

## Developmental phonagnosia: A selective deficit of vocal identity recognition

Lúcia Garrido<sup>a,b,\*</sup>, Frank Eisner<sup>a,c</sup>, Carolyn McGettigan<sup>a,c</sup>, Lauren Stewart<sup>d</sup>, Disa Sauter<sup>e</sup>, J. Richard Hanley<sup>f</sup>, Stefan R. Schweinberger<sup>g</sup>, Jason D. Warren<sup>h</sup>, Brad Duchaine<sup>a,b</sup>

<sup>a</sup> Institute of Cognitive Neuroscience, University College London, UK

<sup>b</sup> Department of Psychology, University College London, UK

<sup>c</sup> Department of Human Communication Science, University College London, UK

<sup>d</sup> Department of Psychology, Goldsmiths College, London, UK

<sup>e</sup> Brain and Behaviour Laboratory, Birkbeck College, UK

<sup>f</sup> Department of Psychology, University of Essex, UK

<sup>g</sup> Institute of Psychology, Friedrich-Schiller-University of Jena, Germany

<sup>h</sup> Dementia Research Centre, Institute of Neurology, University College London, UK

### ARTICLE INFO

#### Article history:

Received 4 June 2008

Received in revised form 1 August 2008

Accepted 6 August 2008

Available online 13 August 2008

#### Keywords:

Phonagnosia

Voice recognition

Voice perception

Developmental disorders

### ABSTRACT

Phonagnosia, the inability to recognize familiar voices, has been studied in brain-damaged patients but no cases due to developmental problems have been reported. Here we describe the case of KH, a 60-year-old active professional woman who reports that she has always experienced severe voice recognition difficulties. Her hearing abilities are normal, and an MRI scan showed no evidence of brain damage in regions associated with voice or auditory perception. To better understand her condition and to assess models of voice and high-level auditory processing, we tested KH on behavioural tasks measuring voice recognition, recognition of vocal emotions, face recognition, speech perception, and processing of environmental sounds and music. KH was impaired on tasks requiring the recognition of famous voices and the learning and recognition of new voices. In contrast, she performed well on nearly all other tasks. Her case is the first report of developmental phonagnosia, and the results suggest that the recognition of a speaker's vocal identity depends on separable mechanisms from those used to recognize other information from the voice or non-vocal auditory stimuli.

© 2008 Elsevier Ltd. All rights reserved.

## 1. Introduction

KH finally felt she understood her lifelong social problem while reading an article in a popular scientific magazine. The article discussed prosopagnosia, a condition characterized by severe face recognition difficulties. Prosopagnosics fail to recognize the faces of co-workers, friends and relatives. Some individuals experience these problems throughout their lives in the absence of neurological damage, a condition called developmental or congenital prosopagnosia (e.g. Behrmann & Avidan, 2005; Bornstein, 1963; Duchaine & Nakayama, 2006a; Kress & Daum, 2003; McConachie, 1976). KH, however, has very good face recognition. Her problems with person recognition involve voices rather than faces, and KH realized she might have a vocal analogue of prosopagnosia.

### 1.1. Phonagnosia

The term 'phonagnosia' was first proposed by Van Lancker and Canter (1982) to refer to disorders of familiar voice recognition. In a series of studies looking at voice perception in either right or left hemisphere brain damaged patients, Van Lancker and colleagues found that recognition of famous voices was impaired in many patients with right hemisphere lesions, usually involving the right parietal lobe (Van Lancker & Canter, 1982; Van Lancker, Cummings, Kreiman, & Dobkin, 1988; Van Lancker & Kreiman, 1987; Van Lancker, Kreiman, & Cummings, 1989). In addition, they found that same-different discrimination between two voice samples was impaired by lesions to either hemisphere, often involving the temporal lobes.

Other studies have assessed whether voice recognition impairments are restricted to voices or extend to other auditory stimuli in an effort to understand the organization of voice and high-level auditory processing mechanisms. Up to the date of their review, Assal, Aubert, and Buttet (1981) found 11 patients in the neurological literature whose voice recognition abilities were impaired following brain damage. Those impairments were always associ-

\* Corresponding author at: Institute of Cognitive Neuroscience, Alexandra House, 17 Queen Square, WC1N 3AR London, UK. Tel.: +44 20 7679 8544.

