

ORIGINAL ARTICLE

Voice Recognition in Face-Blind Patients

Ran R. Liu^{1,†}, Raika Pancaroglu^{1,†}, Charlotte S. Hills¹, Brad Duchaine²
and Jason J. S. Barton^{1,3}

¹Human Vision and Eye Movement Laboratory, Departments of Medicine (Neurology), Ophthalmology and Visual Sciences, University of British Columbia, Vancouver, BC, Canada, ²Department of Psychology, Dartmouth University, Hanover, NH, USA, and ³Neuro-ophthalmology Section K, VGH Eye Care Centre, Vancouver, BC, Canada V5Z 3N9

Address correspondence to Jason J. S. Barton, Neuro-ophthalmology Section K, VGH Eye Care Centre, 2550 Willow Street, Vancouver, BC, Canada V5Z 3N9.
Email: jasonbarton@shaw.ca

[†]Ran R. Liu and Raika Pancaroglu are co-first authors who contributed equally to the study.

Abstract

Right or bilateral anterior temporal damage can impair face recognition, but whether this is an associative variant of prosopagnosia or part of a multimodal disorder of person recognition is an unsettled question, with implications for cognitive and neuroanatomic models of person recognition. We assessed voice perception and short-term recognition of recently heard voices in 10 subjects with impaired face recognition acquired after cerebral lesions. All 4 subjects with apperceptive prosopagnosia due to lesions limited to fusiform cortex had intact voice discrimination and recognition. One subject with bilateral fusiform and anterior temporal lesions had a combined apperceptive prosopagnosia and apperceptive phonagnosia, the first such described case. Deficits indicating a multimodal syndrome of person recognition were found only in 2 subjects with bilateral anterior temporal lesions. All 3 subjects with right anterior temporal lesions had normal voice perception and recognition, 2 of whom performed normally on perceptual discrimination of faces. This confirms that such lesions can cause a modality-specific associative prosopagnosia.

Key words: familiarity, multimodal, prosopagnosia, semantic, voice perception

Introduction

Subjects with prosopagnosia do not realize that faces they have seen before are familiar. Prosopagnosia is a family of conditions, rather than a single entity. It can be caused by a variety of occipital and temporal lesions (Barton 2008) and, though there are dissenting opinions (Busigny et al. 2014), it has long been proposed that there are functional variants (Hécaen 1981; Damasio et al. 1990; de Renzi et al. 1991). These proposals reflect cognitive models in which face recognition proceeds through a hierarchy of operations, progressing from coding of facial structure, matching of the structural code to stored facial memories, and access to semantic information and names (Bruce and Young 1986). In apperceptive prosopagnosia, there is a failure of structural coding of

faces, whereas in the associative/amnesic form the defect lies in matching the percept to facial memories (Damasio et al. 1990; de Renzi et al. 1991; Davies-Thompson et al. 2014).

Structure/function correlations suggest that the apperceptive variant is associated with occipital and fusiform lesions and the associative variant with anterior temporal lobe damage (Barton 2008). However, anterior temporal damage has also been associated with a multimodal person recognition disorder that also affects other means of recognizing people, such as names and voices (Ellis et al. 1989; Hanley et al. 1989; Gainotti et al. 2008; Busigny et al. 2009). Although previous studies have shown that prosopagnosic subjects can recognize names (Barton et al. 2001), it has recently been argued that, since name recognition is a left hemisphere function while voice and face recognition may both

lateralize to the right hemisphere (Gainotti 2013a, 2013b), voice processing needs to be examined to exclude a multimodal problem in subjects with right anterior temporal lobe damage (Gainotti 2010, 2013a, 2013b). Whether voice processing and face processing converge is a theoretically important issue for cognitive models of people recognition, and has rarely been addressed in prior patient studies (Gainotti and Marra 2011).

The hierarchy of face processing models is reflected in a logical sequence of testing, from stimulus coding, to recognizing a stimulus as familiar, to identifying it by name or semantic information. (To avoid ambiguity, we will use the term 'recognition' as equivalent to familiarity, with the term 'identification' referring to the ability to name or provide semantic data about the stimulus.) Coding of facial structure is tested by face discrimination, by having subjects match unfamiliar faces (Benton and van Allen 1972), or detects changes in facial structure (Barton et al. 2002; Barton 2008; Ramon et al. 2010). Face recognition can be probed by asking subjects to indicate which of a set of famous and anonymous faces are familiar (Albert et al. 1979; Barton et al. 2001), but this depends on premorbid acquaintance with celebrities and may confound familiarity for the face with familiarity for the person (Haslam et al. 2001). Consequently, though representations of newly encountered faces may not be as rich and robust as those of famous or personally familiar faces (Burton et al. 2011; Natu and O'Toole 2011), assessments of familiarity for recently viewed faces have become the diagnostic standard (Warrington 1984; Duchaine and Nakayama 2006). The status of facial memories can also be probed by tests of facial imagery, which asks subjects questions that they must answer by recalling the image of faces they have seen in the past (Barton and Cherkasova 2003; Davies-Thompson et al. 2014).

To identify parallel apperceptive and associative defects in voice perception, we developed 2 tests: first, a match-to-sample test for voice discrimination to detect apperceptive defects in voice processing and second, a test of short-term memory for recently heard voices. We applied these tests to a series of subjects with impaired face recognition from various lesions. We asked whether these subjects would have a pattern of impaired voice processing that mirrored their deficits in face processing. Specifically, in the case of patients with right anterior temporal lesions, the possibility of a multimodal recognition disorder would be supported by parallel findings of impaired voice and face recognition, despite normal voice and face discrimination.

Methods

Participants

All subjects had normal or corrected-to-normal vision, and provided informed consent to a protocol approved by the University of British Columbia and Vancouver General Hospital Ethics Review Board. We recruited 73 healthy control subjects, who were compensated for their participation. No control subject had a history of neurological or psychiatric diseases, or visual or auditory complaints. Ages ranged from 19 to 70 years [mean 33.6, standard deviation (SD) 15.5]. To match our prosopagnosic cohort, we required subjects to be born in North America, to have English as their first language, and to have lived in North America for 5 years or more. All 73 subjects performed the voice discrimination test, and 54 subjects (mean 37.2 years, SdD 16.4, range 19–70) also performed the voice recognition test, which was developed after the voice discrimination test.

Ten subjects diagnosed with complaints of impaired face recognition after brain lesions participated, with an age range of 23–61 years (Table 1). These subjects were recruited as part of an ongoing study of prosopagnosia through the www.faceblind.org website, and all had extensive neuropsychological testing of vision and memory (Table 2).

Neuroimaging

Structural Imaging

Patients were scanned in a Philips 3.0-T scanner at the UBC MRI Research Centre. A high-resolution T_1 -weighted anatomical image and a T_2 -weighted FLAIR image were collected from each patient (Fig. 1). Four patients had inferior occipitotemporal lobe lesions (designated as IOT), 5 had anterior temporal lobe lesions (AT), and 1 had right anterior temporal and bilateral occipitotemporal lesions (ATOT). The nomenclature for these subjects follows the evidence for tissue loss or hypointensity on T_1 -weighted images. Some had additional complexities to their lesions. B-ATOT2 had bilateral fusiform lesions and a right anterior temporal lesion, as well as posterior periventricular hyperintensity on FLAIR sequences. L-IOT2, who had resection of the left fusiform gyrus for epilepsy, also had atrophy of the right fusiform gyrus and failed to show activation of the right fusiform face area. FLAIR sequences in R-AT3 revealed left medial temporal lobe hyperintensity. Finally, R-AT5 also suffered unilateral right deafness from radiation treatment.

Table 1 Patient demographics

Subject	Current age	Onset age	Etiology	Lesion(s)
R-IOT1 ^{a,b,c}	54	37	Vascular malformation	Right fusiform
R-IOT4 ^{b,c}	62	57	Stroke	Right fusiform
B-IOT2 ^c	60	27	Hemorrhage	Bilateral fusiform
L-IOT2 ^c	59	39	Resection, epilepsy	Left fusiform resection, right fusiform atrophy
B-ATOT2	23	10	Herpes encephalitis	Bilateral fusiform, right anterior temporal
R-AT2 ^{b,c}	34	26	Herpes encephalitis	Right anterior temporal
R-AT3 ^c	37	30	Herpes encephalitis	Right anterior temporal
R-AT5	60	32	Tumor resection	Right anterior temporal
B-AT1 ^{a,b}	25	21	Herpes encephalitis	Bilateral anterior temporal
B-AT2 ^{c,d}	47	24	Trauma	Bilateral anterior temporal

^aReported in Fox et al. (2011, 2013).

^bReported in Barton et al. (2009) and Dalrymple et al. (2011).

^cReported in Davies-Thompson et al. (2014).

^dReported as subject 1 in Barton et al. (2002), subject 8 in Barton et al. (2001); Barton and Cherkasova (2003); Iaria et al. (2008), and TS in Barton et al. (2003).

Table 2. Neuropsychological test results.

Test	Max	R-IOT1	R-IOT4	B-IOT2	L-IOT2	B-ATOT2	R-AT2	R-AT3	R-AT5	B-AT1	B-AT2
Attention											
Trails A	-	39	48 [#]	<u>80</u>	54 [#]	30	21	22	43	18	30
Trails B	-	61	102 [#]	<u>142</u>	117 [#]	93 [#]	44	37	78	25	40
Star Cancellation	54	54	54	53	53	54	54	54	54	54	54
Visual Search	60	54	n/a	56	60	59	59	59	52	59	56
Memory											
Digit span-forward	16	12	8	14	10	<u>7</u>	13	16	10	12	9
Spatial span-forward	16	9	10	8	10	8 [#]	9	12	<u>6</u>	10	9
Word list	48	28	37	35	27	<u>27</u>	35	31	24	<u>27</u>	23 [#]
Visuo-perceptual											
Hooper Visual Organization	30	27	22	22.5	<u>9</u>	<u>12</u>	28	27.5	22	20	28
Benton Judgment of Line Orientation	30	29	24	29	23	22	28	30	21	28	28
Visual Object and Spatial Perception											
Object: Screening	20	20	18	20	20	20	20	20	17	20	20
Incomplete Letters	20	19	19	19	17	19	20	19	20	19	19
Silhouettes	30	21	18	<u>12</u>	<u>3</u>	<u>4.5</u>	18	22	19	<u>10</u>	25
Object Decision	20	16	19	14	<u>13</u>	<u>10</u>	20	17	14	16	18
Progressive Silhouettes	20	9	13	15	10	4	10	11	<u>17</u>	<u>17</u>	8
Spatial: Dot Counting	10	10	9	10	10	9	10	10	10	10	10
Position Discrimination	20	20	19	19	19	<u>15</u>	20	19	18	19	20
Number Location	10	10	10	10	10	8	9	10	10	10	10
Cube Analysis	10	10	10	10	10	9	10	10	8	10	9
Imagery											
Mental Rotation	10	10	10	10	7	10	9	10	10	10	<u>5</u>

Note: Underlining denotes impaired values.

