinical and functional neuroimaging investigations have shown that highly specific brain structures underlie the identification of familiar people through face and voice, even if the neuro-anatomical substrates of face processing and the neuropsychological investigations of face recognition disorders have been much more intensively and extensively investigated than the homologous aspects of voice processing and voice recognition disorders.

As for the brain structures involved in face recognition, they are more represented in right-sided structures, placed along the ventral stream of visual processing (Sergent, Ohta & MacDonald, 1992). The Kanwisher, McDermott & Chun's (1997) seminal paper argued for the existence in the Fusiform Gyrus of a processing module called the Fusiform Face Area (FFA), but further investigations, guided by an influential neural model proposed by Haxby, Hoffman & Gobbini (2000), suggested that face processing may involve a larger network, that includes, in addition to the FFA, the occipital face area /OFA (Gauthier, Tarr, Moylan, Skudlarski, Gore & Anderson, 2000), the superior temporal sulcus /STS (Puce, Allison, Bentin, Gore & McCarthy, 1998) and their

interconnections (Fox, Iaria & Barton, 2009). In more recent years, new findings have suggested the need of adding to these areas the anterior temporal lobe/ATL (Barton, 2008; Gainotti & Marra, 2011; Duchaine & Yovel, 2015). These different regions should play a different, hierarchically organized role in face processing. According to this model, the OFA should show sensitivity to physical properties of a face (Pitcher, Walsh & Duchaine, 2011), the FFA could shape an holistic face configuration (Duchaine & Yovel, 2015) and the anterior temporal structures could be involved in perceiving identity of a familiar person. (Gainotti & Marra, 2011; Gainotti, 2013; Duchaine & Yovel, 2015).

Even if Belin, Bestelmeyer, Latinus & Watson (2011) have stressed the 'auditory face' model of voice perception, our knowledge of the structures involved in voice recognition is much less detailed than that of structures involved in face recognition. Voice-selective cortical areas (which show a greater response to vocal sounds than to non-vocal sounds) have, indeed, been identified by fMRI studies along the middle and anterior superior temporal sulcus (STS) and the superior temporal gyrus (STG), (Belin, Zatorre, Lafaille, Ahad, & Pike, 2000; Belin, Zatorre & Ahad, 2002; Warren, Scott, Price, & Griffiths, 2006; Bethmann, Scheich & Brechmann, 2012). These areas have been called 'temporal voice areas' (TVA) and some authors have suggested that, just as in face processing, there may be a gradient between posterior and the anterior components of the TVA. According to this model, the posterior areas surrounding the STS might be mainly engaged in acoustic-based voice processing (Von Kriegstein & Giraud, 2004; Mathias & Von Kriegstein, 2014), whereas speaker identification might be sustained by cortical structures in the ATL (Von Kriegstein & Giraud, 2004; von Kriegstein, Kleinschmidt, Sterzer & Giraud, 2005; Bethmann et al., 2012). Furthermore, as in the case of face recognition, a more important role of the right TVA in voice recognition has been documented by von Kriegstein, Eger, Kleinschmidt & Giraud (2003) and by von Kriegstein & Giraud (2004). Yovel & Belin (2013) have therefore suggested that highly similar mechanisms subsume the cognitive and neural processing of face and voice recognition, but Pernet, McAleer, Latinus, Gorgolewski, Charest, Bestelmeyer, Watson, Fleming, Crabbe, Valdes-

Sosa & Belin (2015) have recently stressed the high inter-individual variability in precise anatomical location of the human voice areas.

The prevalence of our knowledge of face over that of voice recognition is even greater if we pass from the neuroanatomical to the clinical data, because a great number of studies have investigated face recognition disorders, labelled 'prosopagnosia' (from the Greek: "prosopon" = "face", "agnosia" = "not knowing") by Bodamer (1947), whereas only a handful of papers have been devoted to voice recognition disorders, labelled 'phonagnosia' by Van Lanker & Canter (1982). An important analogy that has been reported between 'prosopagnosia' and 'phonagnosia' concerns the distinction between an 'apperceptive' and an 'associative' form of these recognition disorders. In both cases the term 'apperceptive' indicates a high-level perceptual impairment, which provokes a defect not only in the recognition of familiar faces or voices, but also in the discrimination of unfamiliar faces or voices. On the other hand, the term 'associative', which has been attributed to a memory or associative disturbance, consists of a specific defect in the recognition of familiar faces or voices, in the absence of problems in the discrimination of unfamiliar faces or voices and of other perceptual disorders (De Renzi, Faglioni, Grossi & Nichelli, 1991; Van Lancker, Cummings, Kreiman & Dobkin, 1988; Peretz, Kolinsky, Tramo, Labrecque, Hublet, Demeurisse & Belleville, 1994; Hailstone, Crutch, Vestergaard, Patterson & Warren, 2010). Two recent papers have taken into account the general problem of voice recognition disorders in brain-damaged patients. The first is a review of voice recognition disorders in various kinds of brain pathology conducted by Gainotti (2011), whereas the second is a research paper in which Hailstone, Ridgway, Bartlett, Goll, Buckley, Crutch & Warren (2011) studied voice processing and the correlative neuroanatomical data in Semantic Dementia (SD) and in Alzheimer's disease. Results of both Gainotti's (2011) general review and Hailstone et al.'s (2011) investigation of demented patients showed that voice recognition disorders are mainly due to right temporal lesions, similarly to face recognition disorders and that famous voice recognition disorders can be dissociated from unfamiliar voice discrimination impairments. This dissociation between impaired recognition of familiar voices and

defective discrimination of unfamiliar voices had already been described by Van Lanker et al. (1988) and by Perez et al. (1994) in patients with focal brain damage. If we focus attention on patients with ‘associative’ forms of prosopagnosia or of phonagnosia, we see that there is a striking contrast between the high number of patients with associative prosopagnosia reported in the literature (see reviews in Gainotti & Marra, 2011 and in Gainotti, 2013) and the unique case of ‘Progressive associative phonagnosia’ reported by Hailstone et al. (2010) in a patient (QR) with behavioral variant of Fronto-Temporal Dementia (bvFTD). From the neuroanatomical point of view, this patient showed a bilateral fronto-temporal atrophy somewhat accentuated in the right anterior temporal lobe, but extending posteriorly within the temporal lobe and including the superior temporal sulcus. Since we recently had the chance of studying a patient who showed a highly selective breakdown of familiar voice recognition after an ischemic stroke of the right temporal pole and of a right subcortical region involving the lenticular and caudate nuclei, we intend to present this case here, because of his interest from the clinical and neuroanatomical point of view.