E-mail address: [m.garrido@ucl.ac.uk](mailto:m.garrido@ucl.ac.uk) (L. Garrido).

ated with other disorders of auditory processing, such as amusia, pure word deafness or auditory agnosia. For example, Assal et al. (1981) reported the case of the patient RB who, following a cerebral vascular accident, presented a global auditory agnosia. After nine months, language impairments and difficulties in recognizing sounds seemed to have disappeared. Nevertheless, RB continued to show impairments in music tasks in addition to voice recognition difficulties (in recognizing familiar voices, discriminating between speakers, and perceiving different intonations in speech). This case suggests that voice recognition impairments can co-occur with normal sound recognition. Similarly, Neuner and Schweinberger (2000) have more recently presented the cases of four patients with voice recognition impairments who performed normally with the recognition of environmental sounds. Patient RB (Assal et al., 1981), though, showed impairments in recognizing music at all times of testing and there are no reports of phonagnosia with normal music processing (note, however, that this has rarely been examined). Peretz et al. (1994) presented two amusic patients who also performed below the normal range on voice recognition tasks, and suggested that certain tasks, such as the recognition of musical instruments, could depend on similar mechanisms as voice discrimination as both depend on timbre.

### 1.2. Models of voice recognition

The study of dissociations of voice processing abilities in patients has been crucial in formulating models of voice recognition. These models have been strongly influenced by models of face processing, and important parallels between the two abilities have been suggested. Ellis, Jones, and Mosdell (1997) based their proposal on the face recognition model by Burton, Bruce, and Johnston (1990). Ellis et al. (1997) suggested that vocal feature information is first encoded at a very basic level (*auditory structural encoding*), after which it is further processed and aggregated in *voice recognition units*. The links between these units and the *person identity nodes* (multi-modal nodes) contribute to familiarity decisions. Finally, semantic information (from the *semantic information units*) is associated with the stimulus.

In a recent review, Belin, Fecteau, and Bédard (2004) proposed a model that not only considers how the recognition of speaker is achieved, but also how other vocal information is processed. This model is largely based on Bruce and Young (1986) influential face recognition model. In their view, voices are first analyzed at a generic low-level, common to other auditory stimuli, and then *structurally encoded*. From this stage, voice information is further analysed in parallel pathways that separately process *speech information*, *vocal affective information* and *vocal identity information*. Because it involves parallel mechanisms, this model predicts functional dissociations between these three types of processing and some cases of phonagnosia with normal perception of speech information or vocal affective information. Consistent with this prediction, RB appears to have had preserved speech perception despite his difficulties with voice recognition (Assal et al., 1981). However, that study provided insufficient detail about the testing so evidence for this dissociation is weak. In addition, no functional dissociations between speaker recognition and the perception of vocal affective information have been reported.

### 1.3. KH

KH, a 60-year-old woman who works as a management consultant reports that she has always experienced severe voice recognition difficulties. She has great difficulty recognizing people who call her on the phone, even if those people are close relatives such as her daughter. KH avoids answering the phone, and for many

years she has only answered 'booked calls'. She books calls with friends or co-workers, so she knows who to expect when the telephone rings at a certain time. In a telling anecdote, KH reported that, in the 1980s, she had a job in which she introduced herself with a different form of her first name so she would know that it was someone related to her job when they called and asked for her using that name.

Our first aim in this study was to establish whether KH had an impairment with voice recognition under laboratory conditions. Demonstrating that KH has such problems when compared to age-matched controls would indicate that KH is the first reported case of developmental phonagnosia. Belin et al. (2004) specifically considered the possibility of the existence of this condition. Our second aim was to shed light on the organization of voice and auditory processing by examining how specific KH's deficit is by testing her ability to perceive other information from the voice such as emotion and speech information and other high-level auditory stimuli.

## 2. Methods and results

### 2.1. Participants

#### 2.1.1. Case KH

KH is a right-handed woman with a master's degree and an uncompleted Ph.D. (20 years of formal education). English is her native language. She reports that she suffered from perinatal asphyxia. KH also reports poor sense of direction and mild dyslexia, which were never assessed. KH is medicated with fluoxetine.<sup>1</sup>

**2.1.1.1. Structural MRI.** We carried out a whole brain structural MRI scan on KH on a 3T Trio whole body scanner (Siemens Medical Systems, Erlangen, Germany). Two sequences were acquired, one used for T1 weighting and one for T2 weighting. A neurologist (JW) inspected the images for gross abnormalities. The scans showed only microvascular changes in the basal ganglia and cerebral white matter which were considered mild and in keeping with KH's age, and thus not considered relevant for the present findings.