[#]Denotes borderline performance.

Volumetric Analysis of Lesions

Lesions were mapped as regions of interest from T₁-anatomical images, and their volumes were measured using MRICron. The anterior tip of the middle fusiform sulcus (Weiner et al. 2014), which roughly falls at the midpoint between the anterior temporal and occipital poles, was used as the demarcation line (Talairach $y = -30$). Regions anterior to the anterior tip of the middle fusiform sulcus were designated as anterior temporal and regions posterior were designated as occipitotemporal. The superior borders of occipitotemporal cortex are not precise: Because face-selective areas such as the anterior inferior temporal region, fusiform face area, and occipital face area are located in inferior cortex, we restricted analysis to regions in slices below that containing the first appearance of the posterior horn of the lateral ventricles (Table 3).

Functional Neuroimaging

All subjects had functional magnetic resonance imaging to localize the core components of the face processing network, using the HVEM dynamic face localizer protocol (Fox et al. 2009). T₂*-weighted functional scans were used to collect data from 36 interleaved axial slices (time repetition = 2000 ms, time echo = 30 ms, field of view = 240 × 216 mm, 3 mm thickness with 1 mm gap, voxel size 3 × 3 mm, 128 reconstruction matrix, reconstructed voxel size = 1.875 × 1.65 mm). The functional slices were coregistered onto a T₁-weighed anatomical image for each patient.

The HVEM dynamic face localizer scan consisted of grayscale video clips of faces and objects. Each stimulus block included 6 video clips lasting 1.5 s each, separated by a 500-ms blank screen. Stimulus blocks were separated by a 12-s fixation block. Each condition (faces or objects) was repeated 8 times per run. Attention was sustained by asking the patients to press a button on an MRI-compatible button-box when the same video was presented

twice in a row. Functional data were analyzed using the Brain-Voyager QX software. Preprocessing steps included slice time correction (cubic spline interpolation), 3D motion correction (trilinear/sinc interpolation), and high-pass temporal filtering (GLM-Fourier, 2 sines/cosines). Face-selective regions were determined for each patient individually with the contrast Faces > Objects at $P < 0.05$. Subjects with an OT designation did not show activation by faces of the right fusiform face area, whereas those with an AT designation alone showed activation of all core areas, namely the fusiform face area, occipital face area, and superior temporal sulcus (STS), with the exception of R-AT5, who did not show activation of the right STS (Fig. 2 and Table 4).

Evaluation of Face and Name Processing

The accuracy of perception or structural coding of the face was evaluated with 2 discriminative tests involving anonymous faces. These were the Benton Face Recognition Test (Benton and van Allen 1972) and a test of the perception of the spatial relationships and features in faces (Barton et al. 2002). In the latter, perception of interocular distance has been shown to be a reliable marker of impaired structural coding of the face in subjects with occipitotemporal lesions (Barton 2008), perhaps reflecting a particular vulnerability of the eye region in apperceptive prosopagnosia (Caldara et al. 2005). Thus, we used the interocular accuracy score as the index for perception of facial configuration, though on any given trial in the test subjects were not aware whether the altered aspect of the face was feature color or spacing, or if it was located in the eye or mouth region.

Face recognition was evaluated with 2 standard tests of short-term familiarity with anonymous faces, the Warrington Recognition Memory Test (Warrington 1984) and the Cambridge Face Memory Test (Duchaine and Nakayama 2006), and a Famous Faces Test (Barton et al. 2001).

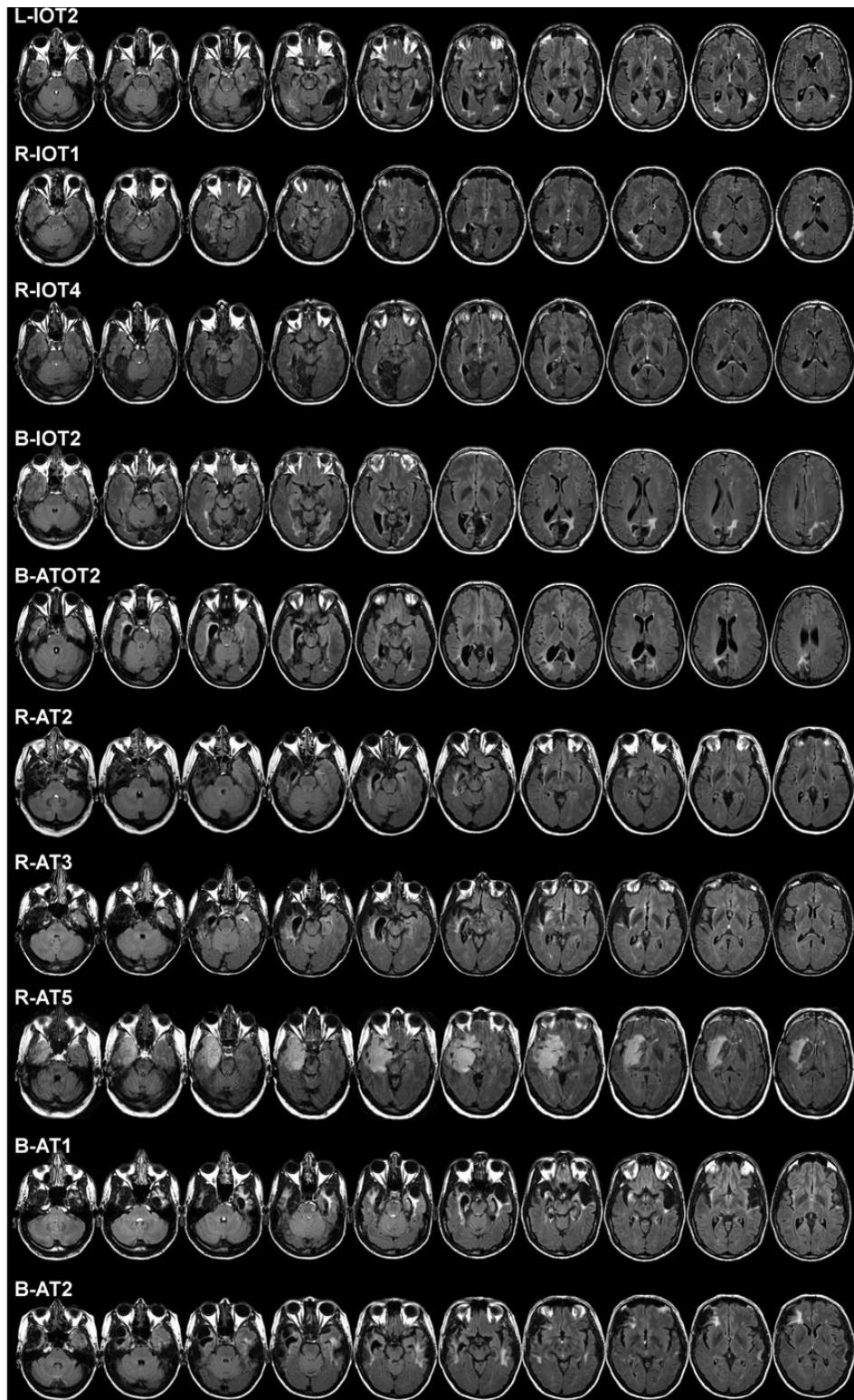


Figure 1. Lesions of subjects. Axial FLAIR MR images are shown.

The status of facial memories was probed with an imagery test that presented subjects with the names of 2 celebrities and asked them to make a judgment about their facial appearance, such as which one had the rounder face (Barton and Cherkasova 2003).

Familiarity with names was evaluated by a test that asked subjects to indicate which of 2 names were familiar, one belonging to a celebrity and the other not (Barton et al. 2001). The ability to access semantic knowledge from names was evaluated by

Table 3 Volumetric analysis of lesions

Patient	Right hemisphere				Left hemisphere			
	Anterior temporal		Inferior occipitotemporal		Anterior temporal		Inferior occipitotemporal	
	Lesion size (cm ³)	% loss	Lesion size (cm ³)	% loss	Lesion size (cm ³)	% loss	Lesion size (cm ³)	% loss
L-IOT2							12.4	30.2
R-IOT1			11.0	24.8				
R-IOT4	3.1	3.8	20.3	45.5				
B-IOT2			12.4	30.3	22.0	54.7		
B-ATOT2	6.1	14.4	7.7	37.7			4.1	15.8
R-AT2	29.0	59.2						
R-AT3	44.5	84.8						
R-AT5	53.1	70.1						
B-AT1	25.5	47.4			22.5	45.9		
B-AT2	27.3	48.7	3.3	10.3	10.4	21.8	3.4	8.0

having subjects sort names by occupation, namely acting or politics (Barton et al. 2001).

Apparatus for Testing Voice Processing

All tests were presented on an IBM Lenovo laptop with 1280 × 800 pixels resolution, using Superlab (www.superlab.com) software. Participants wore a Panasonic RP-HTX7 headset, which provides adequate sound insulation. Subjects sat 57 cm away from the display screen, in a dimly lit and quiet room, and wore the headset throughout the entire test.

Stimuli

Auditory stimuli were created from volunteers between the ages of 20–31. For the discrimination test, a set of stimuli were generated from 20 male and 20 female volunteers. For the voice recognition test, another set of stimuli were generated from 21 male and 21 female volunteers. For all tests, each stimulus was used no more than once as a target or as a distractor.