2. CASE REPORT

2.1. Clinical history

Patient MM, a 48 -year-old right-handed Italian men, with 8 years of education, came to our attention complaining of a disturbance in recognition of voices of singers appeared after an ischemic stroke.

He had been admitted to the hospital for the sudden onset of sensitive and motor disturbances on the left side of the body about two months before coming to our attention. Neurological evaluation at admission in the hospital revealed a mild superficial hypoesthesia in his left limbs and facial district together with mild left hemiparesis. CT scan at the admission did not reveal pathological changes.

He underwent intravenous fibrinolytic therapy with alteplase. Clinically he improved in the following days and the neurological examination at discharge was reported to be normal. A repeated CT scan showed a right subcortical ischemic lesion in the internal capsule involving the caudate nucleus.

Twenty-four hours after fibrinolytic therapy a platelet aggregation inhibitor treatment with acetylsalicylic acid was started. Furthermore, since a patent foramen ovale had been found, he underwent percutaneous closure. After two weeks the patient was discharged apparently without neurological focal signs. When he returned home he spent a few days before getting back to work. He used music in his free time. He noticed he was not able to recognize the voice of known singers and in several occasions he asked to his relatives to confirm if the singer was the original singer or someone else. Worried by this unusual symptom he came to our attention to understand why his favourite singers now were perceived as “strangers”. He referred he did not have problems in perceiving music and sounds. In particular he said that he was able to recognize famous songs only from the melody and he was able to sing some of them. The only altered feature was in the quality of the singer’s voice. He was able to recognize the sex and approximately the age of the singers but was unable to recognize their identity. He did not note problems in identifying the voices of his relatives, but he had never identified them exclusively by their voice, because he interacted with them face-to-face. He was asked if he was able to recognize their voice at the telephone, but he answered that every time some of his relatives or friends gave him a ring call, their name appeared first on his mobile screen so that he knew the identity of the person who was calling him.

2.2. Neurological examination

Neurological examination revealed a mild proprioceptive problem in his left hand fingers (6/8 Hz) in absence of detectable superficial sensitivity problems. No motor problems were present neither sign of pyramidal impairment were detectable (Babinski and Hoffman signs were absent, profound and superficial reflexes were normal, tone was normal).

2.3. Neuropsychology

2.3.1 Background neuropsychology

Background neuropsychological evaluation aimed to explore the main neuropsychological domains. Tests and performances are shown in Table 1 and included the MMSE (Folstein, Folstein & McHugh, 1975) as a measure of general cognitive status, the Bisyllabic Word Span and the Corsi Blocks (Spinnler & Tognoni, 1987) to assess short term memory, the Rey's auditory verbal learning test/Rey AVLT (Spinnler & Tognoni, 1987), and the delayed recall of the Rey-Osterrieth Complex Figure B (Luzzi, Pesallaccia, Fabi, Muti, Viticchi, Provinciali & Piccirilli, 2011) to evaluate long term memory. Furthermore, the Luria's Motor Sequences (Luzzi et al., 2011), the Stroop test (Luzzi et al., 2011) and the phonological fluency test (Caltagirone, Gainotti, Masullo & Miceli, 1979) were used to assess executive functions. The Visual Object and Space Perception Battery (Warrington & James, 1991) was used to evaluate visuo-perceptual and visuo-spatial abilities, whereas constructional praxis was assessed by means of the Rey-Osterrieth Complex Figure B copy version (Luzzi et al., 2011) and ideo-motor praxis using the test described by Luzzi, Piccirilli, Pesallaccia, Fabi & Provinciali (2010). Semantic-lexical functions were assessed using the easy picture naming test and the forced-choice word-picture matching test proposed by Snowden, Thompson & Neary (2004) and a test of categorical word fluency (Luzzi et al., 2011).

Results: in all tests the patient obtained normal performance.

2.3.2. Language

Although a first level language examination did not reveal problems, language was examined in detail with focus on single words naming and comprehension, semantic association tests and linguistic prosody. Laiacona Test (Laiacona, Trivelli, Barbarotto & Capitani, 1993) was used to explore naming and single-word comprehension. Conceptual-semantic associations were studied by means of the verbal and pictorial versions of the Pyramids and Palm trees test/PPTT (Howard & Patterson, 1992), using the Italian normative data, provided by Gamboz, Coluccia, Iavarone and

Brandimonte (2009). Linguistic prosody was explored by means of the Prosody subtest of the Battery for Language of the Right Hemisphere [Batteria sul linguaggio dell'emisfero destro /BLED] (Rinaldi, Marangolo & Lauriola, 2006): the subject was asked to listen to 10 meaningful couples of sentences. In five couples the sentence was read by an actor with the same linguistic prosody, while in 5 pairs the same sentence was read with a different linguistic prosody. A two-item example was provided to verify if the patient could understand the meaning of the task proposed. The maximum score was 10. These data too are reported in Table 1.

Results: single word naming and comprehension, semantic association and linguistic prosody recognition were within the normal range.

