



Research report

The processing of voice identity in developmental prosopagnosia



Ran R. Liu ^a, Sherryse L. Corrow ^{a,*}, Raika Pancaroglu ^a,
Brad Duchaine ^b and Jason J.S. Barton ^a

^a Human Vision and Eye Movement Laboratory, Departments of Medicine (Neurology), Ophthalmology and Visual Sciences, University of British Columbia, Eye Care Centre, Vancouver, BC, Canada

^b Psychological and Brain Sciences, Dartmouth College, Hanover, NH, USA

ARTICLE INFO

Article history:

Received 30 March 2015

Reviewed 21 May 2015

Revised 17 June 2015

Accepted 20 July 2015

Action editor Paolo Bartolomeo

Published online 4 August 2015

Keywords:

Voice perception

Familiarity

Semantic

Multi-modal

Person recognition

ABSTRACT

Background: Developmental prosopagnosia is a disorder of face recognition that is believed to reflect impairments of visual mechanisms. However, voice recognition has rarely been evaluated in developmental prosopagnosia to clarify if it is modality-specific or part of a multi-modal person recognition syndrome.

Objective: Our goal was to examine whether voice discrimination and/or recognition are impaired in subjects with developmental prosopagnosia.

Design/methods: 73 healthy controls and 12 subjects with developmental prosopagnosia performed a match-to-sample test of voice discrimination and a test of short-term voice familiarity, as well as a questionnaire about face and voice identification in daily life.

Results: Eleven subjects with developmental prosopagnosia scored within the normal range for voice discrimination and voice recognition. One was impaired on discrimination and borderline for recognition, with equivalent scores for face and voice recognition, despite being unaware of voice processing problems.

Conclusions: Most subjects with developmental prosopagnosia are not impaired in short-term voice familiarity, providing evidence that developmental prosopagnosia is usually a modality-specific disorder of face recognition. However, there may be heterogeneity, with a minority having additional voice processing deficits. Objective tests of voice recognition should be integrated into the diagnostic evaluation of this disorder to distinguish it from a multi-modal person recognition syndrome.

© 2015 Elsevier Ltd. All rights reserved.

* Corresponding author. Human Vision and Eye Movement Laboratory, Departments of Medicine (Neurology), Ophthalmology and Visual Sciences, University of British Columbia, Eye Care Centre, 360A – 2550 Willow Street, Vancouver, BC, V5Z 3N9, Canada.

E-mail addresses: r.liu@alumni.ubc.ca (R.R. Liu), sherryse.corrow@eyecarecentre.org (S.L. Corrow), raikap@gmail.com (R. Pancaroglu), bradley.c.duchaine@dartmouth.edu (B. Duchaine), jasonbarton@shaw.ca (J.J.S. Barton).

<http://dx.doi.org/10.1016/j.cortex.2015.07.030>

0010-9452/© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Prosopagnosia is the inability to recognize faces. In line with cognitive models of face recognition as a hierarchy of operations (Bruce & Young, 1986), neuropsychological studies of acquired prosopagnosia suggest at least two functional variants, an apperceptive form in which subjects have difficulty perceiving facial structure, and an associative/amnesic form in which perception is relatively spared, but the ability to access memories of previously seen faces is disrupted (Barton, 2008; Davies-Thompson, Pancaroglu, & Barton, 2014). In acquired prosopagnosia, the apperceptive variant has been linked to right or bilateral occipitotemporal lesions, and the associative variant to right or bilateral anterior temporal lesions (Barton, 2008; Damasio, Tranel, & Damasio, 1990).

The fact that voice recognition also activates the right anterior temporal lobes (Belin & Zatorre, 2003) has prompted questions as to whether some cases of associative prosopagnosia may instead have a multi-modal person recognition disorder, since voice recognition is seldom tested (Gainotti, 2013). This is not only a question with theoretical implications, but also potential clinical relevance: as there is some evidence that simultaneous voice processing can enhance face learning, (Bulthoff & Newell, 2015), impaired voice processing could accentuate prosopagnosic deficits while intact voice processing might offer rehabilitative avenues. To address this, a recent study of acquired prosopagnosia found that voice recognition was unaffected after right anterior temporal lesions but impaired by bilateral anterior temporal lesions (Liu, Pancaroglu, Hills, Duchaine, & Barton, 2014). These data tentatively suggest that impaired face recognition can be divided into an apperceptive form linked to occipitotemporal lesions, an associative form linked to right anterior temporal lesions, and a multi-modal person recognition syndrome, with parallel associative deficits in face and voice processing, typically after bilateral anterior temporal lesions. These preliminary function-structure correlations require replication.

These observations are also relevant to the developmental form of prosopagnosia, whose functional and structural basis remains uncertain. Recent studies suggest the possibility of similar apperceptive, associative and other variants in developmental prosopagnosia (Dalrymple, Garrido, & Duchaine, 2014; Stollhoff, Jost, Elze, & Kennerknecht, 2011; Susilo & Duchaine, 2013). Less known is whether it is a modality-specific disorder or part of a multi-modal syndrome: currently there is only one report that studied voice recognition in one developmental prosopagnosic subject, which found some impairment (von Kriegstein, Kleinschmidt, & Giraud, 2006), and brief mention of inferior voice recognition in another who also had Asperger syndrome (Kracke, 1994). Addressing these functional questions in developmental prosopagnosia may help guide and refine studies of this condition.