**2.1.1.2. Audiogram.** Pure tone audiometry showed that KH has normal peripheral hearing. For all frequencies tested (250, 500, 1000, 2000, 4000 and 8000 Hz), and for both ears, her mean hearing level was less than or equal to 20 decibels.

**2.1.1.3. Neuropsychological testing.** KH has a normal forward digit span of 6 (Lezak, 1995). She scored 10 on the Raven's Advanced Progressive Matrices (Raven, Court, & Raven, 1981), a measure of abstract reasoning, which places her at the 95th percentile in her age group. To investigate her auditory memory using verbal material, KH was given the Rey Auditory Verbal Learning Test (Rey, 1964). This test consists of a learning phase in which List A of 15 words is said five times. Immediately after each presentation, the participant is asked to recall as many words as possible from that list. Following the fifth presentation of List A, List B of 15 different words is read to the participant, who is asked to recall as many words from this list as possible. After recalling List B, the participant is asked to recall List A again. Twenty minutes after this, the participant is asked once more to recall as many words as possible from List A. KH correctly recalled 72 words during the learning phase of List A (she recalled 12 words after the first presentation and all 15 words following each of the other four presentations). After interference from List B, KH correctly recalled all the words from List A, and the same occurred after a 20 min delay. All these results are within the range of normal performance (Lezak, 1995).

<sup>1</sup> Fluoxetine is a widely used serotonin reuptake inhibitor. A recent study found no difference between depressed patients treated with fluoxetine and those taking placebo on a range of neuropsychological measures (Strik, Honig, Klinkenberg, Dijkstra, & Jolles, 2006). Another study (Gopal, Briley, Goodale, & Hendea, 2005) found that when a group of depressed patients was taking serotonin reuptake inhibitors, they performed significantly worse on a few low-level auditory tests than when they were not medicated. These results suggest that some early auditory processing may be affected by fluoxetine, but we would expect this to have an impact on all auditory tasks we used. Conversely, as we will see, KH's impairment is very selective and she performed well on several high level auditory tasks. Moreover, KH reports lifelong difficulties in recognizing voices and therefore these precede her taking fluoxetine. Finally, we note that one of the matched controls was also medicated with fluoxetine and she performed well on all tasks.

### 2.1.2. Control participants

In all experimental tasks described below, we tested KH and an age-matched control group. It consisted of eight women between 46 and 64 years old (mean = 56.63; S.D. = 5.50). Their mean years of education was 15.88 (S.D. = 1.96). All were native British English speakers. Six were right-handed and two were left-handed. All reported normal hearing. One control (control 6) was taking fluoxetine, as well as ropinirole for treatment of restless leg syndrome. The other controls reported no neurological or psychiatric history.

## 2.2. Material, procedure and results

Participants were tested on the tasks described below in three different sessions. Unfortunately control 1 was not able to participate in the third session. Unless otherwise specified, stimuli were presented and responses were collected in Cogent 2000 (Cogent 2000 team—<http://www.vislab.ucl.ac.uk/Cogent2000>) running on MATLAB 7.3.0 (R2006b) (Mathworks—<http://www.mathworks.com>). Auditory stimuli were presented through headphones. The volume was set to be the same for all participants using QuickMix 1.06 (Product Technology Partners—<http://www.ptpart.co.uk/quickmix>) (i.e. this program preserved the same intensity for all participants).

KH's scores on each task were compared to controls' results using the modified *t*-test for single case studies developed by Crawford and Howell (1998). Differences between KH and control participants were considered significant when the one-tailed probability was equal to or below 0.05. Only significantly different results are reported.

### 2.2.1. Voice recognition

The following tasks investigated memory for voices. The first tested recognition of famous voices, and then a series of tasks assessed recognition of unfamiliar speakers.

**2.2.1.1. Recognition of famous voices. Material and procedure:** In this task (created by Damjanovic & Hanley, 2007), participants were presented with 96 voice samples taken from television interviews, each of which lasted approximately seven seconds. Half the samples were from famous people (familiar to UK residents) and half from non-famous people. For each trial, participants decided whether they recognized the voice or not. Each sample was presented using iTunes (Apple—<http://www.apple.com/itunes>).