2.3.3. Not musical sound recognition

Not musical sounds were explored by means of a new experimental battery composed by 70 environmental sounds. The tasks composing the 'battery' and the corresponding performances are reported in Table 2, together with other high level auditory tasks and with tasks of face and voice perception and recognition. Seven categories of not musical sounds, concerning 3 living, 3 non-living categories and musical instruments and consisting each of 10 items were explored. The living categories concerned: - human sounds (e.g. sneeze), - pets sounds (e.g. cat meowing) and - wild animals (e.g.: lion roaring); the non living categories concerned: - environmental noises (e.g.: noise of a waterfall), - vehicles sounds (e.g. motorbike chatter), and - appliances (e.g. noise of a vacuum cleaner). The battery included the following tasks:

- a) sounds naming: the subject was asked to produce the name of the sound listened;
- b) sound-to-picture matching task and 3) sound-to-word matching task: to the subject was asked to listen to a sound and match it to the target stimulus (picture or written word respectively) from an array of 4 items. For each item three distractors were provided: a visual, a semantic and an unrelated distractor

The maximum score was 70 both for naming and for the two matching tasks.

c) sound production, i.e. the ability of the patient to imitate a specific sound, was explored as well. A subset of 35 sounds were selected to this purpose, on the basis of a pilot study on 42 normal subjects. The patient performance was recorded and scored by three examiners who were asked to give a score of 1 for each sound which resembled the real one.

Results: Results obtained by M were at ceiling in all the non-musical sound recognition and sound production tests.

2.3.4. Emotional prosody

This test was described in detail by Piccirilli Sciarma & Luzzi (2000). Briefly, 10 meaningful and 10 meaningless sentences with different emotional prosody were presented. The patient had to recognize the emotional tone.

Results: Results obtained by M were at ceiling in the recognition of the emotional tone with which meaningful and meaningless sentences had been presented.

2.3.5. Musical abilities

Musical abilities were explored by a detailed battery previously described (see Piccirilli et al 2000). Briefly three sets of test were applied:

2.3.5.1. Musical perception tests: four tests were applied: 1) pitch discrimination (comparison between two notes); 2) tonal memory (comparison between two five-note motifs that were different for one note) 3) chord analysis (two, three and four note chords); 4) rhythmic memory (comparison between two rhythmic figures with four beats and without any change in pitch).

2.3.5.2. Recognition of musical characteristics: these tests aimed to check if the patient was able to identify musical characteristics and were composed by recognition of intensity, timbre (musical instruments), ascending and descending scales, familiar rhythms (e.g.: waltzer, tango etc).

Normative data, allowing to distinguish normal from pathological data were available for both the musical perception and the recognition of musical characteristics tests.

2.3.5.3. Recognition of familiar melodies: twenty familiar melodies (music from songs without words) were shown. They were selected with the collaboration of the patient's relative.

Results: music perception and recognition of musical characteristics were in the normal range. M was able to recognize all the familiar melodies that were presented.

2.3.6. Recognition of singer voice

Songs utilized for the recognition of familiar melodies were used. The test was composed by 40 items; in 20 items the song was chanted by the original artist (e.g.: “Marinella” from Fabrizio De Andrè), and in 20 items it was sung by another artist (e.g.: “Marinella” sung by Lucio Dalla). In this way, for each song there was one original version, i.e. the one chanted by the original artist and a second version sung by another artist. The patient was asked if the voice he was listening was or not the voice of the original singer. The question addressed to the patient was explicit (i.e.: Is the voice you heard the one of Fabrizio De Andrè ?”) so that possible misunderstandings were excluded. The total score was 40; scores of omissions (number of times the patient did not recognize the author) and false recognitions (the number of times the patient mistook another artist voice with the one of the author) were recorded. It was hard to find controls who shared the same background. His wife and his son were used as normal controls whereas the songs selection was performed with the help of his daughter. The two controls were able to identify correctly all the items.

MM was asked to classify the known singers in two groups according to the degree of familiarity: he was asked to classify the singers in “familiar” and “very familiar”. He classified as “familiar” 7 singers and “very familiar” 13 singers.

Results: Scores obtained by M in the recognition of singers’ voices were extremely low, with a prevalence of omissions over false recognition errors. Ten out of 12 of the singers who MM recognized were classified by him as “very familiar” while 3 out of 8 of the singers he did not recognize were classified as “very familiar”. Fisher exact test approached significance ($p= 0.062$).

2.3.7. Face and voice perception and recognition

Face perception was explored by means of the Benton and Van Allen (1968) face recognition test (BFRT), which requires identification of the same or of different unknown faces, seen from different perspectives or under different lighting, whereas voice perception and recognition of voices and faces of famous people were investigated by means of a new battery described and standardized on an Italian population by Quaranta, Piccininni, Carlesimo, Luzzi, Marra, Papagno, Trojano & Gainotti (2016).

2.3.7.1. Unknown voices discrimination: twenty stimuli, composed by two audio files, recording the voice of the same or different unknown persons and lasting about 15 s were consecutively presented. The subject was asked to say if the voices belonged to the same or to different persons. To reduce the possible interference of the sentence content, the stimuli were subdivided into four groups of five stimuli: same voice/ same phrase; same voice/different phrase; different voice/ same phrase; different voice/different phrase. The total score was obtained by summing the correct answers (range: 0-20).

2.3.7.2. Familiarity evaluation and famous person identification from voices (VO-REC): The VO-REC was formed by 60 items (40 famous voices and 20 non famous voices). In ‘familiarity evaluation and person identification from voices’ (VO-REC) each subject was asked to carefully listen to the audio fragment and provide a familiarity judgment (“is this voice familiar to you?”). The *familiarity score* was obtained by summing the number of voices correctly identified as famous or non-famous (score range: 0–60); a *false alarms score* (number of non-famous voices identified as famous; score range: 0–20) was also obtained.

If the answer concerning the familiarity judgment was positive and the voice belonged to a famous person, subjects were asked three further questions. The first two questions had a multiple choice format and explored the general and specific categories to which famous persons belonged. An example of general information is the following: “is this person involved in (a) politics; (b) entertainment; (c) sport; (d) civil society?”; an example of specific information is the following:

“is this entertainer involved in (a) cinema; (b) theatre; (c) music; (d) TV?”. The third question was open and the subject was asked to provide univocally identifying information about the person (i.e., movies’ titles, political roles and parties, etc.). One point was assigned to each correct answer (semantic score for each voice: 0–3; total semantic score: 0–120).