Voice Discrimination

This used a match-to-sample strategy. Audio stimuli for each individual consisted of 2 different texts that the volunteers read. For the initial sample, the subjects read the phrase: “This is by far one of the most amazing books I have ever read, it tells the story of a Colombian family across generations.” For the target and distractor stimuli that followed, both voices read the phrase: “After a hearty breakfast, we decided to go for a walk on the beach. It was a lovely morning with the crisp smell of the ocean in the air.” Volunteers were asked to speak both texts at the same speed. All recordings were 10 s in duration.

Voice Recognition

Two sets of audio stimuli were made, with each volunteer contributing 2 samples to both sets, one of which could be used in a learning phase and the other in a recognition phase. The “question component” was recorded in interview style, where all individuals were asked to read silently 2 questions and voice their personal responses to these questions. For the learning phase of the question component, we asked, “What was your favorite childhood activity?” For the testing phase, we asked, “What was your favorite vacation?”

The “passages component” was recorded in narrative style, so that for the learning phase all individuals read a passage chosen

at random from the short story collection “Too Much Happiness” and for the testing phase another randomly chosen passage from the short story collection “Friend of My Youth,” both by Alice Munro. All individuals read different passages. For both the question and passages components, stimulus duration for audio clips was set at 12.5 s.

Protocol

Voice Discrimination

Each trial began with a screen that read “Target Voice.” Simultaneously, the subject heard the sample voice for 10 s. After a 1.5-s pause, there was a ring tone lasting 875 ms, which served as an auditory mask and to separate the sample from the match choices. Next, a screen that displayed “Choice 1,” at which time the subject heard simultaneously the first of the 2 choice voices. After this was completed, the screen displayed “Choice 2,” and the subject heard the second choice voice at the same time. One of the 2 choice voices was the match voice, from the same person as the sample voice, and the other was a distractor voice of the same gender. Matches and distractors were given in a random order. Both sample and choice voices began with 1 s of silence, ran for 10 s of the voice, and ended with 1.5 s of silence. Following the 2 choice voices, a screen prompted the subjects to indicate which of the 2 choice voices matched the target voice, by a key press. There were a total of 40 trials divided into 2 blocks, one block of 20 trials with male voices and one block of 20 trials with female voices.

Voice Recognition

The voice recognition test had learning and testing phases. The test was designed so that subjects had to retain memory of a voice over a short interval and with intervening interference from stimuli, in a manner similar to standard tests of face familiarity. However, slight differences in design are unavoidable. First, voices need to be presented sequentially: Hence, choice displays cannot show multiple items simultaneously, as is usually done with faces. Second, because of their temporal dynamics, voices require more time for stimulus presentation. Third, testing difficulty needs to be adjusted because humans are poorer at voice than face recognition (Latinus and Belin 2011), probably reflecting the fact that the visual modality is the main source of information for identifying biological objects (Gainotti et al. 2013). Even when exposure is carefully matched, face identification remains superior to voice identification (Barsics and Bredart 2012).

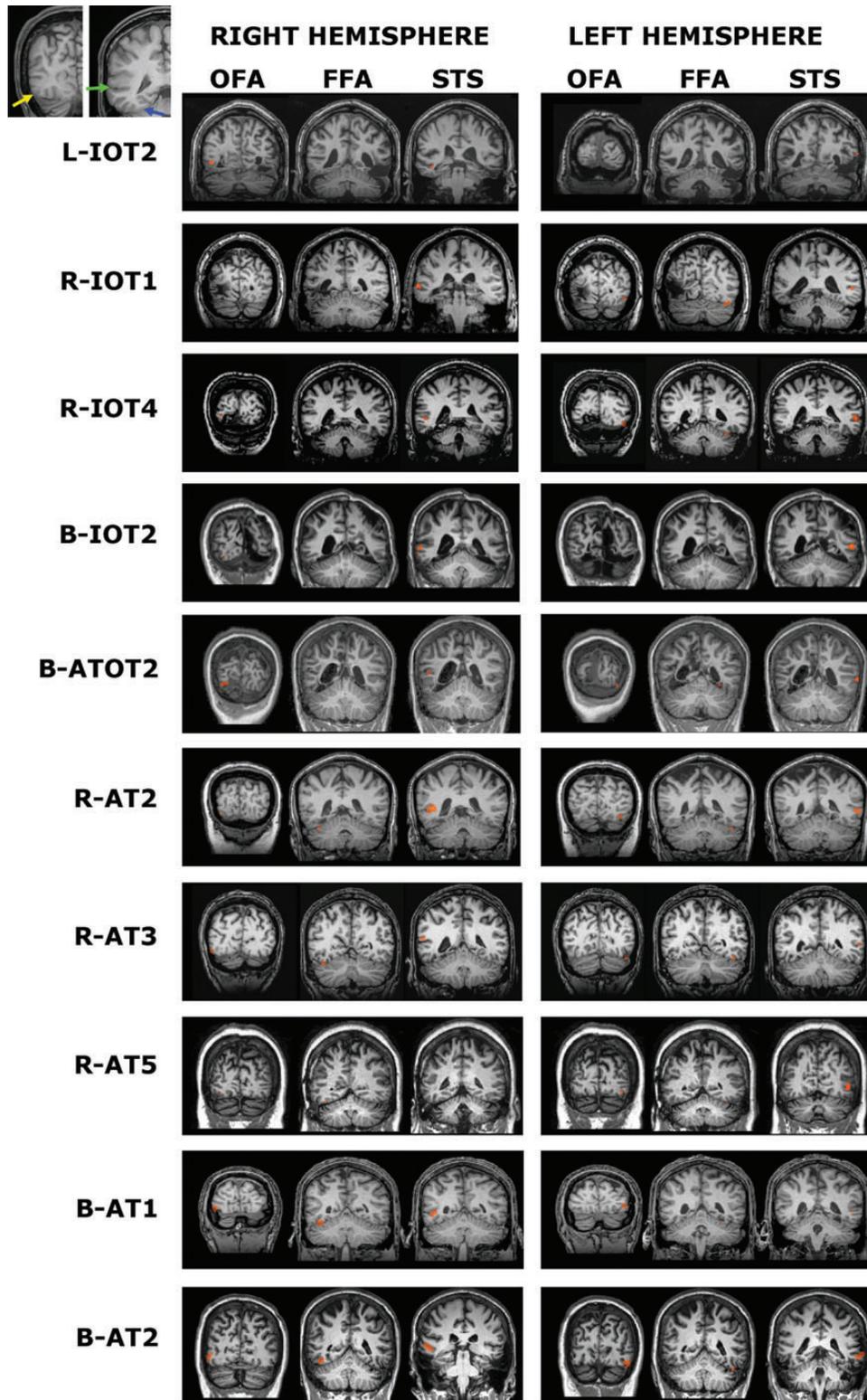


Figure 2. Results of functional MRI in subjects. Activations of the 6 core regions of the face-processing network are depicted in orange and overlaid on coronal T_1 -weighted images. These include the occipital face area (OFA), fusiform face area (FFA), and superior temporal sulcus (STS), in the left and right hemispheres. In subjects in whom no activation was found in a given region, a representative slice at the approximate expected location is given to show the lesion. Inset in the top right corner shows in a representative intact right hemisphere anatomical landmarks, namely the inferior occipital gyrus (yellow arrow), the STS (green arrow), and the fusiform gyrus (blue arrow).

The inferiority of voice perception is evident in the control data of some reports that have studied both face and voice processing in subjects (Hailstone et al. 2010; Hoover et al. 2010). In our pilot

work, we found that healthy subjects performed close to chance when asked to recall 5 unfamiliar voices sequentially. Also, while the mechanisms for short- and long-term voice familiarity are

Table 4 Functional MRI results

Patient	Region	Peak t-value	Cluster size	Coordinates			Patient	Region	Peak t-value	Cluster size	Coordinates		
				X	Y	Z					X	Y	Z
L-IOT2	rOFA	6.66	146	43	-65	-5	R-AT2	rOFA	3.48	51	29	-86	-9
	rFFA	-						rFFA	8.34	626	38	-41	-22
	rpSTS	4.54	48	40	-32	-6		rpSTS	12.59	1825	43	-38	3
	lOFA	3.43	5	-28	-92	-15		lOFA	8.10	184	-32	-79	-11
	lFFA	Lesion						lFFA	6.02	97	-43	-43	-22
R-IOT1	lpSTS	3.72	10	-57	-56	-8	R-AT3	lpSTS	7.79	343	-58	-44	4
	rOFA	Lesion						rOFA	8.76	187	45	-78	-11
	rFFA	Lesion						rFFA	8.11	289	40	-55	-20
	rpSTS	5.52	146	57	-40	13		rpSTS	5.74	462	59	-44	12
	lOFA	4.98	51	-36	-79	-14		lOFA	3.78	84	-39	-72	-19
R-IOT4	lFFA	6.71	281	-33	-68	-23	R-AT5	lFFA	12.82	491	-41	-54	-17
	lpSTS	6.32	785	-57	-28	-2		lpSTS	3.42	5	-58	-47	8
	rOFA	4.54	88	27	-89	-5		rOFA	4.63	7	26	-74	-15
	rFFA	Lesion						rFFA	4.13	7	35	-50	-18
	rpSTS	4.34	194	51	-36	3		rpSTS	-				
B-IOT2	lOFA	9.01	786	-32	-83	-13	B-AT1	lOFA	5.88	49	-34	-76	-15
	lFFA	7.34	169	-33	-43	-20		lFFA	4.04	5	-34	-47	-19
	lpSTS	7.47	414	-57	-38	2		lpSTS	5.2	259	-44	-61	-5
	rOFA	5.45	45	26	-81	-14		rOFA	12.37	3956	30	-88	-5
	rFFA	Lesion						rFFA	13.09	1064	39	-52	-20
B-ATOT2	rpSTS	10.37	966	58	-42	1	B-AT2	rpSTS	9.67	329	46	-49	-2
	lOFA	Lesion						lOFA	9.43	1543	-30	-85	-8
	lFFA	Lesion						lFFA	5.96	57	-39	-55	-26
	lpSTS	9.94	731	-50	-51	1		lpSTS	5.90	50	-60	-46	4
	rOFA	4.57	439	25	-88	-14		rOFA	6.77	359	43	-68	-13
B-ATOT2	rFFA	Lesion					B-AT2	rFFA	12.76	679	39	-46	-19
	rpSTS	3.82	93	47	-45	8		rpSTS	11.64	1464	50	-29	3
	lOFA	4.94	132	-29	-92	-23		lOFA	8.32	738	-42	-76	-26
	lFFA	4.25	123	-29	-55	-10		lFFA	5.81	175	-44	-46	-30
	lpSTS	7.08	252	-59	-51	2		lpSTS	4.54	396	-55	-43	-7

not completely identical, healthy subjects perform poorly on tests with famous voices. On one extensive test of famous voice recognition, controls could only name on average about 50% of 79 voices (Meudell et al. 1980), and on another only 19% of 96 voices (Garrido et al. 2009), which contrasts with a mean hit rate of 85% on one test of famous faces (Barton 2008), for example.