If participants reported familiarity with a voice, they were asked to name the person or provide any other specific information about them. However, it could be that some participants had previously heard the voices of the people on the test very often, whereas others had not. To assess exposure to the famous voices, we asked each participant after the task whether they thought they had heard the voices of the unidentified famous people enough in everyday life to recognize them.

**Results:** Two control participants (C3 and C8) did not perform this task because they reported that they had not had a television for over 30 years and did not watch films. For the six controls who did the task, we calculated *A'* (McMillan & Creelman, 1990) as a measure of discrimination between famous and non-famous people. They were all able to successfully discriminate famous from non-famous voices, showing a mean *A'* of 0.80 (S.D. = 0.06) (see Fig. 1A). KH's *A'* was 0.62, which was significantly lower than controls ( $t(5) = -2.78, p = 0.020$ ). She correctly discriminated famous vs non-famous voices on only 54 of the 96 trials, which is not significantly different from chance ( $z = 1.12, p = 0.13$ ).

Controls were able to name or provide specific information for an average of 17.67 of the voices (S.D. = 4.72). For example, nearly all controls identified the voices of Margaret Thatcher, David Beckham, Dawn French, Chris Tarrant, Joanna Lumley, Sean Connery and Ann Widdecombe. In contrast, KH identified only one person's

voice (Sean Connery), and her score is far below and significantly different from the control mean ( $t(5) = -3.27, p = 0.011$ ). Controls correctly identified an average of 47.40% (S.D. = 9.81) of the voices to which they were exposed. In contrast, KH correctly identified only 3.85% of the voices to which she reported she was highly exposed. This percentage is also significantly different from controls ( $t(5) = -4.11, p = 0.005$ ) (see Fig. 1B).

### 2.2.1.2. Learning and recognition of six different speakers.

**2.2.1.2.1. Learning phase. Material and procedure:** Participants were asked to learn the voices of six unfamiliar young female speakers. All were native British English speakers and had similar accents. Speakers read sentences with three key words taken from the BKB Sentence List (Bench, Kowal, & Bamford, 1979). All samples were recorded in an anechoic chamber using Cool Edit 96 (Syntrillium Software Corporation—<http://www.syntrillium.com>). The stimuli were normalised for peak amplitude using the programme PRAAT (Boersma & Weenink, 2005).

Participants were asked to learn name-voice pairings, though they were told that the knowledge of the names would only be necessary for a later task. They were first presented with the name of a speaker, each of which started with a different letter from A to F, and then heard a sentence said by that speaker. After that, they heard a number of sentences and for each one they decided if it was said by the same speaker or not. Half the sentences were said by the target speaker, while the other half were said by different speakers. This procedure was repeated for each of the six speakers.

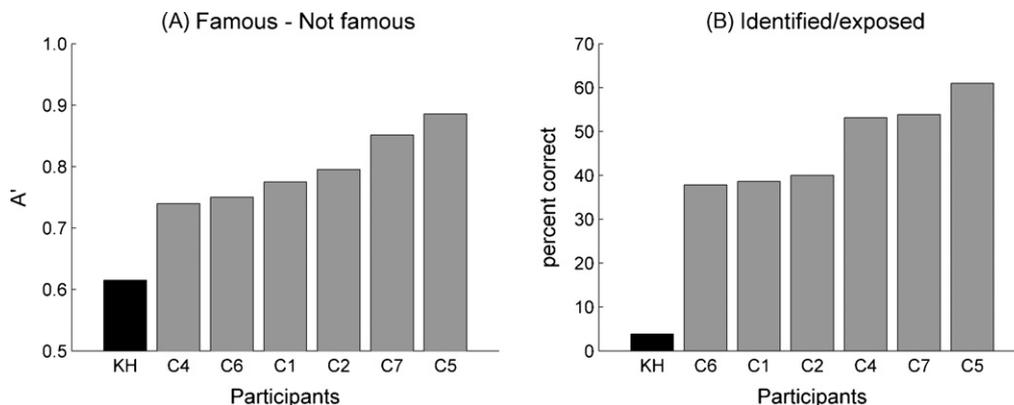
There were four blocks. In the practice block, participants were presented with two test sentences after attempting to memorize the voice of each of the target speakers. In the first test block, six test sentences followed the sentence presenting each speaker's voice. In the second and third test blocks, there were ten test sentences. No sentences were repeated.