2.3.7.3. Familiarity evaluation and famous person identification from faces (FA-REC): The test was formed by black-and-white photos of the same famous persons who entered the VO-REC, along with twenty black-and-white photos of non-famous persons. Photos of famous and non-famous persons were comparable as for quality (definition and shading) of the material. The FA-REC was formed by 60 items (40 famous faces and 20 non famous faces). To avoid any interference with the VO-REC, the order of presentation of faces was random, and different from the order of presentation of voices.

Familiarity evaluation and person identification from faces (FA-REC) is structurally similar to the familiarity evaluation and person identification from voices.

M underwent the VO-REC and FA-REC in separate sessions with an interval of 2 months to avoid priming and interference effects. The audio files were presented using an audio-player software; both the examiner and the examinee used headphones to avoid environmental interferences.

Results: M performance were in the upper part of the normal range in tasks exploring voice and face discrimination and in famous face recognition. By contrast he showed problem in famous voice recognition with very low scores both in the familiarity evaluation and in the semantic questionnaire. No false alarms were registered.

2.3.7.4 Familiar voices recognition: this test was developed with the collaboration of the patient’s relatives which recordered a set of 16 familiar voices. Voices were recorded from relatives (wife, parents, daughters), friends, colleagues and neighbors. To all persons was asked to read the same sentence (a phrase from “The little prince”, Antoine de Saint-Exuperì). Their voice was recordered

and was matched to the voice of 16 unknown persons which shared similar age and sex with the M's familiar persons.

Results: M failed to recognize four familiar persons (two colleagues, a friend of him and a neighbor). He did not recognise as familiar unknown persons. M's wife was used as healthy control. She scored 16 out of 16.

=====

Insert about here

Table 1

=====

=====

Insert about here

Table 2

=====

3.NEUROIMAGING

3.1. RM imaging acquisition

MR structural image data were acquired by using a PHILIPS Achieva machine (1.5 Tesla) and a 8-channels head coil. We acquired axial 2D structural images using standard protocols for T1 (TR =

537 ms, TE = 12.5 ms, Slice thickness = 5 mm), T2 (TR = 5500 ms, TE = 120 ms, Slice thickness = 3 mm), and FLAIR (TR = 8000 ms, TE = 125 ms, Slice thickness = 5 mm).

3.2. PET imaging acquisition

PET imaging was performed in a silent room, with the subject's head fixed to a head holder with a plastic spacer to minimize the subject's head movement. The subjects received a mean dose of [18F] 2- deoxy-2-fluoro-D-glucose (FDG) of 185 ± 10 MBq (5 ± 0.27 mCi).

PET emission scan started approximately 30 min after FDG injection using a GE Discovery PET/CT 690 VCT scanner, with spatial resolutions of 4.8, 4.8 and 5.0 mm at full-width-half-maximum (FWHM) in radial, tangential and axial directions, respectively. A three-dimensional emission scan was performed for 15 min, and a post-injection transmission scan was performed using CT for tissue attenuation correction (Helical Full 0.6 s, 3.75 mm, 47 slices). PET imaging data was corrected for tissue attenuation. PET imaging data were reconstructed into 256×256 matrices based on an Iterative Reconstruction Algorithm: VPFX-S (VUE Point FX, 3.2 mm, 32/8).

3.3. PET data analysis

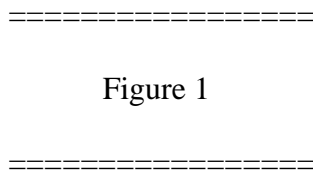
Imaging analysis was performed using SPM12. Preprocessing included rigid-body transformation (realignment) to correct for head movement. The images were then normalized to MNI space using non-linear warping implemented in SPM12 and finally smoothed with a Gaussian filter of 12 mm full-width at half maximum (FWHM) to increase the signal-to-noise ratio and to facilitate group analyses. Finally, individual global counts were normalized by proportional scaling to a mean value of 50 mg/100 mL/min. Linear contrasts were used to test for regionally specific differences between the single case and the healthy control group ($n = 34$), producing t-statistic maps in MNI standard space. We considered significant effects having $p < 0.05$ FWE at voxel level. We also reported results using a less conservative statistical threshold with $p < 0.05$ FWE corrected for multiple comparisons

at cluster level. Clusters were identified by using a voxel level threshold $p < 0.001$ not corrected for multiple comparisons.

3.3. Neuroimaging results

We evaluated where in the brain the patient a decrease in FDG uptake could be observed compared to a healthy control group. As expected from the location of the lesions visible in the structural MR images (Fig 1B), we observed a significant reduction in the right anterior middle temporal gyrus, the temporal pole, the right putamen and pallidum (Fig. 1A). When a less conservative statistical threshold was used ($p < .05$ FEW at cluster level) the FDG uptake reduction extended more posteriorly in the middle temporal gyrus and also involved the right insula, the right precentral cortex, the right middle cingulate and the right supplementary motor cortex.

Insert about here



4. DISCUSSION

Our patient (MM) showed a very selective defect of famous voice recognition both at the clinical level and on highly controlled neuropsychological tests. At the clinical level his unique complaint was the inability to identify by the quality of their voice singers who were very familiar to him, in spite of being able to recognize famous songs from their melody only and to sing some of them. In the general neuropsychological examination he obtained normal scores on tests of short-term and long-term verbal and visuo-spatial memory, executive functions, visuo-perceptual and spatial skills, praxis, language and conceptual thinking. Equally normal were results obtained on tasks involving comprehension and naming of non-musical sounds, interpretation of prosodic emotional tones,

musical perception, recognition of musical characteristics and of familiar melodies, unfamiliar faces and voices discrimination and familiar faces recognition. All tasks of non-verbal auditory perception and recognition (including unfamiliar voices discrimination) and all tasks of famous faces recognition were, therefore, perfectly carried out. He obtained, on the contrary, pathological scores on a controlled task in which he was requested to recognize the voice of famous singers and on a task requiring recognition (familiarity assessment) and identification (personal semantic evaluation) of famous persons through their voice. On the basis of these results, patient MM (being selectively unable to recognize familiar people from their voice) can be considered as a very clear and clean case of associative phonagnosia (Van Lancker et al., 1988; Hailstone et al., 2010).