Because of the foregoing, the test was divided into sets of 3 trials. Each set began with a screen that displayed "Learning Phase" for 3 s. During the learning phase, the subject heard 3 target voices. With the first target voice, the screen simultaneously displayed "Voice A," during the second target voice, the screen read "Voice B," and during the third target voice, the screen read "Voice C." The audio clip of each voice began with 200 ms of silence followed by 12.5 s of the person speaking. This was then followed by 700 ms of silence and finally a ring tone lasting 875 ms. After completion of this learning phase, which lasted about 45 s, the screen displayed "Testing Phase" for 3 s. Participants then heard 3 pairs of choice voices. In the first pair, one of the 2 choice voices was the same as "Voice A" in the learning phase, while the other was a distractor of the same gender. The temporal order of matches and distractors was randomized. While the first choice voice was playing, the screen displayed "Voice A Choice 1," and while the second choice voice was playing it displayed "Voice A Choice 2." Each of these choice clips lasted 12.5 s, and, like the target voice, was preceded by 200 ms of silence and followed by 700 ms of silence, and then a ring tone lasting 875 ms. A screen then prompted the subjects to indicate

by a key press which of the 2 choice voices matched the Voice A they had heard in the learning phase. This was then followed by the choices for Voice B, and finally by the choices for Voice C. Each set of 3 trials had at least one trial for each gender. Subjects completed 7 sets (21 trials) with stimuli from the "question component," and then 7 sets (21 trials) with stimuli from the "passages component," to give a total of 42 trials.

All subjects completed the voice discrimination test first, with subjects randomized to start with either male or female stimuli, and the option for a short break between the two. For the 49 control subjects who also performed the voice recognition test, there was a break of at least 10 min before they did this second test. The voice recognition test presented the "question component" first and the "passages component" second.

Data Collection and Analysis

For each subject, the accuracy score (percent correct) in each test was calculated. For healthy controls we compared the effects of gender by t-tests for each of the 2 tests. For age, we performed a linear regression of scores against age, after our inspection of the data did not reveal any inflection that suggested a better fit from a non-linear function. To evaluate subjects, we performed an age-adjusted analysis by regressing out the variance due to age, and used the residual variance in the function to calculate the 95% prediction intervals appropriate for single-subject comparisons for the regression against age.

Questionnaire About Face and Voice Identification

Many prosopagnosic subjects state that they rely on voice to identify people, but objective tests do not always corroborate this (Boudouresques et al. 1979). To determine how our subjects' experience related to our testing, we administered a questionnaire about face and voice identification in daily life. There were 5 questions about identifying people by face and 5 similar questions for identification by voice (Supplementary Material, Appendix), with subjects asked to indicate responses on a 7-point Likert scale. Forty-eight control subjects (mean age 37.9 years, SD 15.7, range 19–70) and all prosopagnosic subjects completed the questionnaire. Scores from the 5 face questions and from the 5 voice questions were summed separately to give a face and a voice score out of 35, with a higher score indicating more difficulty.

Additional Testing

Because B-ATOT2 had problems discriminating voices, we performed more testing to exclude a primary auditory problem or a general auditory agnosia. First, she had standard clinical pure tone audiometry. Second, we created an on-line sound recognition test that presented 27 audio clips of a variety of animal, object, environmental, or human non-verbal sounds, taken from the internet, ranging in duration from 1 to 9 s (<http://www.neuroophthalmology.ca/UBCNeuroOp/JBarton/soundquizHVEM/soundintroduction.html>). Subjects had to write down what they had heard.

Results

Prosopagnosic Subjects, Face, and Name Processing

All subjects were impaired on the Famous Faces Test (Table 5), and on either the faces component of the Warrington Recognition Memory Test or the Cambridge Face Memory Test, most on both. In contrast, all subjects performed normally on the word component of the Warrington Recognition Memory Test.

Assessments of structural coding of faces showed that discrimination of facial configuration cosegregated most reliably with lesion site, being impaired in all subjects with right fusiform damage, and spared in all but one subject with anterior temporal

damage alone (R-AT5, whose damage extended to the anterior fusiform cortex). The Benton Face Recognition Test was less sensitive, with performance impaired in only 2 subjects. This is consistent with concerns expressed by others about the adequacy of the Benton Face Recognition Test as a probe of face perception (Farah 1990; Gainotti 2010).

All subjects with anterior temporal lesions were impaired on the test of facial imagery. B-AT1 could not do the test because he did not recognize enough of the celebrity names. All subjects with inferior occipitotemporal lesions alone had normal scores, with the exception of L-IOT2.

All subjects except B-AT1 had high familiarity with celebrity names and could sort them by occupation, thus demonstrating intact ability to access semantic information from names.

In summary, these data indicate an apperceptive prosopagnosia in all subjects with occipitotemporal lesions, and in R-AT5. The remaining subjects with anterior temporal lesions have intact face discrimination but impaired face familiarity, consistent with an associative prosopagnosia, which is also supported by their impaired imagery for faces. Of these, only B-AT1 has impaired familiarity with names, raising the possibility of a multimodal problem with person recognition even before testing of voice perception.

Control Subjects and Voice Processing

There was no effect of gender in the voice discrimination test (male mean score = 34.5, SD 2.81, female = 34.9, s.d. 3.57, $t(71) = 0.57$, $p = 0.57$) or the voice recognition test (male mean score = 33.5, s.d. 4.05, female = 33.1, s.d. 3.95, $t(42) = 0.30$, $p = 0.76$). Performance declined with age on voice discrimination ($r = 0.32$, slope = -0.07 , $F(1,72) = 8.15$, $p < 0.006$) and voice recognition ($r = 0.33$, slope = -0.08 , $F(1,53) = 6.33$, $p < 0.015$, Figure 3). Hence we collapsed control data across gender and derived age-adjusted 95% prediction intervals.

Prosopagnosic Subjects and Voice Processing

On voice discrimination, only 2 subjects were impaired, B-ATOT2 and R-AT5, who had right-sided deafness as a complication of radiation therapy (Fig. 3, top).

On voice recognition (Fig. 3, bottom), 3 subjects were impaired, B-ATOT2, B-AT1, and B-AT2. Because B-ATOT2 was impaired in both voice discrimination and recognition, her

Table 5. Results on tests of face and name processing.

	R-IOT1	R-IOT4	B-IOT-2	L-IOT2	B-ATOT2	R-AT2	R-AT3	R-AT5	B-AT1	B-AT2
FACES										
Structural coding										
Eye configuration	<u>0.39</u>	<u>0.44</u>	<u>0.28</u>	<u>0.28</u>	<u>0.28</u>	1.00	1.00	<u>0.56</u>	1.00	0.94
BFRT	45	46	38	<u>31</u>	37	47	38	<u>33</u>	45	40
Familiarity										
Famous faces d'	<u>1.96</u>	<u>1.29</u>	<u>1.31</u>	<u>0.00</u>	<u>-0.15</u>	<u>0.65</u>	<u>0.90</u>	<u>1.52</u>	<u>-0.36</u>	<u>0.68</u>
WRMT face	<u>33</u>	39	<u>21</u>	<u>27</u>	<u>19</u>	<u>27</u>	<u>31</u>	<u>28</u>	<u>27</u>	<u>31</u>
WRMT word	41	50	42	42	39	47	47	46	45	46
CFMT	44	<u>27</u>	<u>24</u>	<u>21</u>	<u>24</u>	<u>33</u>	<u>31</u>	<u>35</u>	<u>30</u>	<u>31</u>
Face imagery	0.82	0.84	0.86	<u>0.41</u>	<u>0.48</u>	<u>0.73</u>	<u>0.49</u>	<u>0.81</u>	*	<u>0.50</u>
NAMES										
familiarity	1.00	0.95	1.00	0.95	0.90	0.95	1.00	0.95	<u>0.65</u>	1.00
occupation sorting	1.00	0.98	1.00	0.88	0.73	1.00	0.98	1.00	<u>0.54</u>	1.00

Note: Underlining indicates an abnormal result.

BFRT: Benton face recognition test; WRMT: Warrington recognition memory test; CFMT: Cambridge face memory test.

*B-AT1 did not recognize enough celebrity names to perform the imagery test.