**Results:** Controls responded correctly on 71.53% (S.D. = 4.13) of trials in the first block, 74.38% (S.D. = 5.27) in the second block, and 78.13% (S.D. = 5.45) in the third block. KH scored 72.22%, 63.33% and 70% in each block respectively (see Fig. 2A). Only the result on the second block was significantly different from controls' performance ( $t(7) = -1.98, p = 0.044$ ). This learning phase depended highly on matching voices with recently heard targets and, as we will see again below, KH's performance in such tasks does not compare as unfavourably with controls as it does in tasks that require memory for voices. Especially in the first block, in which there were only six test trials after listening to each voice, KH performed well within controls' performance. However, after more test trials in the second and third blocks, KH's performance dropped, whereas controls improved with more exposure to the voices.

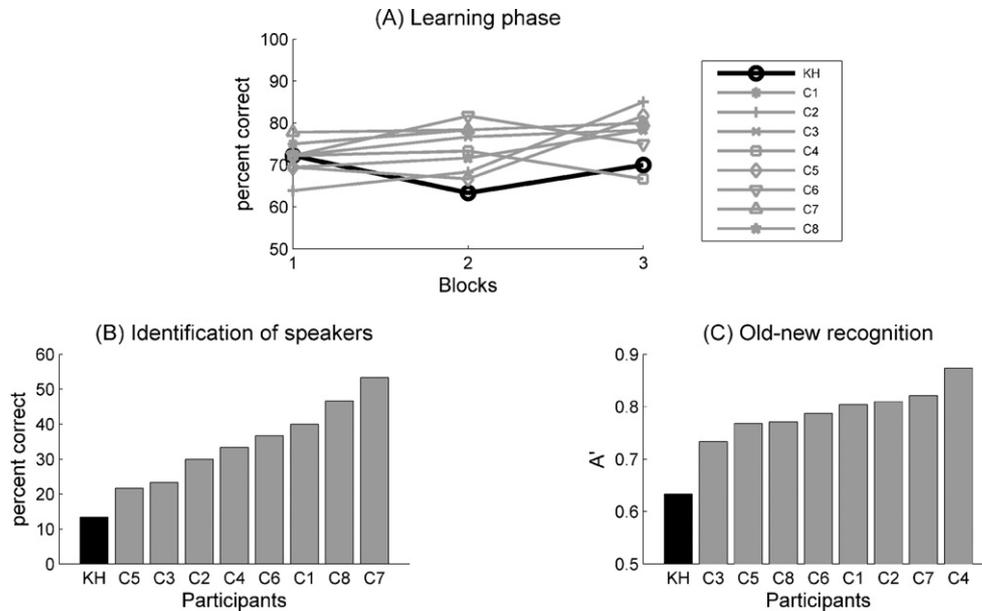
**2.2.1.2.2. Identification of speakers. Material and procedure:** Immediately after the learning task described above, participants were presented with 60 sentences (new sentences, 10 by each speaker in randomized order) and were asked to select the name corresponding to each speaker. Names appeared on the computer screen. Feedback (one beep) was given for incorrect responses.

**Results:** With six choices, chance performance on this task was 16.67%. The mean percent correct for controls was 35.63% (S.D. = 10.94). KH correctly identified only 13.33% of the test items. This result is significantly different from controls ( $t(7) = -1.92, p = 0.048$ ). Fig. 2B shows the individual scores on this task.

**2.2.1.2.3. Old–new recognition. Material and procedure:** KH's impaired performance on the previous task could result from problems associating the names to the voices rather than a failure to recognize the voices. In our next experiment, participants were not required to identify voices with names. For each sentence, participants had to decide whether it was said by one of the six speakers whose voices they had learnt during the two previous tasks or whether it was said by a new speaker. Like the target speakers, new speakers were young females with accents similar to the target speakers. New recordings were done for all speakers (targets and distracters), this time in a silent room, using Cool Edit 96. Peak amplitude for



**Fig. 1.** KH and controls' performance on *Recognition of famous voices* task. (A) *A'* for the discrimination of famous vs non-famous voices. *A'* varies between 0.5 (chance) and 1.0 (perfect discrimination). (B) The percent of voices correctly identified of the voices to which each participant was exposed.