The neuroanatomical correlates of this selective form of phonagnosia were less clean, because, in addition to a cortical lesion encroaching upon the posterior part of the right temporal pole, patient MM also showed another right subcortical lesion, involving lenticular and caudate nuclei. Since these subcortical lesions had probably provoked the sensori-motor disturbances shown just after his ischemic stroke by patient MM, but there are no reason to think that they had played a role in his familiar voice recognition disorders, we will focus our attention on the cortical lesion involving the right temporal pole. This lesion affected the anterior part of the right middle temporal gyrus, near the temporal pole, involving the antero-superior and sparing the infero-medial part of the right temporal lobe and extended caudally to all the anterior half of the right middle temporal gyrus. These neuroimaging data are quite consistent with the selective impairment of famous voice recognition shown by our patient, because the anterior-inferior banks of the STS, subsuming speaker identification (Von Kriegstein & Giraud, 2004; von Kriegstein et al., 2005; Bethmann et al., 2012) were selectively affected, whereas both the posterior parts of STS, involved in voice discrimination (Von Kriegstein & Giraud, 2004; Mathias & Von Kriegstein, 2014) and the infero-temporal and fusiform gyri, involved in face processing (Haxby, Hoffman & Gobbini, 2000; Barton, 2008; Gainotti & Marra, 2011) were completely spared. Our neuroimaging data are also in agreement with the neuroanatomical data reported by Hailstone et al. (2010) in a patient (QR) with

‘Progressive associative phonagnosia’, who showed a bilateral fronto-temporal atrophy more severe in the right anterior temporal lobe, which extended posteriorly within this lobe and included the superior temporal sulcus. The same data are, on the contrary, at variance with anatomo-clinical data reported by Van Lancker et al. (1988), who had suggested that recognition of familiar voices is impaired by damage to inferior and lateral parietal regions of the right hemisphere, whereas impairment of voice discrimination abilities is associated with temporal lobe damage of either hemisphere. Since the neuroimaging data reported by Van Lancker et al. (1988) had been obtained with a computer tomography (CT) scan, provided of poor discriminative power, we can say that anatomo-clinical data are consistent with functional neuroimaging experiments to suggest that voice-selective cortical areas can be identified in man along the lower bank of the superior temporal sulcus (STS) and the superior temporal gyrus (STG) and that speaker identification may be sustained by the rostral parts of these temporal voice areas. Taken together, these data are relevant with respect to the general organization of perceptual (face and voice) channels subsuming people recognition within the temporal lobes, because they suggest that the rostral parts of the right (superior) ‘temporal voice areas’ are the locus where familiar person are identified through their voice, and whose lesion produces associative phonagnosia, just as the ventral parts of the right ATL (vATLs) are the apex of a hierarchically organized face processing system, extending from the inferior occipital gyri to the vATLs and whose lesion can elicit ‘associative prosopagnosia’ (Gainotti & Marra, 2011; Collins & Olson, 2014; Liu, Pancaroglu, Hills, Duchaine & Barton, 2016). When the lesion of the ATLs prevails on the right side, but is bilateral (as is the case of the cortical atrophy in the right temporal variant of fronto-temporal degeneration), the breakdown affects not only the visual (face) and auditory (voice) channels through which familiar people are identified, but also their semantic representations or, in any case, those embodied (visual and auditory) components of person-specific semantic knowledge which are mainly subsumed by the right ATL. In these patients, face recognition disorders and voice recognition disorders cannot be considered as instances of ‘prosopagnosia’ and of ‘phonagnosia’, but of a ‘multimodal people

recognition disorder', because they are not restricted to the visual and , respectively, to the auditory modality, but affect both of them (Gainotti, 2007 and 2013; Josephs, Whitwell, Vemuri, Senjem, Boeve, Knopman et al., 2008; Busigny, Robaye , Dricot &, Rossion, 2009). Furthermore, in these patients the multimodal defect in familiar people recognition often coexists with a general disruption of conceptual representations, that is best explained with the 'hub-and-spokes' model of semantic knowledge (Patterson , Nestor & Rogers, 2007). In particular, the 'graded' version of this model (Rice, Hoffman & Lambon Ralph, 2015; Lambon Ralph, Jefferies , Patterson & Rogers, 2017) explains the differences between atrophy of the right and left ATL as a consequence of their differential connectivity with lateralized language and face processing networks. An alternative interpretation of these differences was advanced by Gainotti (2012 and 2015), who suggested that the format of conceptual representations and of person-specific knowledge may be mainly verbal at the level of the left ATL and mainly embodied (visual and auditory) at the level of the right ATL. Data supporting this position have been gathered, among others, by Snowden et al. (2004), Snowden, Thompson & Neary (2012) and Luzzi, Baldinelli, Ranaldi, Fabi, Cafazzo, Fringuelli et al. (2017). These authors have shown that semantic information accessed through faces and names is different depending on the prevalent side of atrophy (being lower through face when the atrophy prevails on the right side and through name when it prevails on the left side) and that an important overlap exists between disorders affecting famous people and disorders concerning conceptual representations. As a matter of fact, semantic dementia patients with predominantly left temporal lobe atrophy identified faces better than names and performed better on the picture than on the word version of the semantic memory 'Pyramids and Palm Trees' test (Howard and Patterson, 1992), whereas patients with right temporal lobe atrophy showed the opposite pattern of performance. Anyway, whatever may be the explanation of the differences between right and left ATL's atrophy with respect both to familiar people recognition and to general conceptual disorders, these differences are probably due to the bilateral representation of semantic knowledge. In our patient, where the vascular lesion was restricted to the right ATL, the defect of people recognition was

modality-specific, only affecting the vocal channel with a form of associative phonagnosia, and no defect was observed on the verbal or pictorial versions of the PTT test.