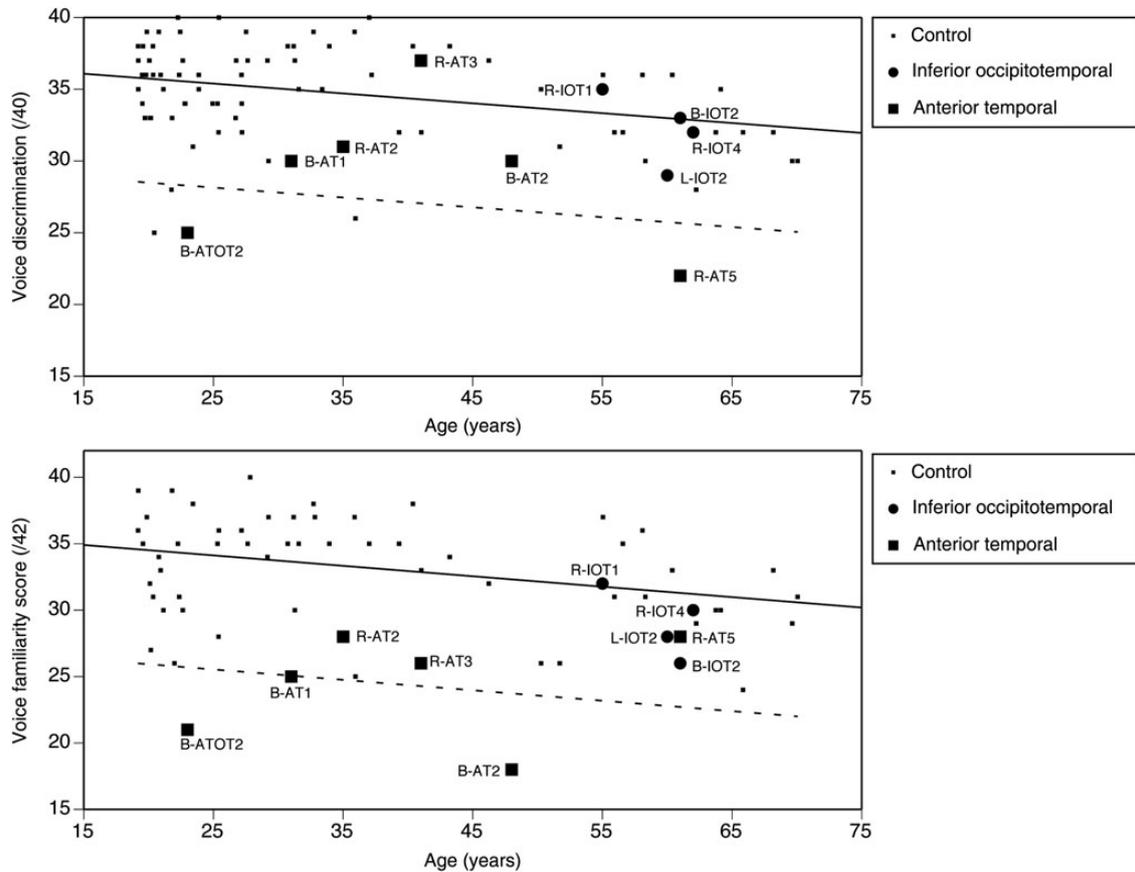


Figure 3. Voice scores plotted as a function of age. Top graph shows voice discrimination, and bottom graph shows voice recognition. Control subjects (small black discs) in both tests show a significant decline with age. Solid line shows the linear regression and the dotted line shows the age-adjusted lower 95% prediction limit. Subjects falling below the dotted line are impaired. Subject B-ATOT2 is impaired on both voice discrimination and recognition. Only B-AT1 and B-AT2 are impaired on recognition, but normal on voice discrimination. L: left; R: right; IOT: inferior occipitotemporal lesion; AT: anterior temporal lesion; ATOT: anterior temporal and bilateral occipitotemporal lobe lesions.

difficulties are consistent with an aperceptive phonagnosia. B-AT1 and B-AT2 have the pattern of preserved voice and face discrimination and impaired voice and face recognition that could point to a multimodal person recognition disorder. None of the 3 subjects with unilateral right anterior temporal lesions had impaired voice recognition. (Interestingly, R-AT5, who had right hearing loss and an impaired score for voice discrimination, performed normally on voice recognition. We can only speculate that some of the auditory cues that aid immediate discrimination may not contribute to the encoding of short-term voice representations, and that perception of those cues are more impaired by her hearing loss.)

We also compared the severity of the voice recognition deficit with scores on 2 standard tests of short-term memory for faces. We converted all accuracy scores into age-adjusted z-scores, and plotted voice recognition against face recognition (Fig. 4). Only subject B-AT2 had a deficit in voice recognition that was of similar severity to her difficulty with face recognition.

Questionnaire

Control subjects showed no effect of age or gender. They reported better identification by face than by voice ($t_{(42)} = 6.31, P < 0.0001$). Compared with 95% prediction limits, all prosopagnosic subjects indicated difficulty with face identification (Table 6), and only B-AT1 reported difficulty with voice identification.

Additional Tests for B-ATOT2

B-ATOT2 had normal hearing thresholds of 0–10 dB for all frequencies between 250 and 8000 Hz, in either ear. On a 27-item test of recognition of sounds made by animals, objects, or environments, 7 control subjects (6 female, age 23–36 years) obtained a mean score of 92% correct (SD 6.5%), whereas B-ATOT2 scored 85% correct, well above the lower 95% prediction limit of 75%. Inspection of her MRI (Fig. 1) shows lesions of right and left fusiform gyri, right lateral temporo-occipital cortex, and right medial occipitoparietal cortex, damage to the medial aspect of the temporal pole and inferior temporal cortex, and hyperintensity of the white matter around the posterior horns of both lateral ventricles and the periventricular white matter underlying the lateral temporal cortex.

Discussion

Our study produced 3 main findings. First, subjects with occipitotemporal lesions causing loss of the right fusiform face area have an aperceptive prosopagnosia, consistent with prior work (Damasio et al. 1990; Barton 2008; Fox et al. 2011). These subjects have both intact name familiarity and semantic access from names, and intact voice discrimination and recognition (Fig. 5): Hence, their problem is modality-specific. Investigating voice processing in these subjects is important because functional neuroimaging studies show that familiar voices are associated

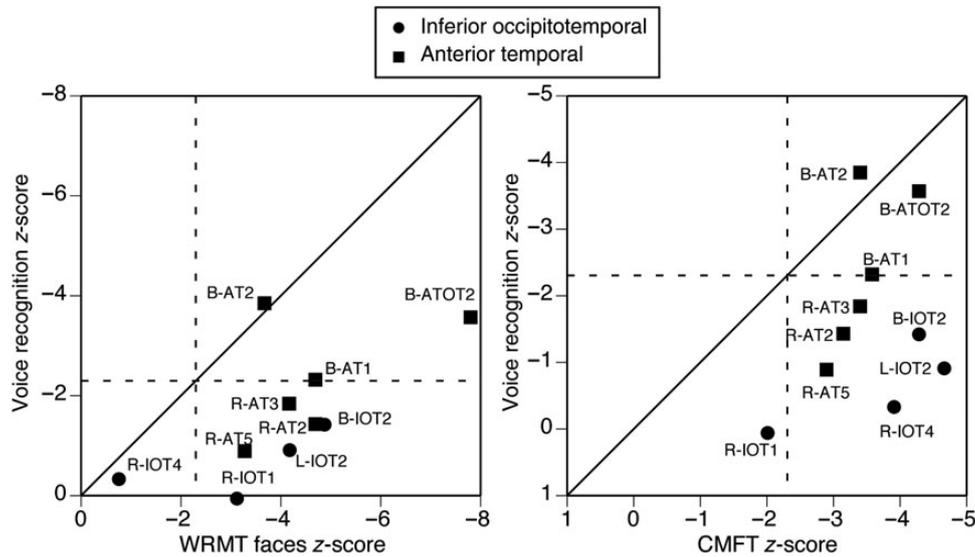


Figure 4. Comparison between voice recognition and face recognition in prosopagnosia. Scores on our voice recognition and standard neuropsychological face tests have been converted into age-adjusted z-scores, where a more negative z-score indicates greater impairment. Left graph shows voice recognition plotted against Warrington Recognition Memory Test (face component—WRMT). Right graph shows voice recognition plotted against Cambridge Face Memory Test (CMFT). Horizontal and vertical lines show the upper limits of normal scores, while the diagonal line indicates equivalent performance on face and voice recognition. Only B-AT2 shows deficits of similar severity on both face and voice recognition in both graphs.

Table 6. Questionnaire results (score 0-35).

	Faces	Voices
Control mean	10.08	13.06
s.d.	3.72	4.10
95% upper limit	18.87	22.75
R-IOT1	<u>27</u>	11
R-IOT4	<u>21</u>	9
B-IOT2	<u>35</u>	7
L-IOT2	<u>31</u>	22
B-ATOT2	<u>35</u>	14
R-AT2	<u>28</u>	18
R-AT3	<u>30</u>	16
R-AT5	<u>29</u>	19
B-AT1	<u>28</u>	<u>32</u>
B-AT2	<u>35</u>	7

Note: Underlining denotes impaired values.

with signal changes in the fusiform face area, possibly through top-down activation from supramodal regions (von Kriegstein et al. 2005), and tractography with diffusion tensor imaging has shown connections between the fusiform face area and temporal voice areas (Blank et al. 2011). Along with behavioral (von Kriegstein et al. 2006) and electrophysiological (Schweinberger et al. 2011) evidence for face-voice interactions, these have led to the inclusion of interconnectivity between face and voice processing streams in some recent models (Gainotti 2014). Hence, these suggest at least a theoretical possibility that fusiform lesions could adversely affect familiarity for voices. Our results show that they do not.

Second, the data of our 3 subjects with right anterior temporal damage address the question of whether such lesions cause a modality-specific prosopagnosia (Gainotti 2010, 2013a, 2013b). Voice testing had seldom been done in older reports, and it was noted (Gainotti 2010) that an older case of purported associative

prosopagnosia had an asymptomatic voice recognition deficit (Boudouresques et al. 1979). However, none of our 3 subjects with right anterior temporal damage alone showed impaired name or voice familiarity. Furthermore, both R-AT2 and R-AT3 had the pattern of preserved face discrimination with impairments in face familiarity and face imagery that points to an associative/amnesic variant of prosopagnosia. Our data on R-AT2 and R-AT3 thus establish an important theoretical point, that right anterior temporal damage can cause a modality-specific associative prosopagnosia.

Third, problems with familiarity across multiple modalities occurred in 3 subjects, all with bilateral lesions. B-ATOT2 had impaired voice and face discrimination, indicating apperceptive defects. As it is implausible that an apperceptive voice defect has the same basis as an apperceptive face defect, these are likely independent impairments. Only the 2 subjects with bilateral anterior temporal lesions had a pattern of spared voice and face perception with impaired voice and face recognition that would be consistent with either combined associative deficits for face and voice or a multimodal post-perceptual disorder.

Testing of both voice discrimination and recognition was important to discriminate apperceptive from associative defects. Only a few studies (Neuner and Schweinberger 2000; Garrido et al. 2009; Hailstone et al. 2010) have probed the cognitive hierarchy of voice processing in the manner done for face processing. In particular, voice discrimination has rarely been assessed. Some have noted normal perception of the speaker's gender, size, or age (Gentileschi et al. 2001; Hailstone et al. 2010), but given that perception of facial gender and age can be dissociated from perception of facial identity (Tranel et al. 1988), this may not be an adequate indicator of intact coding of voice identity. Better has been the use of a same/different task with unfamiliar voices (Van Lancker and Kreiman 1987; Neuner and Schweinberger 2000; Garrido et al. 2009; Hailstone et al. 2010).