**Fig. 2.** KH and controls' performance on *Learning and recognition of six different speakers* task. (A) The percent of correct responses on the learning phase. (B) The percent of correctly identified voices after learning. (C)  $A'$  for the discrimination of 'old' speakers and 'new' speakers.  $A'$  varies between 0.5 (chance) and 1.0 (perfect discrimination).

all stimuli was matched using PRAAT. The test trials included six sentences by each of the old speakers (36 'old' trials) and four sentences said by each of the nine new speakers (36 'new' trials).

**Results:** Controls discriminated between the voices of old and new speakers with a mean  $A'$  of 0.80 (S.D. = 0.04). Individual results can be seen in Fig. 2C. KH performed much lower than any control, and her  $A'$  of 0.63 is significantly different from the controls' performance ( $t(7) = -4.01, p = 0.003$ ). With 72 trials, KH's total correct of 40 is not significantly different from chance ( $z = 0.83, p = 0.20$ ).

## 2.2.2. Voice discrimination

KH's performance on the previous tasks shows that she clearly has a problem with memory for voices. The following tasks investigated whether she also has difficulties discriminating between two sequentially presented voices. Van Lancker and colleagues (Van Lancker & Kreiman, 1987; Van Lancker et al., 1988) showed that there was a double dissociation between the recognition of famous voices and the discrimination of two unfamiliar voices. Our first task used German speakers because we thought that using samples in German would cause participants to be more reliant on the acoustic properties of the voices, rather than on accent information. The second and third tasks used English speakers, and two different types of distortion were applied to the voices to avoid ceiling effects.

**2.2.2.1. Voice matching test. Material and procedure:** This test (created by Neuner & Schweinberger, 2000) used 50 pairs of voice samples lasting about two seconds each from unfamiliar German speakers. Participants decided whether pairs were spoken by the same or different speakers (the sentences were always different). Half the pairs used the same speaker and half different speakers. Stimuli were presented using iTunes.

**Results:** The eight controls correctly matched the voices on 81.50% (S.D. = 2.56) of trials. KH's score of 80% is within the controls' range and shows that she did not have difficulties with this task.

**2.2.2.2. Discrimination of voices using samples with combined noise. Material and procedure:** Participants again made same-different judgements on voice pairs, though these were English voices with white noise added. New sentences uttered by the same six female speakers used in task 2.2.1.2. were used (This discrimination task was performed before task 2.2.1.2.). White noise was added to the sentences using MATLAB. The script combined noise and signal waveforms at four different signal-to-noise ratios: -6, 0, 6, and 12 dB. There were 96 trials. Half the pairs were uttered by the same speaker and half by different speakers. Sentences in each pair were always different. There were 24 sentences pairs for each signal-to-noise ratio.

**Results:** Controls correctly matched 66.15% (S.D. = 9.03), 82.29% (S.D. = 5.34), 72.40% (S.D. = 8.31) and 80.21% (S.D. = 5.34) of the trials for each of the signal-to-noise ratios respectively (-6, 0, 6, 12). A pilot study had shown that performance increased with higher levels of signal-to-noise, but this did not happen for this group. KH scored 62.50%, 66.67%, 70.83% and 70.83%. Only the score for the second signal-to-noise ratio was significantly different from controls ( $t(7) = -2.76, p = 0.014$ ), though the difference for the fourth level approached significance ( $t(7) = -1.66, p = 0.071$ ).

**2.2.2.3. Discrimination of voices using noise-vocoded speech. Material and procedure:** In this task participants discriminated two samples of noise-vocoded speech. Stimuli were created by having 21 native British English speakers read sentences aloud from the BKB list. Different speakers read different sets of sentences. Stimuli were recorded using a microphone and the program Cool Edit 96. Stimuli were normalised for peak amplitude using PRAAT and sentences were noise-vocoded using the same programme. In the noise-vocoding transformation (Shannon, Zeng, Kamath, Wygonski, & Ekelind, 1995), the original speech stimulus is divided into several frequency bands. The amplitude envelope is extracted from each of these bands and used to modulate wideband noise (in the corresponding bandwidth). The modulated noise-bands are then summed to give a stimulus with impoverished spectral detail and sense of pitch. The more frequency channels used during vocoding, the more spectral information is preserved. For the current task, stimuli with six, 16, and 48 frequency channels were created to cover a range of spectral clarities.