The goal of this study was to apply the same two tests of voice discrimination and recognition used in the study of acquired prosopagnosia (Liu et al., 2014) to a cohort with developmental prosopagnosia. We asked a) whether this condition is a modality-specific disorder of face recognition or part of a

multi-modal person recognition syndrome, and b) whether the results were homogenous or heterogeneous in the cohort.

2. Material and methods

2.1. Subjects

We studied 12 subjects with developmental prosopagnosia (9 females, age range 27–67 years, mean 41.82 years, SD 12.91 years). These were local residents and native English speakers recruited from www.faceblind.org. Diagnostic criteria included an in-person clinical interview that revealed a subjective report of life-long difficulty in face recognition, and objective confirmation of impaired face recognition that included a score at least two standard deviations below the control mean on the Cambridge Face Memory Task (Duchaine & Nakayama, 2006) and a discordance between preserved word memory and impaired face memory on the Warrington Recognition Memory Test for Faces and Words (Warrington, 1984) that was in the bottom 5th percentile (Table 1). One subject had a unilateral retinal detachment, since corrected and now with normal acuity and fields; no other subject reported perceptual problems. Apart from mild concussions years prior in four subjects, none had other neurologic problems. None of these subjects reported any significant changes in face perception associated with their mild concussions or other medical events. All subjects had best corrected visual acuity of better than 20/60, normal visual fields on Goldmann perimetry, and normal color vision on the Farnsworth-Munsell hue discrimination test. To exclude autism spectrum disorders, all subjects had a score less than 32 on the Autism Questionnaire (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001).

Although it is not part of current diagnostic criteria (Behrmann & Avidan, 2005; Bate et al., 2014; Susilo & Duchaine, 2013), eight subjects had magnetic resonance brain imaging with T1-weighted and FLAIR sequences to exclude structural lesions that would have indicated early acquired rather than developmental prosopagnosia (Barton, Cherkasova, Press, Intriligator, & O'Connor, 2003); in one subject MRI was contraindicated and in the remaining three it was declined because of time limitations (Table 1). To evaluate face discrimination, subjects also completed the Cambridge Face Perception Test (CFPT; Duchaine, Yovel, & Nakayama, 2007).

Seventy-three control subjects completed the voice discrimination test (50 females, age range 19–70 years, mean 33.6 years, SD 15.5 years), 54 of whom also completed the voice recognition test (41 females, age range 19–70 years, mean age 37.2, SD 16.4 years). All subjects were born in North America, lived in North America for at least five years, and spoke English as their first language. All control subjects reported normal or corrected-to-normal vision, normal hearing, and no history of brain damage. All control subjects were remunerated ten dollars per hour for their participation.

All subjects gave informed consent to a protocol approved by the University of British Columbia and Vancouver General Hospital Ethics Review Boards.

Table 1 – Demographics and face processing scores.

| | Gender | Age (yr) | CFPT | | CFMT | WRMT | | | MRI |
|-------|--------|----------|-----------|-----------|-----------|-------|-------|-----------|--------|
| | | | Upright | Inverted | | Faces | Words | W–F | |
| DP003 | F | 36 | <u>74</u> | 90 | <u>35</u> | 28 | 50 | <u>22</u> | No |
| DP008 | F | 62 | 48 | <u>96</u> | <u>36</u> | 36 | 49 | <u>13</u> | Yes |
| DP014 | M | 43 | <u>64</u> | 60 | <u>32</u> | 30 | 48 | <u>18</u> | Yes |
| DP016 | F | 52 | 48 | 72 | <u>41</u> | 37 | 49 | <u>12</u> | Yes |
| DP021 | F | 29 | 36 | 80 | <u>37</u> | 33 | 50 | <u>17</u> | Yes |
| DP024 | F | 35 | 62 | 74 | <u>41</u> | 38 | 50 | <u>12</u> | Yes |
| DP029 | F | 28 | 44 | 76 | <u>42</u> | 36 | 49 | <u>13</u> | No |
| DP032 | M | 67 | <u>68</u> | 84 | <u>42</u> | 37 | 47 | <u>10</u> | No |
| DP033 | F | 47 | 52 | 82 | <u>29</u> | 39 | 50 | <u>11</u> | Contra |
| DP035 | M | 40 | <u>86</u> | 68 | <u>36</u> | 35 | 49 | <u>14</u> | Yes |
| DP038 | F | 28 | 32 | 66 | <u>39</u> | 36 | 49 | <u>13</u> | Yes |
| DP044 | F | 37 | <u>68</u> | 82 | <u>40</u> | 34 | 49 | <u>15</u> | Yes |

CFMT – Cambridge face memory test.

CFPT – Cambridge face perception test.

WRMT – Warrington recognition memory test, W–F = words score minus faces score.

Bold/Underline – Impaired performance.

2.2. Apparatus for testing voice processing

All tests were constructed with Superlab (www.superlab.com) software and performed on an IBM Lenovo laptop with 1280 × 800 pixels resolution. Subjects wore Panasonic RP-HTX7 headset throughout the test and were placed approximately 57 cm away from the screen, in a quiet and dimly lit room with the door closed.

2.3. Stimuli

Audio-clips were generated from volunteers aged 20–31 years (Liu et al., 2014): 20 males and 20 females for the discrimination test, and 21 males and 21 females for the recognition test. Each audio-clip was never used more than once as a target or as a distractor.

2.3.1. Voice discrimination

This used a match-to-sample design. Each volunteer read two different texts. For the initial sample, the volunteer 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 match and distractor stimuli, both voices read: “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”. All audio-clips began with 1 sec of silence, then approximately 10 sec of the voice, and ended with 1.5 sec of silence.