COMPETING INTEREST STATEMENT

There are not competing interests of any author of this paper.

FUNDING

The study was not sponsored.

REFERENCES

Barton, J. (2008). Structure and function in acquired prosopagnosia: lessons from a series of ten patients with brain damage. *Journal of Neuropsychology*, 2:197–225.

Belin, P., Bestelmeyer, P.E., Latinus, M. & Watson R. (2011). Understanding voice perception. *British Journal of Psychology*, 102(4), 711-25

Belin, P., Zatorre, R.J. & Ahad, P. (2002). Human temporal-lobe response to vocal sounds. *Brain Research Cognitive Brain Research*, 13, 17-26

Belin, P., Zatorre, R.J., Lafaille, P Ahad, P & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature* 403, 309-12.

Benton, A.L. & Van Allen, M.W. (1968) Impairment in facial recognition in patients with cerebral disease. *Transactions of the American Neurological Association* 93:38-42

Bethmann, A. Scheich, H. & Brechmann, A. (2012) The Temporal Lobes Differentiate between the Voices of Famous and Unknown People: An Event-Related fMRI Study on Speaker Recognition. *PLoS ONE* 7(10): e47626

Bodamer, J. (1947). Die Prosopagnosic. *Archiv für Psychiatrie und Nervenkrankheiten*, 179 (1-2), 6-53.

- Busigny, T., Robaye L, Dricot L, Rossion, B. (2009). Right anterior temporal lobe atrophy and person-based semantic defect: a detailed case study. *Neurocase*, 15, 485-508.
- Caltagirone, C., Gainotti, G., Masullo, C. & Miceli, G. (1979) Validity of some neuropsychological tests in the assessment of mental deterioration. *Acta Psychiatr Scandinavica*, 60, 50–56.
- Collins, J.A. & Olson, I.R. (2014). Beyond the FFA: The role of the ventral anterior temporal lobes in face processing. *Neuropsychologia*, 61, 65-79.
- De Renzi, E., E., Faglioni, P., Grossi, D., Nichelli, P. (1991). Apperceptive and associative forms of prosopagnosia. *Cortex*, 27(2), 213-221.
- Duchaine, B. & Yovel, G. (2015) A Revised Neural Framework for Face Processing. *Annual Review of Vision Science*, 1, 393-416, DOI: 10.1146/annurev-vision-082114-035518
- Folstein, M.F., Folstein, S.E. & McHugh, P.R. (1975). “Mini-mental state”. A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12, 189–198.
- Fox, C.J., Moon, S.Y., Iaria, G. & Barton, J.J. (2009). The correlates of subjective perception of identity and expression in the face network: an fMRI adaptation study. *Neuroimage*, 44(2), 569-80.
- Gainotti, G. (2007). Different patterns of famous people recognition disorders in patients with right and left anterior temporal lesions: A systematic review. *Neuropsychologia*, 45, 1591–1607.

Gainotti, G. (2011). What the study of voice recognition in normal subjects and brain-damaged patients tells us about models of familiar people recognition. *Neuropsychologia*, 49, 2273-2282.

Gainotti, G. (2012). The Format of Conceptual Representations Disrupted In Semantic Dementia: A Position Paper. *Cortex*, 48 . 521-529.

Gainotti, G. (2013). Is the Right Anterior Temporal Variant of Prosopagnosia a Form of 'Associative Prosopagnosia' or a Form of 'Multimodal Person Recognition Disorder'? *Neuropsychology Review*, 23 (2), 99-110.

Gainotti G. (2015). Is the difference between right and left ATLs due to the distinction between general and social cognition or between verbal and non-verbal representations? *Neuroscience and Biobehavioral Reviews*, 51, 296-312.

Gainotti, G. & Marra, C. (2011) Differential contribution of right and left temporo-occipital and anterior temporal lesions to face recognition disorders. *Frontiers in human neuroscience* 5:55.
doi:10.3389/fnhum.2011.00055

Gamboz, N., Coluccia, E., Iavarone, A. & Brandimonte, M.A. (2009). Normative data for the Pyramids and Palm Trees Test in the elderly Italian population. *Neurological Sciences*, 30(6), 453-458.

Gauthier, I., Tarr, M.J., Moylan, J., Skudlarski, P., Gore, J.C. & Anderson A.W. (2000). The fusiform “face area” is part of a network that processes faces at the individual level. *Journal of Cognitive Neuroscience*, 12, 495–504.

Hailstone, J.C., Crutch, S.J., Vestergaard, M.D., Patterson, R.D. & Warren, J.D. (2010). Progressive associative phonagnosia: a neuropsychological analysis. *Neuropsychologia*, 48(4), 1104-14.

Hailstone, J.C., Ridgway, G.R., Bartlett, J.W., Goll, J.C., Buckley, A.H., Crutch, S.J. & Warren, J.D. (2011). Voice processing in dementia: a neuropsychological and neuroanatomical analysis. *Brain*, 134(9), 2535-2547.

Haxby, J.V., Hoffman, E.A., Gobbini, M.I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, 4(6), 223-233.

Howard, D., & Patterson, K. (1992). *Pyramids and Palm Trees: access from pictures and words*. Bury St Edmunds (UK), Thames Valley Test Company,

Josephs, K.A., Whitwell, J.L., Vemuri, P., Senjem, M.L., Boeve, B.F., Knopman, D.S., Smith, G.E., Ivnik, R.J., Petersen, R.C., Jack, C.R. Jr. (2008). The anatomic correlate of prosopagnosia in semantic dementia. *Neurology*, 71(20), 1628-33.

Kanwisher, N., McDermott, J., Chun, M.M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, 17(11), 4302-4311.

Lambon Ralph MA, Jefferies E, Patterson K & Rogers, T.T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, 18, 42-55.

Liu, R.R., Pancaroglu, R., Hills, C.S., Duchaine, B. & Barton, J.J. (2016). Voice Recognition in Face-Blind Patients. *Cerebral Cortex*, 26(4), 1473-1487. doi: 10.1093/cercor/bhu240.