Participants decided whether pairs of sentences were said by the same speaker or different speakers, with 50% being same pairs. For each pair, sentences were always said by speakers of the same sex with half said by male speakers. There were 28 pairs for each of the three channel levels making a total of 84 trials.

**Results:** Control 1 did not perform this task because she did not participate in the last testing session. The other seven controls correctly matched an average of 62.24% (S.D. = 8.95), 77.55% (S.D. = 5.73) and 77.04% (S.D. = 3.49) of trials for each of the noise-vocoded levels respectively (6, 16, 48). KH scored 60.71%, 64.29% and 71.43%. There was a significant difference between KH's and controls' scores for the 16-band noise-vocoded items ( $t(6) = -2.17, p = 0.037$ ). As in the previous task, KH's performance was almost always lower than controls' results but without being consistently significantly different from their scores.

## 2.2.3. Perception of emotion from the voice

To examine whether KH's difficulties are restricted to the recognition of speaker's identity or extend to the processing of other information from the voice, the following tasks investigated emotion recognition from the voice using both non-verbal and verbal stimuli (created by Sauter, 2006; Sauter & Scott, 2007). Belin et al.'s (2004) model of voice processing predicts that there will be cases of phonagnosia with spared processing of vocal affective information.

**2.2.3.1. Recognition of vocal expressions of emotion. Material and procedure:** Stimuli were 90 non-verbal emotion sounds expressing one of the following emotions: achievement/triumph, amusement, anger, disgust, fear, pleasure, relief, sadness and surprise (10 stimuli for each emotion). For each stimulus, participants had to select the corresponding adjective presented on the screen (from a choice of the nine above possibilities) that best described the emotion in the voice.

**Results:** Controls selected the correct adjective on 82.22% (S.D. = 6.06) of trials. KH percent correct of 83.33% places her score slightly above the control mean.

**2.2.3.2. Recognition of emotion via paralinguistic cues in speech. Material and procedure:** Stimuli consisted of emotionally inflected spoken three-digit numbers. Like the previous task, there were ten stimuli for each of the nine emotion categories. There were 10 additional stimuli conveying contentment, for a total of 100 trials. All

10 adjectives were presented on the computer screen on each trial, and participants selected the adjective that best described the emotion in the voice.

*Results:* Controls correctly identified the emotions on 72.13% (S.D. = 5.46) of trials. KH's score of 70% was within the controls' range.

#### 2.2.4. Perception of gender from the voice

Another type of information that can be extracted from the voice is the gender of the speaker. We tested KH on a task requiring gender identification using noise-vocoded stimuli to increase the difficulty of the task.

**2.2.4.1. Perception of gender in noise-vocoded speech. Material and procedure:** Twenty-six native English speakers read sentences aloud from the BKB Sentence List (stimuli were collected using a microphone and the program Cool Edit 96). Different speakers read different sets of sentences. Sentences were matched for peak amplitude and noise-vocoded using PRAAT. Three, six, and twelve frequency channels were used.

On each trial, participants listened to one sentence and were asked to decide whether the speaker was a man or a woman. Half the sentences were from male and half from female speakers. There were 30 sentences for each frequency channel making a total of 90 trials.

*Results:* Control 1 did not perform this task because she did not participate in the last testing session. The other seven controls correctly identified the gender of the speaker on 52.38% (S.D. = 3.71), 74.76% (S.D. = 13.86) and 90.95% (S.D. = 14.36) of the trials for each of the noise-vocoded levels respectively (3, 6, 12). KH scored 53.33%, 76.67% and 100%, all of which were above the control means.

#### 2.2.5. Speech perception

The following tasks examined KH's perception of speech. In Belin et al.'s (2004) model, perception of speech information is processed separately from information about speaker's identity. We tested KH on four tests, increasing the demands of the tasks by adding noise and using noise-vocoded speech.

**2.2.5.1. Verbal sound processing. Material and procedure:** This task (created by Saygin, Dick, Wilson, Dronkers, & Bates, 2003; Saygin, Dick, & Bates, 2005) required rapid matching of sentences to pictures on the screen. Participants listened to 90 sentences and matched each one to one of two pictures on the screen. On half the trials, the distracter was related to the target and on the other half it was unrelated. For example, when the target was 'cow', the distracters were either a sheep or a violin. Verbal and non-verbal (as described in Section 2.2.6.1.) trials were randomized. Stimuli were delivered using PsyScope software (Cohen, MacWhinney, Flatt, & Provost, 1993). Participants were asked to respond as quickly as possible.