2.3.2. Voice recognition

Two sets of audio-clips were made, with each volunteer contributing two recordings to each set. The first “question component” was recorded in interview style, and the volunteers answer two questions. For the learning phase, we asked, “What was your favorite childhood activity?”, while for the test phase, we asked, “What was your favorite vacation?” The second “passages component” was recorded in narrative style. Volunteers read a random passage from Alice Munro's short story “Too Much Happiness” for the learning phase and for the

testing phase a random passage from the short story “Friend of My Youth”. All volunteers read different passages. For both components, each audio-clip began with 200 msec of silence followed by 12.5 sec of speech, then 700 msec of silence.

2.4. Protocol

2.4.1. Voice discrimination

Each trial started with a screen displaying “Target Voice” that was accompanied by the audio-clip of the sample voice. After a 1.5 sec pause, there was a ring tone lasting 875 msec, which served as an auditory mask. This was followed by the audio-clip of the first choice voice accompanied by a screen displaying “Choice 1”, and then the audio-clip of the second choice voice with the visual display “Choice 2”. One of the two choice voices was the match, from the same person to whom the sample voice belonged, and the other a distractor, from a different person of the same gender. Matches and distractors were given in random order. After the last choice voice, a screen prompted the subject to indicate with a keypress which choice voice matched the sample voice. There were 40 trials given in two blocks, one with 20 male and one with 20 female trials.

2.4.2. Voice recognition

This tested the ability to remember voices over a short interval and with intervening interference, similar to standard tests of face familiarity such as the Warrington Recognition Memory Test (Warrington, 1984) and the Cambridge Face Memory Test (Duchaine & Nakayama, 2006). Our pilot work found that healthy subjects performed at chance when asked to recall five voices sequentially. Hence our recognition test used sets of three trials. Each set started with a screen that displayed “Learning Phase” for 3 sec. During this phase, the subject heard audio-clips of three target voices in a row. During the first target voice the screen displayed “Voice A”, with the second it displayed “Voice B”, and with the third “Voice C”. Each audio-clip was followed by an 875 msec ring tone. After completion of the learning phase, which lasted approximately 46 sec, the screen displayed “Testing Phase” for 3 sec. Subjects

then heard three pairs of choice voices. In the first pair, one of the two choice voices was from the person who was “Voice A” in the learning phase, while the other was a distractor from a different person of the same gender. The temporal order of matches and distractors was randomized. When the first choice voice was playing, the screen displayed “Voice A Choice 1”, and with the second it displayed “Voice A Choice 2”. Each audio-clip was followed by an 875 msec ring tone. A screen then prompted subjects to indicate by keypress which of the two choice voices matched Voice A. This was then followed by a similar sequence for Voices B and C. Each set of three trials contained at least one trial for each gender. Subjects completed seven sets (21 trials) with stimuli from the ‘question component’, and another seven with stimuli from the ‘passages component’, for 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 blocks. There was a break of at least 10 min before the voice recognition test, in which the ‘question component’ was presented first and the ‘passages component’ second.

2.5. Analysis

We calculated accuracy for each subject. Because our prior evaluation of these control subjects showed an effect of age but not gender (Liu et al., 2014), our individual-subject analysis regressed out the variance due to age in controls, and used the residual variance in the function to calculate the 95% prediction intervals appropriate for single-subject comparisons.

Second, we performed group analyses. To generate age-matched control groups we used two methods to match our prosopagnosic group. The first ‘age-paired’ method paired each prosopagnosic subject with the control subject closest in age, giving 12 subjects in each group. For voice discrimination, this yielded a control group with mean age of 41.82 years (SD 12.7), equivalent to that of our prosopagnosic group [$t_{(11)} = .0002$, $p = .99$]. For voice recognition, this yielded a control group with mean age of 41.84 years (SD 12.7), also equivalent to that of our prosopagnosic group [$t_{(11)} = .005$, $p = .99$].

The second ‘age-range’ method capitalized on the statistical advantage of the large number of control subjects. We included all controls whose ages fell within the range spanning the minimum and maximum ages of our prosopagnosic subjects. For voice discrimination this gave a control group of 35 subjects whose mean age of 42.24 years (SD 13.0) did not differ from the prosopagnosia group [$t_{(20)} = .09$, $p = .93$], while for voice recognition this gave a control group of 32 subjects with mean age of 43.33 years (SD 13.0), again not differing from the prosopagnosia group [$t_{(19)} = .35$, $p = .73$]. Contrasts in voice scores were performed using t-tests for both the voice discrimination and voice recognition test.

2.6. Questionnaire about face and voice identification

Subjects with prosopagnosia often report that they rely on voices for people recognition. However, some have noted that these subjective reports may not correlate with objective testing (Boudouresques, Poncet, Cherif, & Balzamo, 1979; Liu et al., 2014). To examine the relationship between self-report

and our objective testing, we asked all subjects to complete the questionnaire about their daily life experiences in face and voice identification that we had administered to subjects with acquired prosopagnosia (Liu et al., 2014). Five questions addressed face identification and five voice identification, and subjects responded using a 7-point Likert scale, giving a face and voice score out of 35 each, with higher scores indicating more difficulty. All developmental prosopagnosia subjects and 43 control subjects (mean age 36.2 years, SD 15.6, range 19–70) completed the questionnaire. Our prior study found no significant effect of age or gender on questionnaire scores in control subjects (Liu et al., 2014).