Luzzi, S., Baldinelli, S., Ranaldi, V., Fabi, K., Cafazzo, V., Fringuelli, F. et al. (2017). Famous faces and voices: Differential profiles in early right and left semantic dementia and in Alzheimer's disease. *Neuropsychologia*, 94, 118-128.

Luzzi, S., Pesallaccia, M., Fabi, K., Muti M, Viticchi G, Provinciali L & Piccirilli M (2011) Non-verbal memory measured by Rey-Osterrieth Complex Figure B: normative data. *Neurological Sciences*, 32, 1081-1089.

Luzzi, S., Piccirilli M, Pesallaccia, M., Fabi, K. & Provinciali, L. (2010) Dissociation apraxia secondary to right premotor stroke. *Neuropsychologia*, 48, 68-76.

Mathias, S.R. & von Kriegstein, K. (2014). How do we recognise who is speaking? *Frontiers in Bioscience* (Schol Ed). 2014 Jan 1;6:92-109.

Patterson, K., Nestor, P.J. & Rogers, T.T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, 8(12), 976-987.

Peretz, I., Kolinsky, R., Tramo, M., Labrecque, R., Hublet, C., Demeurisse, G. et al. (1994). Functional dissociations following bilateral lesions of auditory-cortex. *Brain*, 117, 1283–1301.

Pernet, C.R., McAleer, P., Latinus, M., Gorgolewski, K.J., Charest, I., Bestelmeyer, P.E. et al. (2015). The human voice areas: Spatial organization and inter-individual variability in temporal and extra-temporal cortices. *Neuroimage*, 119, 164-174.

Piccirilli, M., Sciarma, T. & Luzzi S. (2000). Modularity of music: evidence from a case of pure amusia. *Journal of Neurology, Neurosurgery and Psychiatry*, 69, 541-5.

Pitcher, D., Walsh, V. & Duchaine, B. (2011). The role of the occipital face area in the cortical face perception network. *Experimental Brain Research*, 209(4), 481-493.

Puce, A., Allison, T., Bentin S., Gore, J.C. & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *Journal of Neuroscience*, 18, 2188–2199.

Quaranta, D., Piccininni, C., Carlesimo, G.A., Luzzi, S., Marra, C., Papagno, C. et al. (2016). Recognition disorders for famous faces and voices: a review of the literature and normative data of a new test battery. *Neurological Sciences*, 37, 345-352.

Rice, G.E., Hoffman, P., Lambon Ralph, M.A. (2015). Graded specialization within and between the anterior temporal lobes. *Annals of the New York Academy of Sciences*, 1359, 84-97.

Rinaldi, M.C., Marangolo, P., Lauriola, M. (2006). BLED Santa Lucia (*Batteria sul Linguaggio dell'Emisfero Destro*). Organizzazioni Speciali, Firenze.

Sergent, J., Ohta, S. & MacDonald, B. (1992). Functional neuroanatomy of face and object processing. A positron emission tomography study. *Brain*, 115,15-36.

Snowden, J.S., Thompson, J.C., Neary, D. (2004). Knowledge of famous faces and names in semantic dementia. *Brain*, 127(4), 860-872.

Snowden, J.S., Thompson, J.C., Neary, D., 2012. Famous people knowledge and the right and left temporal lobes. *Behavioural Neurology*, 25, 35-44.

Spinnler, H. & Tognoni, G. (1987) Standardizzazione e Taratura Italiana di Test Neuropsicologici. *Italian Journal of Neurological Sciences*, Suppl 8, 1-120.

Van Lancker, D.R. & Canter, G.J. (1982). Impairment of voice and face recognition in patients with hemispheric damage. *Brain and Cognition*, 1(2), 185-95.

Van Lancker D.R., Cummings J.L., Kreiman J. & Dobkin B.H. (1988). Phonagnosia: A dissociation between familiar and unfamiliar voices. *Cortex*, 24, 195–209.

von Kriegstein, K., Eger, E., Kleinschmidt, A. & Giraud, A.L. (2003). Modulation of neural responses to speech by directing attention to voices or verbal content. *Brain Research Cognitive Brain Research*, 17(1), 48-55.

Von Kriegstein, K. & Giraud, A.L.(2004). Distinct functional substrates along the right superior temporal sulcus for the processing of voices. *Neuroimage*, 22(2), 948-55.

von Kriegstein, K., Kleinschmidt, A., Sterzer, P. & Giraud, A.L. (2005) Interaction of face and voice areas during speaker recognition. *Journal of Cognitive Neuroscience*, 17, 367–376

Warren, J.D., Scott, S.K., Price, C.J. & Griffiths, T.D. (2006). Human brain mechanisms for the early analysis of voices. *Neuroimage* 31, 1389-97

Warrington, E.K. & James, M. (1991). *The Visual Object and Space Perception Battery -VOSP-* Thames Calley Test Company, Bury St Edmunds (UK).

Yovel, G & Belin P. (2013). A unified coding strategy for processing faces and voices. *Trends in Cognitive Sciences*, 17, 263-271

Figure 1 caption

Panel A: Comparison between ^{18}F 2-deoxy-2-fluoro-D-glucose (FDG) uptake in the single case and the healthy controls ($n = 34$). The highlighted brain regions are those in which the patient shows a significantly lower FDG-uptake ($p < .05$ FEW corrected at voxel level). Panel B: MRI axial sections acquired using a FLAIR sequence. We selected all sections in which a lesion was visible. The showed sections are contiguous in the z axis.