Otherwise, most studies reported on voice identification alone, with or without an assessment of familiarity, and the degree of detail in these reports varies. There are anecdotal

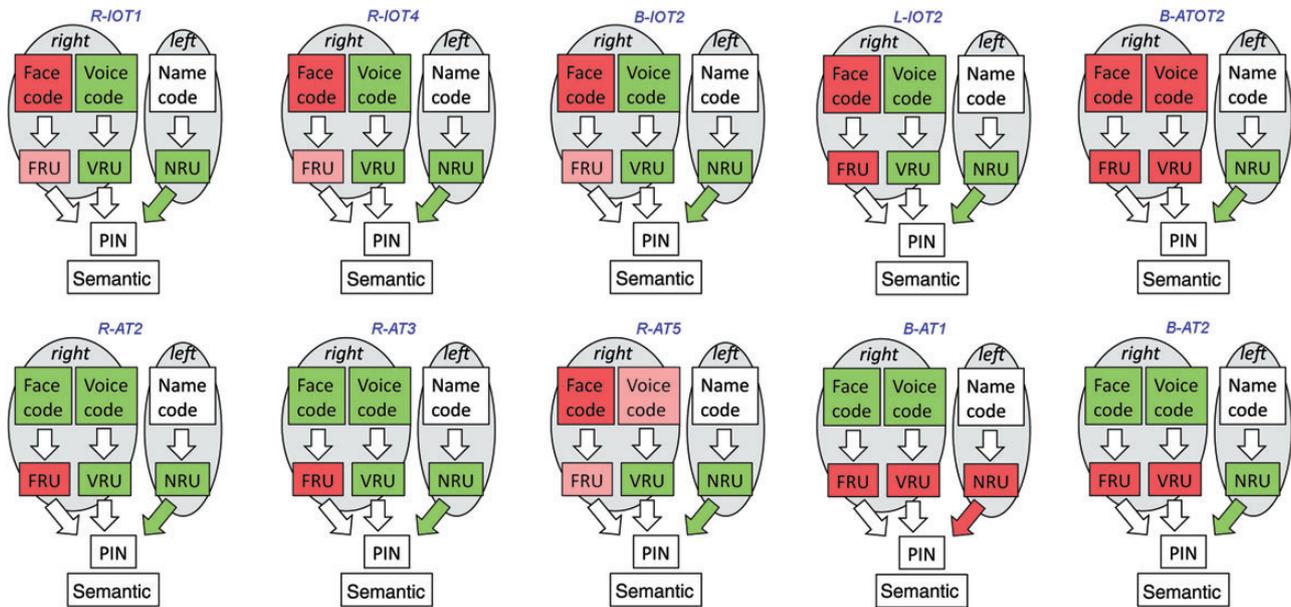


Figure 5. Schematic of the subjects' deficits in a person-processing hierarchy for faces, voices, and names. The first level, coding, is assessed by discriminative tests, namely the perception of facial configuration and our match-to-sample voice test. The output of the recognition units (FRU, VRU, and NRU for faces, voices, and names respectively) level is stimulus familiarity, as assessed by the Warrington Recognition Test for faces, our voice recognition test, and our name familiarity test. Identification of names, as indexed by occupation sorting, is represented by the arrow showing access from NRUs to person identity nodes (PINs). Green shades indicate normal performance, and red shades indicate impairment: the first appearance of a red shade indicates the primary impairment. Pink-shaded FRUs indicate impaired familiarity but preserved face imagery, suggesting that the familiarity defect is a downstream consequence of impaired face discrimination, rather than damage to FRUs. The pink shade of voice coding in R-AT5 reflects the fact that it is not certain whether her impaired voice discrimination is all or partly due to her unilateral deafness.

comments that subjects could identify people by voice (Evans et al. 1995; Mendez and Ghajarnia 2001; Joubert et al. 2003), or that voices did not help (De Renzi 1986). Some have constructed tests with the voices of family members to probe both familiarity and identification (Boudouresques et al. 1979; Gentileschi et al. 1999, 2001; Nakachi et al. 2007; Busigny et al. 2009). These invariably suffer from having few test items and either no control data (Gentileschi et al. 1999; Busigny et al. 2009) or a single family control (Gentileschi et al. 2001; Nakachi et al. 2007). Other studies have turned to famous voices: 2 case reports (Ellis et al. 1989; Hanley et al. 1989) assessed identification with a standardized famous voice test (Meudell et al. 1980), while smaller samples of famous voices have been used to test naming (Boudouresques et al. 1979; Gainotti et al. 2003) or familiarity and identification (Gainotti et al. 2008), but with no anonymous foils and only one or no controls. Only 3 investigations comprehensively assessed both familiarity and naming with larger samples of famous and anonymous voices, data from multiple control subjects, as well as same/different tests of voice discrimination (Neuner and Schweinberger 2000; Garrido et al. 2009; Hailstone et al. 2010).

Less is known about the neural substrate of voice than face recognition (Belin et al. 2004; Latinus and Belin 2011). Functional neuroimaging has revealed "temporal voice areas" that respond more to vocal than environmental sounds in the middle and anterior STS, bilaterally but more on the right (Belin et al. 2000). Adaptation fMRI studies have revealed sensitivity to voice identity in the right anterior (Belin and Zatorre 2003) or bilateral posterior STS (Warren et al. 2006). Familiarity for voices correlated with signal changes in more anterior parts of the STS and the superior middle temporal gyrus (Bethmann et al. 2012). In a training study, long-term sensitivity to voice identity localized to the anterior STS (Andics et al. 2010). Functional imaging in monkeys has shown analogous voice-sensitive areas in

the superior temporal plane (Petkov et al. 2008), and single-cell recordings confirm voice-selective cells in these regions (Perrodin et al. 2011).

The anatomic relation of voice-selective to face-selective areas in the anterior temporal lobe is important. Studies show a face-selective area in the anterior inferior temporal cortex that is sensitive to facial identity (Kriegeskorte et al. 2007; Tsao et al. 2008; Rajimehr et al. 2009). One study of face familiarity found adaptation effects in the right anterior middle temporal gyrus, left anterior STS, and bilateral temporal poles (Sugiura et al. 2001). However, these face-selective regions are more ventral than the location of temporal voice areas (Belin et al. 2000). Studies that have examined both face and voice stimuli have not found overlap in the anterior temporal cortex (Shah et al. 2001; Joassin et al. 2011).

Neuropsychological studies also support independence of face and voice recognition. One study found among 9 subjects with right hemispheric damage, 3 with deficits in naming either faces or voices, but one with a deficit in face naming alone and one with impaired voice naming (Van Lancker and Canter 1982). Another large study concluded that recognition deficits for faces, voices, and names are dissociable (Neuner and Schweinberger 2000): Only one subject was impaired on all three. Four had impaired voice but normal face recognition, 3 with normal voice matching. These could be considered as having a modality-specific associative phonagnosia. One subject with a right-sided lesion (case 22) was impaired on face familiarity, but did well on face matching, voice matching, and voice recognition: This subject is the most similar to R-AT2 and R-AT3, although he also had impaired name familiarity.

The evidence for independence of voice and face familiarity from prior case reports is highly variable in quality. Subject SO, with developmental prosopagnosia and intact face-matching

consistent with an associative variant (Kress and Daum 2003), had impaired short-term memory for voices (von Kriegstein et al. 2006). Subject SB with acquired prosopagnosia from age 3 actually showed superior voice recognition (Hoover et al. 2010); however, he had severe apperceptive defects and more generalized visual agnosia due to bilateral occipital lesions. Subject MT with acquired prosopagnosia due to temporal atrophy could still recognize and identify personally familiar voices (Nakachi et al. 2007), but was not tested on face discrimination to prove that his prosopagnosia was associative. The same is true for subject MS, who had prosopagnosia from bilateral fusiform and anterior temporal pole lesions, and yet had preserved ability to discriminate familiar from unfamiliar voices (Arnott et al. 2008). Subject KH with developmental phonagnosia had intact voice discrimination, face perception, and face recognition (Garrido et al. 2009), consistent with a modality-specific associative deficit in voice recognition. On the other hand, there are at least 6 subjects with impaired familiarity for both voices and faces, 5 of whom had temporal atrophy (Gentileschi et al. 1999, 2001; Gainotti et al. 2008; Hailstone et al. 2010), the other with encephalitis (Boudouresques et al. 1979). However, in some of these subjects, voice discrimination was either not tested (Boudouresques et al. 1979; Gentileschi et al. 1999; Gainotti et al. 2008) or may have been affected (Gentileschi et al. 2001), whereas intact face discrimination was claimed mainly on the basis of performance on the Benton Face Recognition Test, which as we have seen in our subjects may not be definitive.

The question of whether right anterior temporal lesions always cause a combined voice and face processing deficit raises the important issue of what is meant by a multimodal person recognition disorder. The use of the singular form would suggest a defect in one amodal cognitive operation. In cognitive models, this might correspond to the person identity node (Bruce and Young 1986; Burton et al. 1999). On the other hand, the proximity of neural substrates for voice and face processing suggests the potential for co-occurrence of modality-specific voice and face processing disorders. In this case, the better term may be a multimodal person recognition syndrome, with impairments in separate, modality-specific operations. In the cognitive models, these would correspond to face and voice recognition units. The distinction is critical on empirical and theoretical grounds. Empirically, only a syndrome would allow for dissociations between face and voice recognition beyond apperception. Theoretically, in a syndrome, the mechanism underlying a primary defect in face familiarity would not differ whether voice familiarity was also impaired or not.