3. Results

3.1. Voice discrimination

11 of 12 prosopagnosic subjects were normal, while one, DP035, had an abnormal score, at the 96.5th percentile (Fig. 1A). In the group analysis (Fig. 2), there was no difference compared to either the age-paired [$t_{(11)} = .15$, $p = .88$] or age-range controls [$t_{(30)} = .26$, $p = .73$].

3.2. Voice recognition

All 12 prosopagnosic subjects were normal, although DP035 was borderline, at the 94th percentile (Fig. 1B). The group

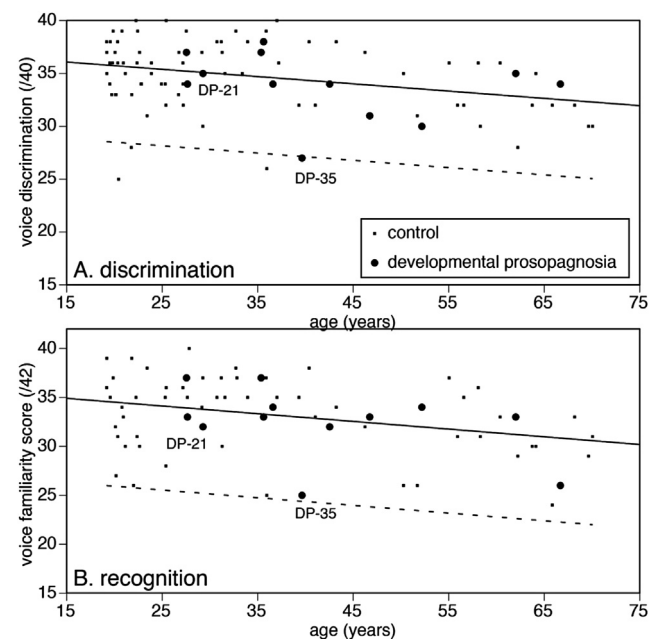


Fig. 1 – Voice scores plotted as a function of age. A. voice discrimination, B. voice recognition. Control subjects (small dots) in both tests show a significant decline with age. Solid line shows the mean of the linear regression and the dotted line shows the age-adjusted lower 95% prediction limit. Subject DP035 is impaired on voice discrimination and has borderline performance on voice recognition. Subject DP021 performs normally, despite reporting voice recognition difficulties on the questionnaire.

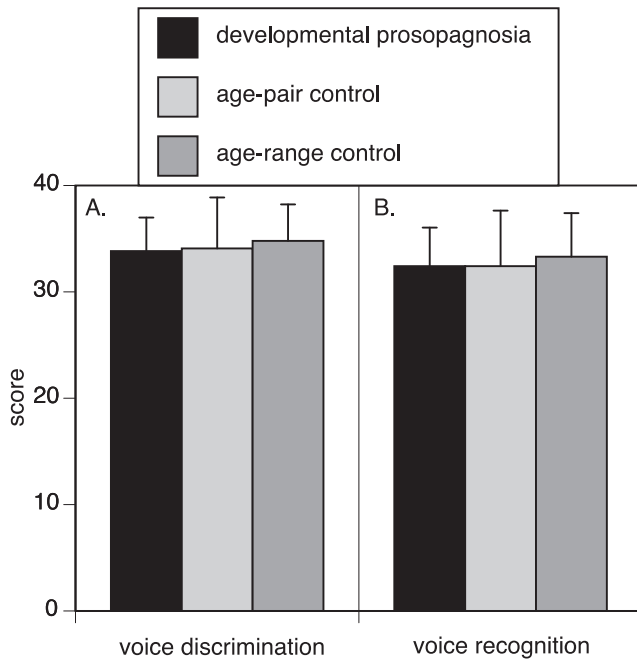


Fig. 2 – Age-matched group analyses. A. voice discrimination, B. voice recognition. The prosopagnosia group is compared with the 12 control subjects closest in age (age-pair controls), and with all control subjects falling in the range of 27–67 years (age-range controls). Group means are shown, with error bars showing one standard deviation.

analysis (Fig. 2) showed no difference between the prosopagnosic group and either age-paired [$t_{(11)} = .00, p = 1.00$] or age-range controls [$t_{(21)} = .71, p = .48$].

To compare the severity of short-term familiarity deficits for voices and faces, we correlated the z-scores on the voice recognition and Cambridge Face Memory tests (Fig. 3). Subject DP035 has equivalent deficits for both faces and voices, but as a group, developmental prosopagnosic subjects were worse on face than voice recognition [$t_{(11)} = 9.11, p < .0001$].

3.3. Questionnaire

All but two subjects with developmental prosopagnosia reported impairment on the questionnaire's face recognition portion: DP029 and DP008 had borderline scores (Table 2). Only one subject (DP021) reported difficult voice recognition. However, this subject scored close to control age-adjusted mean score for both voice discrimination and recognition (Fig. 1). Of note, DP035 was unaware of problems with voice recognition.