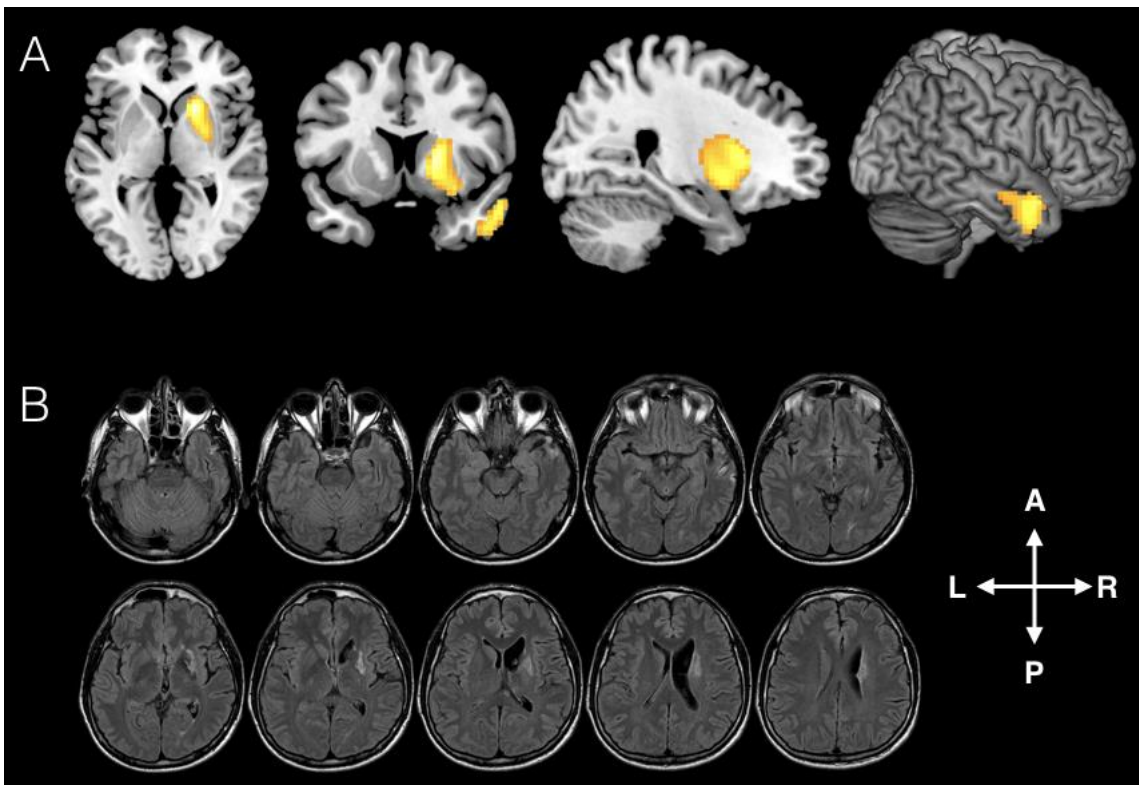


Table 1. General neuropsychology

Test (maximum score)	M's performance
<hr/> <i>General abilities</i>	
MMSE (30)	30
<i>Memory</i>	
Bi-syllabic word span	6
Corsi block span	6
Rey figure B delayed recall (31)	28
AVTL test long term recall (15)	9
<i>Executive functions and attention</i>	
Weigl's sorting test (15)	13
Luria's Motor Sequences (30)	30
Raven's Progressive Matrices (36)	33
Stroop Test I time (seconds)	20
Stroop Test I errors	0
Stroop Test II time (seconds)	50
Stroop Test II errors	0
Phonological fluency	37
<i>Perceptuo-spatial skills</i>	
VOSP	
- Shape detection test (20)	20
- Incomplete letters (20)	20
- Silhouettes (30)	24
- Object decision (20)	18
- Dot counting (10)	10
- Number location (10)	10

- Position discrimination (20)	20
- Cube analysis (10)	9

Praxis

Bucco-linguo-facial praxis (20)	20
Ideomotor praxis (20) right hand	20
Ideomotor praxis (20) left hand	20
Ideomotor praxis (10) right lower limb	20
Ideomotor praxis (10) left lower limb	20
Rey figure B copy (31)	31

Language

“Easy” picture naming (40)#	40
“Easy” word-picture matching (40)#	40
Reading (40)#	40
Writing (40)#	40
Repetition (40)#	40
Verbal fluency (3 categories 1 minute each)#	48
Pyramids and Palm tree test verbal version (30)#	30
Pyramids and Palm tree test visual version (30)#	30
Laiacona Test: naming (80)	79
Laiacona Test: comprehension (80)	80
Laiacona Test: comprehension within categories (80)	80
Questionnaire (480)	478
Linguistic prosody comprehension (10)	10

*= abnormal performance

Table 2. Examination of non musical sounds, musical abilities, faces, and voices perception and recognition

Test (maximum score)	M's performance
<i>Not musical sounds</i>	
Naming (70)	68
Sound-to-picture matching test(70)	70
Sound-to-name matching test (70)	70
Sound production (35)	35
<i>Prosodic emotional tone interpretation</i>	
Meaningful sentences (10)	10
Meaningless sentences (10)	10
<i>Musical Perception Tests</i>	
Pitch discrimination (20)	19
Tonal memory (10)	10
Chord Analysis (20)	18
Rhythmic memory (10)	10
<i>Recognition of musical characteristics</i>	
Musical instruments (10)	10
Intensity (10)	15
Ascending and descending scales (10)	10
Familiar rhythms (10)	10
<i>Recognition of familiar melodies</i>	
Recognition of familiar songs (20)	20
<i>Recognition of the singer voice</i>	
Recognition of the artist (20)	12**
Omissions	8
Recognition of another artist (20)	15**
False recognition	5
<i>Unfamiliar faces and voices discrimination</i>	
BFRT face discrimination (54)	52
Unknown voice discrimination (20)	18
<i>Famous voice and face recognition</i>	
(VO-REC)	
Familiarity score (40)	24**
Familiarity evaluation and person identification from voices (120)	56**
(FA-REC)	
Familiarity score (40)	40

Highlights

A selective impairment in famous voice recognition was found after right ATL stroke
No defect was found in voice perception, face perception and famous face recognition
PET and MRI revealed small lesions in right subcortical nuclei and temporal pole
From the temporal pole the lesion extended posteriorly in the middle temporal gyrus

.