Highly relevant to this issue is the debate about where familiarity for people is generated. In the model of Bruce and Young (1986), familiarity was viewed as a product of successful matching to modality-specific recognition units. An alternative model placed familiarity decisions at a supramodal level, in person identity nodes (Burton et al. 1990 1999). The possibility that familiarity might be generated for both the stimulus and the person has also been entertained (Haslam et al. 2001; Gainotti 2007), and one might indeed serve to reinforce the other. Nevertheless, it has been argued that the neuropsychological data are inconsistent with familiarity arising in amodal person identity nodes (Gainotti 2007). First, since person identity nodes are also responsible for access to name and semantic data, this would suggest a close correlation between familiarity and identification. However, it is clear that familiarity can occur without identification in neuropsychological cases (Warrington and McCarthy 1988; Geva et al. 1997; Haslam et al. 2001; Mendez and Ghajarnia 2001). Second, since person identity nodes are amodal,

familiarity in one modality should correlate with familiarity in another, and dissociations should not occur. However, cases 5 and 24 of Neuner and Schweinberger (2000) show an associative phonagnosia with sparing of face and name recognition, while case 22 (Neuner and Schweinberger 2000), MT (Nakachi et al. 2007), and our cases R-AT2 and R-AT3 show the reverse, an associative prosopagnosia with sparing of voice recognition. Such dissociations support an attribution of familiarity to modality-specific recognition units rather than amodal person identity nodes.

What are we to make of our 3 subjects with impaired familiarity for both faces and voices? In B-ATOT2, impaired familiarity is likely secondary to a combined apperceptive prosopagnosia and apperceptive phonagnosia, due to separate auditory and visual processing defects. We are aware of one other subject with a pattern indicative of apperceptive phonagnosia, case 1 of Neuner and Schweinberger (2000). As with that case, our additional tests confirmed that B-ATOT2 has intact hearing and recognition of sounds other than voices. Hence, she does not have a general auditory agnosia, a potential consequence of childhood herpes simplex encephalitis (Kaga et al. 2000, 2003). Rather, her phonagnosia is a selective auditory agnosia, just as her apperceptive prosopagnosia is a selective visual agnosia. While fusiform damage likely accounts for her apperceptive prosopagnosia (Barton et al. 2002; Fox et al. 2011), it is less clear which aspects of her complex lesions account for apperceptive phonagnosia. Adaptation fMRI studies show signals relevant to voice discrimination in the right STS (Belin and Zatorre 2003; Warren et al. 2006), while a morphometric analysis of Alzheimer's disease suggests that voice discrimination impairments are associated with right inferior parietal and right parahippocampal cortical loss (Hailstone et al. 2011).

B-AT2 and B-AT1 have intact discrimination of voices and faces, consistent with associative deficits. B-AT2 had normal familiarity and identification for names, and thus is most similar to QR (Hailstone et al. 2010) and case 37 (Neuner and Schweinberger 2000). We would argue that these 3 cases likely have a combined associative prosopagnosia/phonagnosia, particularly given their preserved ability to recognize names and/or identify people from their names. On the other hand, B-AT1, like KL (Hailstone et al. 2010), Emma (Gentileschi et al. 2001), and case 13 (Neuner and Schweinberger 2000), has a familiarity problem that affects faces, voices, and names. Although this might suggest an amodal dysfunction, a key point is that it is familiarity rather than identification alone that is affected in these subjects, and familiarity can be dissociated between faces, voices, and names in other subjects with associative deficits. Thus, the most parsimonious explanation is that they too have a multiple modality syndrome of person recognition, rather than a single cognitive dysfunction. Perhaps indirectly supporting this conclusion is the observation that all but one (case 37) of the above subjects with associative deficits in multiple modalities had either bilateral lesions or frontotemporal degeneration, a rare condition that even when asymmetric is almost certain to have diffuse bilateral effects, and therefore highly likely to affect multiple cognitive operations.

Our data and the anatomic, neuropsychologic, and imaging findings reviewed above suggest that voice and face recognition are independent functions. Given the proximity of their neural substrates, though, it would not be surprising to find an occasional subject with impairments of both, but that would not prove that a single cognitive deficit causes both, any more than would the frequent association between achromatopsia and apperceptive prosopagnosia imply a common origin for those 2 deficits. Subjects like R-AT2, R-AT3, case 22 (Neuner and

Schweinberger 2000), and possibly MT (Nakachi et al. 2007) show that impaired face recognition with spared face perception can occur with normal voice recognition, indicating that associative prosopagnosia is not just theoretical but a real entity, and, in the case of R-AT2 and R-AT3, can follow right anterior temporal damage. Conversely, cases of impaired voice recognition with spared voice perception and normal face recognition show that associative phonagnosia also exists and can be similarly modality-specific (Neuner and Schweinberger 2000; Garrido et al. 2009). Such findings raise questions about whether multimodal problems of person familiarity should be viewed as a single amodal disorder or rather as the co-occurrence of different modality-specific deficits in one syndrome, due to anatomic proximity of their neural substrates.

Supplementary Material

Supplementary material can be found at: <http://www.cercor.oxfordjournals.org/>

Funding

This work was funded by CIHR grant (MOP-102567). R.R.L. was funded by a Fight for Sight summer student fellowship and a medical student summer research scholarship from the American Academy of Neurology. J.J.S.B. was funded by a Canada Research Chair and the Marianne Koerner Chair in Brain Diseases.

Notes

We thank Eleni Nasiopoulos, Alan Kingstone, Sherryse Corrow, and Jodie Davies-Thompson for technical assistance. *Conflict of Interest*: None declared.

References

- Albert M, Butters N, Levin J. 1979. Temporal gradients in retrograde amnesia of patients with alcoholic Korsakoff's disease. *Arch Neurol*. 36:211–216.
- Andics A, McQueen JM, Petersson KM, Gal V, Rudas G, Vidnyanszky Z. 2010. Neural mechanisms for voice recognition. *Neuroimage*. 52:1528–1540.
- Arnott SR, Heywood CA, Kentridge RW, Goodale MA. 2008. Voice recognition and the posterior cingulate: an fMRI study of prosopagnosia. *J Neuropsychol*. 2:269–286.
- Barsics C, Bredart S. 2012. Recalling semantic information about newly learned faces and voices. *Memory*. 20:527–534.
- Barton J. 2008. Structure and function in acquired prosopagnosia: lessons from a series of ten patients with brain damage. *J Neuropsychol*. 2:197–225.
- Barton J, Cherkasova M. 2003. Face imagery and its relation to perception and covert recognition in prosopagnosia. *Neurology*. 61:220–225.
- Barton J, Cherkasova M, O'Connor M. 2001. Covert recognition in acquired and developmental prosopagnosia. *Neurology*. 57:1161–1167.
- Barton J, Press D, Keenan J, O'Connor M. 2002. Lesions of the fusiform face area impair perception of facial configuration in prosopagnosia. *Neurology*. 58:71–78.
- Barton J, Zhao J, Keenan J. 2003. Perception of global facial geometry in the inversion effect and prosopagnosia. *Neuropsychologia*. 41:1703–1711.
- Barton JJ, Hanif H, Ashraf S. 2009. Relating visual to verbal semantic knowledge: the evaluation of object recognition in prosopagnosia. *Brain*. 132:3456–3466.
- Belin P, Fecteau S, Bedard C. 2004. Thinking the voice: neural correlates of voice perception. *Trends Cogn Sci*. 8:129–135.
- Belin P, Zatorre RJ. 2003. Adaptation to speaker's voice in right anterior temporal lobe. *Neuroreport*. 14:2105–2109.
- Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B. 2000. Voice-selective areas in human auditory cortex. *Nature*. 403:309–312.
- Benton A, van Allen M. 1972. Prosopagnosia and facial discrimination. *J Neurol Sci*. 15:167–172.
- 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:e47626.
- Blank H, Anwender A, von Kriegstein K. 2011. Direct structural connections between voice- and face-recognition areas. *J Neurosci*. 31:12906–12915.
- Boudouresques J, Poncet M, Cherif AA, Balzamo M. 1979. Agnosia for faces: evidence of functional disorganization of a certain type of recognition of objects in the physical world. *Bull Acad Natl Med*. 163:695–702.
- Bruce V, Young A. 1986. Understanding face recognition. *Br J Psychol*. 77:305–327.
- Burton A, Bruce V, Hancock P. 1999. From pixels to people: a model of familiar face recognition. *Cogn Sci*. 23:1–31.
- Burton A, Bruce V, Johnston R. 1990. Understanding face recognition with an interactive activation model. *Br J Psychol*. 81:361–380.
- Burton AM, Jenkins R, Schweinberger SR. 2011. Mental representations of familiar faces. *Br J Psychol*. 102:943–958.
- 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.
- Busigny T, Van Belle G, Jemel B, Hosen A, Joubert S, Rossion B. 2014. Face-specific impairment in holistic perception following focal lesion of the right anterior temporal lobe. *Neuropsychologia*. 56:312–333.
- Caldara R, Schyns P, Mayer E, Smith ML, Gosselin F, Rossion B. 2005. Does prosopagnosia take the eyes out of face representations? Evidence for a defect in representing diagnostic facial information following brain damage. *J Cogn Neurosci*. 17:1652–1666.
- Dalrymple KA, Oruc I, Duchaine B, Pancaroglu R, Fox CJ, Iaria G, Handy TC, Barton JJ. 2011. The anatomic basis of the right face-selective N170 in acquired prosopagnosia: a combined ERP/fMRI study. *Neuropsychologia*. 49:2553–2563.
- Damasio A, Tranel D, Damasio H. 1990. Face agnosia and the neural substrates of memory. *Annu Rev Neurosci*. 13:89–109.
- Davies-Thompson J, Pancaroglu R, Barton J. 2014. Acquired prosopagnosia: structural basis and processing impairments. *Front Biosci (Elite Ed)*. 6:159–174.
- De Renzi E. 1986. Current issues on prosopagnosia. In: Ellis HD, editor. *Aspects of face processing*. Dordrecht: Nijoff. p. 243–252.
- de Renzi E, Faglioni P, Grossi D, Nichelli P. 1991. Apperceptive and associative forms of prosopagnosia. *Cortex*. 27:213–221.
- Duchaine BC, Nakayama K. 2006. The Cambridge Face Memory Test: results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic patients. *Neuropsychologia*. 44:576–585.
- Ellis AW, Young AW, Critchley EM. 1989. Loss of memory for people following temporal lobe damage. *Brain*. 112(Pt 6): 1469–1483.