4. Discussion

Most of our subjects with developmental prosopagnosia had intact voice discrimination and recognition for newly-learned voices, indicating a modality-specific visual agnosia rather than a multi-modal person recognition syndrome. On the other hand, one of 12 subjects had borderline results that were

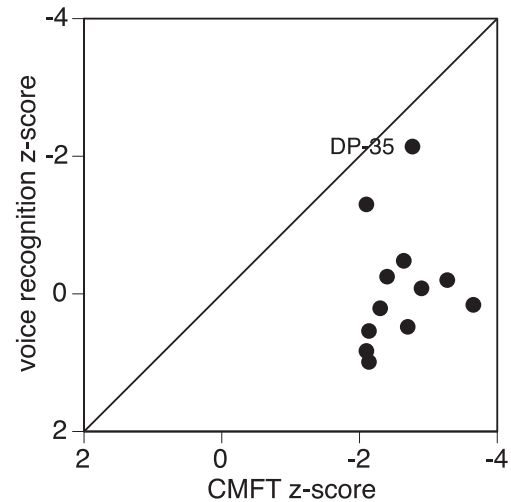


Fig. 3 – Voice versus face recognition in developmental prosopagnosia. Z-scores for the voice recognition test are plotted against z-scores for the Cambridge Face Memory Test (CFMT), with negative z-scores indicating impairment. Subject DP035 has equivalent deficits for face and voice recognition, but other subjects are more impaired for face recognition.

not definitive for either normality or abnormality. Nevertheless, the fact that he scored low on both voice tests, and that his voice recognition deficit was of similar severity to his face recognition deficit, are consistent with a plausible impairment, which could indicate that a subset of developmental prosopagnosic subjects have voice recognition deficits. This is reminiscent of the report of impaired voice recognition in subject SO (von Kriegstein et al., 2006), although the parallel is not exact, since SO was impaired not on a similar test of recognition of newly learned voices, but on recognition of voices from daily life. Compared to our acquired prosopagnosic subjects (Liu et al., 2014), DP035 is most functionally similar to B-ATOT2, a subject with bilateral fusiform and

Table 2 – Voice and face processing questionnaire, results.

| Questionnaire results (score 0–35). | | |
|-------------------------------------|-----------|-----------|
| | Faces | Voices |
| control mean | 10.16 | 13.14 |
| SD. | 3.88 | 4.10 |
| 95% upper limit | 19.39 | 22.89 |
| DP003 | <u>20</u> | 9 |
| DP008 | 18 | 12 |
| DP014 | <u>27</u> | 7 |
| DP016 | <u>22</u> | 13 |
| DP021 | <u>31</u> | <u>27</u> |
| DP024 | <u>24</u> | 8 |
| DP029 | 18 | 8 |
| DP032 | <u>28</u> | 12 |
| DP033 | <u>26</u> | 18 |
| DP035 | <u>32</u> | 14 |
| DP038 | <u>22</u> | 21 |
| DP044 | <u>23</u> | 17 |

Bold underline indicates abnormal scores.

anterior temporal lesions, in that he is impaired not only on both voice and face recognition tests, but also on voice and face discrimination tests. This suggests the co-occurrence of voice and face apperceptive defects, rather than a single amodal recognition deficit.

The potential for voice recognition to be affected in developmental prosopagnosia rests on several grounds. First, some studies have suggested that the fusiform region is anomalous in developmental prosopagnosia (Furl, Garrido, Dolan, Driver, & Duchaine, 2011; Garrido, Furl, et al., 2009), and neuroimaging studies indicate that the fusiform face area is sensitive to voice processing (von Kriegstein et al., 2006; von Kriegstein, Kleinschmidt, Sterzer, & Giraud, 2005; Schweinberger, Kloth, & Robertson, 2011) and has connections with voice-sensitive areas in the superior temporal sulcus (Blank, Anwender, & von Kriegstein, 2011). Against this, our study of acquired prosopagnosia failed to find any impairment of voice discrimination or recognition of recently heard voices when lesions were limited to the fusiform gyri (Liu et al., 2014). Similarly, voice recognition was intact in subject MS, an acquired prosopagnosic subject who had damage to bilateral fusiform gyri and the right anterior temporal lobe (Arnott, Heywood, Kentridge, & Goodale, 2008) and even superior to controls in subject SB, who had extensive right-sided damage including the fusiform gyrus and less extensive left occipitotemporal injury (Hoover, Demonet, & Steeves, 2010).

A second scenario focuses on the anterior temporal lobe. Investigators noting normal activation of the fusiform face area in developmental prosopagnosia (Avidan & Behrmann, 2009; Avidan, Hasson, Malach, & Behrmann, 2005) have performed morphometric studies reporting abnormalities in anterior fusiform cortex and the anterior and middle temporal lobes (Behrmann, Avidan, Gao, & Black, 2007). Functional neuroimaging has revealed temporal voice areas in the middle and anterior superior temporal sulcus, more so on the right (Belin, Zatorre, Lafaille, Ahad, & Pike, 2000), which show sensitivity to voice identity in adaptation studies (Belin & Zatorre, 2003; Warren, Scott, Price, & Griffiths, 2006) and activity correlations with the familiarity of voices (Bethmann, Scheich, & Brechmann, 2012). It must be noted, though, that the anterior temporal regions responding to facial identity are more ventral than temporal voice areas (Belin et al., 2000) and do not overlap (Joassin et al., 2011). Nevertheless, if anterior or middle temporal anomalies underlie developmental prosopagnosia, these could theoretically disrupt voice processing also. However, our previous study found that impairments of both face and voice recognition were more likely to occur with bilateral than just right anterior temporal damage (Liu et al., 2014).

Both of these last points are consistent with the literature on multi-modal person recognition syndromes. First, of the six cases with documented impaired familiarity for both voices and faces with preserved discrimination of these stimuli, all four with lesion information had anterior temporal lesions, including QR and KL (Hailstone, Crutch, Vestergaard, Patterson, & Warren, 2010), as well as B-AT1 and B-AT2 (Liu et al., 2014). Such information was not available for case 13 and 37 (Neuner & Schweinberger, 2000). Second, lesions were either bilateral, as in B-AT1, B-AT2, and case 13 (Liu et al., 2014;

Neuner & Schweinberger, 2000), or if predominantly right-sided as in QR and KL (Hailstone et al., 2010) had atrophic degeneration, in which damage is diffuse and highly likely to involve the left hemisphere to some degree. Case 13 of (Neuner & Schweinberger, 2000) had a right-sided lesion, but their pathology was not given. Even if we extend this list to include subjects in whom the data on face or voice discrimination is insufficient to exclude apperceptive deficits – e.g., Emma (Gentileschi, Sperber, & Spinnler, 2001), Maria (Gentileschi, Sperber, & Spinnler, 1999), CD (Gainotti, Ferraccioli, Quaranta, & Marra, 2008) and case 31 (Neuner & Schweinberger, 2000) – these two points still hold: these have mainly anterior temporal damage with either bilateral lesions or right dominant lesions with etiologies that cannot exclude a degree of bilaterality. Third, in functional neuroimaging of healthy subjects, temporal voice areas are found in the middle and anterior superior temporal sulci, and while these show right dominance, they are found bilaterally (Belin et al., 2000). Likewise, fMRI adaptation studies reveal sensitivity to voice identity in the right anterior superior temporal sulcus in some studies (Belin & Zatorre, 2003) but in bilateral posterior superior temporal sulci in others (Warren et al., 2006).

At this point it is not definitive whether the pathology in developmental prosopagnosia is focal or diffuse, unilateral or bilateral, anterior temporal or fusiform, homogenous or heterogeneous. Do the behavioral data help? First, although it was not our focus, the fact that many but not all of our subjects performed well when discriminating between faces on the CFPT suggests that most have a form of anterior temporal rather than fusiform dysfunction. This inference aligns with the conclusions of some morphometric studies suggesting abnormalities in more anterior aspects of inferotemporal cortex (Behrmann et al., 2007; Garrido, Furl, et al., 2009). Second, the preservation of voice recognition suggests a resemblance to acquired prosopagnosia with right anterior temporal rather than bilateral lesions. Some structural studies of developmental prosopagnosia have found temporal abnormalities only in the right hemisphere (Garrido, Furl, et al., 2009), while others failed to find a hemispheric asymmetry (Behrmann et al., 2007; Thomas et al., 2009). All told, though, observations by us and others (Stollhoff et al., 2011) that some developmental prosopagnosia subjects show face discrimination deficits while others do not, and some have voice recognition deficits (von Kriegstein et al., 2006) while others do not, suggests a spectrum of structural and functional anomalies in developmental prosopagnosia.

Finally, we close with a discussion of two technical points. First, our data underline the value of formal perceptual testing. This point has been made regarding face recognition, that subjective impressions or questionnaires are not always corroborated by formal face processing tests (De Haan, 1999). Similarly, although most prosopagnosic subjects claim to rely on voice to identify others, objective tests have revealed asymptomatic voice recognition deficits in some (Boudouresques et al., 1979; Liu et al., 2014). Here we found one subject with an asymptomatic impairment and another who performed well despite claiming poor voice recognition. Thus objective tests rather than subjective reports are required to clarify if a face-recognition impairment is a modality-specific deficit or part of a multi-modal person recognition syndrome,

to situate the deficit correctly in the hierarchy of cognitive operations involved in recognizing people.

Second, our test of voice recognition involved recently learned voices, as is also used in some other studies (Garrido, Eisner, et al., 2009; Hoover et al., 2010; Roswadowitz et al., 2014), rather than famous or personally familiar voices. These almost certainly have behavioral and neural differences (Blank, Wieland, & von Kriegstein, 2014), given the richer emotional, episodic and semantic associations of known faces and voices. Nevertheless, there are pragmatic advantages to using recently familiarized rather than personally known voices in testing. For one, the degree of exposure to famous face and voice stimuli will vary widely across subjects for individual and cultural reasons, whereas exposure to faces made familiar during testing can be controlled and made uniform in a cohort. Personally familiar people such as relatives and friends can be used in a single case study, but it is impossible to do so in a fashion that is equivalent across subjects and controls in a cohort study. For these reasons, the diagnosis of developmental prosopagnosia has come to rest on two things: first, self-report of lifelong impairment in face recognition and second, impaired performance on a test of face memory, most often the Cambridge Face Memory Test (Duchaine & Nakayama, 2006), which assesses familiarity of recently viewed faces (e.g., DeGutis, Cohan, & Nakayama, 2014; Shah, Gaule, Gaigg, Bird, & Cook, 2015). In many cases, studies also require impairment on an additional test of face processing which could be a test of face discrimination, configural face processing, or face familiarity (famous or newly learned; e.g., Avidan, Tanzer, Hadj-Bouziane, Liu, Ungerleider, 2014; Johnen et al., 2014). Importantly, in either case, it is typically not a requirement that subjects show impairment on a test of famous or personally familiar face memory. With far fewer published cases and fewer objective tests, the diagnostic protocol for developmental phonagnosia is less established. Yet, a recent study used a similar approach for developmental phonagnosia, with inclusion criteria resting on reported lifelong impairment in voice recognition and impaired performance on a test of newly learned voices, but not famous voices (Roswadowitz et al., 2014).