- Evans JJ, Heggis AJ, Antoun N, Hodges JR. 1995. Progressive prosopagnosia associated with selective right temporal lobe atrophy. A new syndrome? *Brain*. 118(Pt 1):1–13.
- Farah MJ. 1990. *Visual agnosia: disorders of visual recognition and what they tell us about normal vision*. Cambridge: MIT Press.
- Fox CJ, Hanif HM, Iaria G, Duchaine BC, Barton JJ. 2011. Perceptual and anatomic patterns of selective deficits in facial identity and expression processing. *Neuropsychologia*. 49:3188–3200.
- Fox CJ, Iaria G, Barton JJ. 2009. Defining the face processing network: optimization of the functional localizer in fMRI. *Hum Brain Mapp*. 30:1637–1651.
- Fox CJ, Iaria G, Duchaine BC, Barton JJ. 2013. Residual fMRI sensitivity for identity changes in acquired prosopagnosia. *Front Psychol*. 4:756.
- Gainotti G. 2014. Cognitive models of familiar people recognition and hemispheric asymmetries. *Front Biosci (Elite Ed)*. 6:148–158.
- Gainotti G. 2007. Face familiarity feelings, the right temporal lobe and the possible underlying neural mechanisms. *Brain Res Rev*. 56:214–235.
- Gainotti G. 2013a. Is the right anterior temporal variant of prosopagnosia a form of “associative prosopagnosia” or a form of “multimodal person recognition disorder”? *Neuropsychol Rev*. 23:99–110.
- Gainotti G. 2013b. Laterality effects in normal subjects’ recognition of familiar faces, voices and names. Perceptual and representational components. *Neuropsychologia*. 51:1151–1160.
- Gainotti G. 2010. Not all patients labeled as “prosopagnosia” have a real prosopagnosia. *J Clin Exp Neuropsychol*. 32:763–766.
- Gainotti G, Barbier A, Marra C. 2003. Slowly progressive defect in recognition of familiar people in a patient with right anterior temporal atrophy. *Brain*. 126:792–803.
- Gainotti G, Ferraccioli M, Quaranta D, Marra C. 2008. Cross-modal recognition disorders for persons and other unique entities in a patient with right fronto-temporal degeneration. *Cortex*. 44:238–248.
- Gainotti G, Marra C. 2011. Differential contribution of right and left temporo-occipital and anterior temporal lesions to face recognition disorders. *Front Hum Neurosci*. 5:55.
- Gainotti G, Spinelli P, Scaricamazza E, Marra C. 2013. The evaluation of sources of knowledge underlying different conceptual categories. *Front Hum Neurosci*. 7:40.
- Garrido L, Eisner F, McGettigan C, Stewart L, Sauter D, Hanley JR, Schweinberger SR, Warren JD, Duchaine B. 2009. Developmental phonagnosia: a selective deficit of vocal identity recognition. *Neuropsychologia*. 47:123–131.
- Gentileschi V, Sperber S, Spinnler H. 2001. Crossmodal agnosia for familiar people as a consequence of right infero polar temporal atrophy. *Cogn Neuropsychol*. 18:439–463.
- Gentileschi V, Sperber S, Spinnler H. 1999. Progressive defective recognition of people. *Neurocase*. 5:407–424.
- Geva A, Moscovitch M, Leach L. 1997. Perceptual priming of proper names in young and older normal adults and a patient with prosopagnosia. *Neuropsychology*. 11:232–242.
- Hailstone JC, Crutch SJ, Vestergaard MD, Patterson RD, Warren JD. 2010. Progressive associative phonagnosia: a neuropsychological analysis. *Neuropsychologia*. 48:1104–1114.
- Hailstone JC, Ridgway GR, Bartlett JW, Goll JC, Buckley AH, Crutch SJ, Warren JD. 2011. Voice processing in dementia: a neuropsychological and neuroanatomical analysis. *Brain*. 134:2535–2547.
- Hanley JR, Young AW, Pearson NA. 1989. Defective recognition of familiar people. *Cogn Neuropsychol*. 6:179–210.
- Haslam C, Cook M, Coltheart M. 2001. “I know your name but not your face”: explaining modality-based differences in access to biographical knowledge in a patient with retrograde amnesia. *Neurocase*. 7:189–199.
- Hécaen H. 1981. The neuropsychology of face recognition. In: Davies G, Ellis H, Shepherd J, editors. *Perceiving and remembering faces*. London: Academic Press. p. 39–54.
- Hoover AE, Demonet JF, Steeves JK. 2010. Superior voice recognition in a patient with acquired prosopagnosia and object agnosia. *Neuropsychologia*. 48:3725–3732.
- Iaria G, Fox CJ, Waite C, Aharon I, Barton JJS. 2008. The contribution of the fusiform gyrus and superior temporal sulcus in processing facial attractiveness: neuropsychological and neuroimaging evidence. *Neuroscience*. 155:409–422.
- Joassin F, Pesenti M, Maurage P, Verreault E, Bruyer R, Campanella S. 2011. Cross-modal interactions between human faces and voices involved in person recognition. *Cortex*. 47:367–376.
- Joubert S, Felician O, Barbeau E, Sontheimer A, Barton J, Ceccaldi M, Poncet M. 2003. Impaired configurational processing in a case of progressive prosopagnosia associated with right temporal lobe atrophy. *Brain*. 126:2537–2550.
- Kaga K, Kaga M, Tamai F, Shindo M. 2003. Auditory agnosia in children after herpes encephalitis. *Acta Otolaryngol*. 123:232–235.
- Kaga M, Shindo M, Kaga K. 2000. Long-term follow-up of auditory agnosia as a sequel of herpes encephalitis in a child. *J Child Neurol*. 15:626–629.
- Kress T, Daum I. 2003. Event-related potentials reflect impaired face recognition in patients with congenital prosopagnosia. *Neurosci Lett*. 352:133–136.
- Kriegeskorte N, Formisano E, Sorger B, Goebel R. 2007. Individual faces elicit distinct response patterns in human anterior temporal cortex. *Proc Natl Acad Sci USA*. 104:20600–20605.
- Latinus M, Belin P. 2011. Human voice perception. *Curr Biol*. 21:R143–R145.
- Mendez MF, Ghajarnia M. 2001. Agnosia for familiar faces and odors in a patient with right temporal lobe dysfunction. *Neurology*. 57:519–521.
- Meudell PR, Northen B, Snowden JS, Neary D. 1980. Long term memory for famous voices in amnesic and normal subjects. *Neuropsychologia*. 18:133–139.
- Nakachi R, Muramatsu T, Kato M, Akiyama T, Saito F, Yoshino F, Mimura M, Kashima H. 2007. Progressive prosopagnosia at a very early stage of frontotemporal lobar degeneration. *Psychogeriatrics*. 7:155–162.
- Natu V, O’Toole AJ. 2011. The neural processing of familiar and unfamiliar faces: a review and synopsis. *Br J Psychol*. 102:726–747.
- Neuner F, Schweinberger SR. 2000. Neuropsychological impairments in the recognition of faces, voices, and personal names. *Brain Cogn*. 44:342–366.
- Perrodin C, Kayser C, Logothetis NK, Petkov CI. 2011. Voice cells in the primate temporal lobe. *Curr Biol*. 21:1408–1415.
- Petkov CI, Kayser C, Steudel T, Whittingstall K, Augath M, Logothetis NK. 2008. A voice region in the monkey brain. *Nat Neurosci*. 11:367–374.
- Rajimehr R, Young JC, Tootell RB. 2009. An anterior temporal face patch in human cortex, predicted by macaque maps. *Proc Natl Acad Sci USA*. 106:1995–2000.
- Ramon M, Busigny T, Rossion B. 2010. Impaired holistic processing of unfamiliar individual faces in acquired prosopagnosia. *Neuropsychologia*. 48:933–944.

- Schweinberger SR, Kloth N, Robertson DM. 2011. Hearing facial identities: brain correlates of face—voice integration in person identification. *Cortex*. 47:1026–1037.
- Shah NJ, Marshall JC, Zafiris O, Schwab A, Zilles K, Markowitsch HJ, Fink GR. 2001. The neural correlates of person familiarity. A functional magnetic resonance imaging study with clinical implications. *Brain*. 124:804–815.
- Sugiura M, Kawashima R, Nakamura K, Sato N, Nakamura A, Kato T, Hatano K, Schormann T, Zilles K, Sato K, et al. 2001. Activation reduction in anterior temporal cortices during repeated recognition of faces of personal acquaintances. *Neuroimage*. 13:877–890.
- Tranel D, Damasio AR, Damasio H. 1988. Intact recognition of facial expression, gender, and age in patients with impaired recognition of face identity. *Neurology*. 38:690–696.
- Tsao DY, Moeller S, Freiwald WA. 2008. Comparing face patch systems in macaques and humans. *Proc Natl Acad Sci USA*. 105:19514–19519.
- Van Lancker D, Kreiman J. 1987. Voice discrimination and recognition are separate abilities. *Neuropsychologia*. 25:829–834.
- Van Lancker DR, Canter GJ. 1982. Impairment of voice and face recognition in patients with hemispheric damage. *Brain Cogn*. 1:185–195.
- von Kriegstein K, Kleinschmidt A, Giraud AL. 2006. Voice recognition and cross-modal responses to familiar speakers' voices in prosopagnosia. *Cereb Cortex*. 16:1314–1322.
- von Kriegstein K, Kleinschmidt A, Sterzer P, Giraud AL. 2005. Interaction of face and voice areas during speaker recognition. *J Cogn Neurosci*. 17:367–376.
- Warren JD, Scott SK, Price CJ, Griffiths TD. 2006. Human brain mechanisms for the early analysis of voices. *Neuroimage*. 31:1389–1397.
- Warrington E. 1984. Warrington Recognition Memory Test. Los Angeles: Western Psychological Services.
- Warrington EK, McCarthy RA. 1988. The fractionation of retrograde amnesia. *Brain Cogn*. 7:184–200.
- Weiner KS, Golarai G, Caspers J, Chuapoco MR, Mohlberg H, Zilles K, Amunts K, Grill-Spector K. 2014. The mid-fusiform sulcus: a landmark identifying both cytoarchitectonic and functional divisions of human ventral temporal cortex. *Neuroimage*. 84:453–465.