Furthermore, while we and others still test famous face recognition despite the above reservations, famous voice recognition is notoriously poor in healthy subjects. On one test of famous voice recognition controls could name only about 50% of 79 voices (Meudell, Northen, Snowden, & Neary, 1980), and on another only 18 of 96 voices (Garrido, Eisner, et al., 2009). In another study (Hailstone et al., 2010), control subjects on a 48-item test of familiarity for famous voices scored as low as 29/48, which by binomial proportions is at the upper limit of chance performance.

On a conceptual level, the use of recognition tests for recently viewed faces reflects the fact that prosopagnosia is not just difficulty recognizing long-known faces, but also the inability to learn to recognize new faces (Damasio et al., 1990). In fact, while anterograde prosopagnosia exists, in which subjects cannot learn to recognize new faces but can still recognize long familiar faces (Tranel & Damasio, 1985; Young et al., 1995), we are not aware of any subject with the converse pattern. In this respect, the behavioral and neural differences between recently learned and long familiar faces are

immaterial for the diagnosis of prosopagnosia for most subjects. For all of these reasons, and until there is evidence to the contrary, we believe that the testing of familiarity for recently learned voices and faces is a justifiable and possibly superior means of assessing for the presence of phonagnosia and prosopagnosia respectively.

Acknowledgments

This work was supported by a CIHR operating grant (MOP-102567) to JB. JB was supported by a Canada Research Chair and the Marianne Koerner Chair in Brain Diseases. BD was supported by grants from the Economic and Social Research Council (UK) (RES-062-23-2426) and the Hitchcock Foundation. SC was supported by National Eye Institute of the National Institutes of Health under award number F32 EY023479-02. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

REFERENCES

- Arnott, S. R., Heywood, C. A., Kentridge, R. W., & Goodale, M. A. (2008). Voice recognition and the posterior cingulate: an fMRI study of prosopagnosia. *Journal of Neuropsychology*, 2(Pt 1), 269–286.
- Avidan, G., & Behrmann, M. (2009). Functional MRI reveals compromised neural integrity of the face processing network in congenital prosopagnosia. *Current Biology*, 19(13), 1146–1150.
- Avidan, G., Hasson, U., Malach, R., & Behrmann, M. (2005). Detailed exploration of face-related processing in congenital prosopagnosia: 1. Functional neuroimaging findings. *Journal of Cognitive Neuroscience*, 17(7), 1150–1167.
- Avidan, G., Tanzer, M., Hadj-Bouziane, F., Liu, N., Ungerleider, L. G., & Behrmann, M. (2014). Selective dissociation between core and extended regions of the face processing network in congenital prosopagnosia. *Cerebral Cortex*, 24, 1565–1578.
- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (aq): evidence from asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, 31(1), 5–17.
- Barton, J. J. (2008). Structure and function in acquired prosopagnosia: lessons from a series of 10 patients with brain damage. *Journal of Neuropsychology*, 2(Pt 1), 197–225.
- Barton, J., Cherkasova, M., Press, D., Intriligator, J., & O'Connor, M. (2003). Developmental prosopagnosia: a study of three patients. *Brain and Cognition*, 51, 12–30.
- Bate, S., Cook, S. J., Duchaine, B., Tree, J. J., Burns, E. J., & Hodgson, T. L. (2014). Intranasal inhalation of oxytocin improves face processing in developmental prosopagnosia. *Cortex*, 50, 55–63.
- Behrmann, M., & Avidan, G. (2005). Congenital prosopagnosia: face-blind from birth. *Trends in Cognitive Sciences*, 9(4), 180–187.
- Behrmann, M., Avidan, G., Gao, F., & Black, S. (2007). Structural imaging reveals anatomical alterations in inferotemporal cortex in congenital prosopagnosia. *Cerebral Cortex*, 17(10), 2354–2363.
- Belin, P., & Zatorre, R. J. (2003). Adaptation to speaker's voice in right anterior temporal lobe. *NeuroReport*, 14(16), 2105–2109.

- Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature*, 403(6767), 309–312.
- 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.
- Blank, H., Anwender, A., & von Kriegstein, K. (2011). Direct structural connections between voice- and face-recognition areas. *Journal of Neuroscience*, 31(36), 12906–12915.
- Blank, H., Wieland, N., & von Kriegstein, K. (2014). Person recognition and the brain: merging evidence from patients and healthy individuals. *Neuroscience & Biobehavioral Reviews*, 47, 717–734.
- Boudouresques, J., Poncet, M., Cherif, A. A., & Balzamo, M. (1979). Agnosia for faces: evidence of functional disorganization of a certain type of recognition of objects in the physical world. *Bulletin de l'Académie Nationale de Médecine*, 163(7), 695–702.
- Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, 77, 305–327.
- Bulthoff, I., & Newell, F. N. (2015). Distinctive voices enhance the visual recognition of unfamiliar faces. *Cognition*, 137, 9–21.
- Dalrymple, K. A., Garrido, L., & Duchaine, B. (2014). Dissociation between face perception and face memory in adults, but not children, with developmental prosopagnosia. *Developmental Cognitive Neuroscience*, 10, 10–20.
- Damasio, A. R., Tranel, D., & Damasio, H. (1990). Face agnosia and the neural substrates of memory. *Annual Review Neuroscience*, 13, 89–109.
- Davies-Thompson, J., Pancaroglu, R., & Barton, J. (2014). Acquired prosopagnosia: structural basis and processing impairments. *Frontiers in Bioscience (Elite Ed)*, 6, 159–174.
- De Haan, E. H. (1999). A familial factor in the development of face recognition deficits. *Journal of Clinical and Experimental Neuropsychology*, 21(3), 312–315.
- DeGutis, J., Cohan, S., & Nakayama, K. (2014). Holistic face training enhances face processing in developmental prosopagnosia. *Brain*, 137(Pt 6), 1781–1798.
- Duchaine, B., & 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 participants. *Neuropsychologia*, 44(4), 576–585.
- Duchaine, B., Yovel, G., & Nakayama, K. (2007). No global processing deficit in the Navon task in 14 developmental prosopagnosics. *Social Cognitive and Affective Neuroscience*, 2(2), 104–113.
- Furl, N., Garrido, L., Dolan, R. J., Driver, J., & Duchaine, B. (2011). Fusiform gyrus face selectivity relates to individual differences in facial recognition ability. *Journal of Cognitive Neuroscience*, 23(7), 1723–1740.
- 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., 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(3), 238–248.
- Garrido, L., Eisner, F., McGettigan, C., Stewart, L., Sauter, D., Hanley, J. R., et al. (2009). Developmental phonagnosia: a selective deficit of vocal identity recognition. *Neuropsychologia*, 47(1), 123–131.
- Garrido, L., Furl, N., Draganski, B., Weiskopf, N., Stevens, J., Tan, G. C., et al. (2009). Voxel-based morphometry reveals reduced grey matter volume in the temporal cortex of developmental prosopagnosics. *Brain*, 132(Pt 12), 3443–3455.
- Gentileschi, V., Sperber, S., & Spinnler, H. (1999). Progressive defective recognition of people. *Neurocase*, 5, 407–424.
- Gentileschi, V., Sperber, S., & Spinnler, H. (2001). Crossmodal agnosia for familiar people as a consequence of right infero-polar temporal atrophy. *Cognitive Neuropsychology*, 18(5), 439–463.
- 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), 1104–1114.
- Hoover, A. E., Demonet, J. F., & Steeves, J. K. (2010). Superior voice recognition in a patient with acquired prosopagnosia and object agnosia. *Neuropsychologia*, 48(13), 3725–3732.
- Joassin, F., Pesenti, M., Maurage, P., Verreclt, E., Bruyer, R., & Campanella, S. (2011). Cross-modal interactions between human faces and voices involved in person recognition. *Cortex*, 47(3), 367–376.
- Johnen, A., Schmukle, S. C., Huttenbrink, J., Kischka, C., Kennerknecht, I., & Döbel, C. (2014). A family at risk: congenital prosopagnosia, poor face recognition and visuo-perceptual deficits within one family. *Neuropsychologia*, 58, 52–63.
- Kracke, I. (1994). Developmental prosopagnosia in asperger syndrome: presentation and discussion of an individual case. *Developmental Medicine and Child Neurology*, 36, 873–886.
- von Kriegstein, K., Kleinschmidt, A., & Giraud, A. L. (2006). Voice recognition and cross-modal responses to familiar speakers' voices in prosopagnosia. *Cerebral Cortex*, 16(9), 1314–1322.
- 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(3), 367–376.
- Liu, R. R., Pancaroglu, R., Hills, C. S., Duchaine, B., & Barton, J. J. (2014). Voice recognition in face-blind patients. *Cerebral Cortex*.
- Meudell, P. R., Northen, B., Snowden, J. S., & Neary, D. (1980). Long term memory for famous voices in amnesic and normal subjects. *Neuropsychologia*, 18(2), 133–139.
- Neuner, F., & Schweinberger, S. R. (2000). Neuropsychological impairments in the recognition of faces, voices, and personal names. *Brain and Cognition*, 44(3), 342–366.
- Roswandowitz, C., Mathias, S. R., Hintz, F., Kreitewolf, J., Schelinski, S., & von Kriegstein, K. (2014). Two cases of selective developmental voice-recognition impairments. *Current Biology*, 24, 2348–2353.
- Schweinberger, S. R., Kloth, N., & Robertson, D. M. (2011). Hearing facial identities: brain correlates of face–voice integration in person identification. *Cortex*, 47(9), 1026–1037.
- Shah, P., Gaule, A., Gaigg, S. B., Bird, G., & Cook, R. (2015). Probing short-term face memory in developmental prosopagnosia. *Cortex*, 64, 115–122.
- Stollhoff, R., Jost, J., Elze, T., & Kennerknecht, I. (2011). Deficits in long-term recognition memory reveal dissociated subtypes in congenital prosopagnosia. *PLoS One*, 6(1), e15702.
- Susilo, T., & Duchaine, B. (2013). Advances in developmental prosopagnosia research. *Current Opinion in Neurobiology*, 23(3), 423–429.
- Thomas, C., Avidan, G., Humphreys, K., Jung, K. J., Gao, F., & Behrmann, M. (2009). Reduced structural connectivity in ventral visual cortex in congenital prosopagnosia. *Nature Neuroscience*, 12(1), 29–31.
- Tranel, D., & Damasio, A. (1985). Knowledge without awareness: an autonomic index of facial recognition by prosopagnosics. *Science*, 228, 1453–1454.
- Warren, J. D., Scott, S. K., Price, C. J., & Griffiths, T. D. (2006). Human brain mechanisms for the early analysis of voices. *NeuroImage*, 31(3), 1389–1397.
- Warrington, E. (1984). *Warrington recognition memory test*. Los Angeles: Western Psychological Services.
- Young, A., Aggleton, J., Hellawell, D., Johnson, M., Brooks, P., & Hanley, J. (1995). Face processing impairments after amygdalotomy. *Brain*, 118, 15–24.