*Results:* Controls performed accurately on 99.37% (S.D. = 0.87) of trials and the mean of their reaction times was 1341 milliseconds (ms) (S.D. = 317). KH correctly matched 100% of the trials and her average reaction time was 1346 ms, both of which were within the controls' range.

**2.2.5.2. Perception of speech in white noise. Material and procedure:** Stimuli consisted of 100 sentences with three key words from the BKB List. The speaker was a female native British English speaker. Participants had not heard the sentences or the speaker's voice before. Recordings were done in an anechoic chamber. White noise was combined with the sentences using MATLAB. Five different signal-to-noise ratios were used (-6, -3, 0, 3 and 6 dB) with 20 sentences in each. Participants listened to each sentence and were asked to say as many words as possible from the sentence. The experimenter typed all answers and initiated the next trial.

*Results:* The eight controls correctly reported an average of 25.21% (S.D. = 8.75), 78.96% (S.D. = 5.49), 92.08% (S.D. = 4.94), 97.29% (S.D. = 2.17) and 98.96% (S.D. = 1.77) of words for each of the signal-to-noise ratios respectively: -6, -3, 0, 3 and 6. KH scored 8.33%, 60%, 90%, 83.33% and 100% for each of the noise levels. The result for the second ( $t(7) = -3.26, p = 0.007$ ) and fourth ( $t(7) = -3.89, p = 0.003$ ) signal-to-noise ratios were significantly different from controls' performance. However, this task was done at the end of a fatiguing session, and we were concerned that this might explain KH's poor performance. To investigate whether fatigue or genuine problems with speech perception account for these results, we tested her with two other speech perception tasks.

**2.2.5.3. Perception of noise-vocoded speech. Material and procedure:** Stimuli consisted of 60 sentences, each with three key words from the BKB List. The speaker was a female native British English speaker. Recordings were done in an anechoic chamber. Sentences were noise-vocoded using PRAAT. Three levels of spectral resolution were used: four, six, and eight channels. Participants listened to each sentence and were asked to say as many words as possible from the sentence. The experimenter typed all answers and initiated the next trial.

*Results:* Control 1 did not perform this task because she did not participate in the last testing session. The other seven controls correctly reported 54.52% (S.D. = 18.20), 70% (S.D. = 13.05) and 95.83% (S.D. = 1.75) of words for each of the noise-vocoded levels respectively (4, 6, 8). KH scored normally with 65%, 68.33% and 95% at the three levels, all of which were within the controls' range.

**2.2.5.4. Identification of vowels. Material and procedure:** Despite KH's normal performance on the previous task, it could be that KH has an impairment of speech perception that is not well captured in the type of tasks we used. In these tasks sentences were presented and some words could be inferred from context. Therefore, we also tested KH on a task that demanded the identification of single vowels.

Stimuli consisted of 18 different vowels in English, with five exemplars of each vowel (see Appendix A for details). White noise was combined with the 90 stimuli using MATLAB. One single signal-to-noise ratio was used (-3 dB) because this was one of the levels at which KH showed impaired performance in task 2.2.5.2. Participants listened to each vowel and were asked to identify to which of the 18 options it corresponded.

*Results:* Eight controls (four females, native English speakers, not age-matched; mean age = 25.29, S.D. = 6.05) correctly identified 79.72% (S.D. = 8.96) vowels. KH's score of 73.33% is within the controls' range.

#### 2.2.6. Recognition of environmental sounds

We also investigated whether KH's impairment is restricted to the voice or extends to other auditory processing, as part of a more general auditory agnosia. The next two tasks examined how KH recognizes environmental sounds.

**2.2.6.1. Non-verbal sound processing. Material and procedure:** In this task (Saygin et al., 2003, 2005), participants listened to 90 environmental sounds and matched them to one of two pictures on the screen. The non-verbal trials were randomized with the verbal ones described in Section 2.2.5.1. In half the trials, the distracter was related to the target and in the other half it was unrelated. Participants were asked to respond as quickly as possible.

*Results:* Controls accurately identified 97.94% (S.D. = 1.00) sounds and the mean of their reaction times was 1293 ms (S.D. = 365). KH correctly matched 98% sounds to the pictures and her average reaction time was 1166 ms, both of which were within the controls' range.

**2.2.6.2. Sound recognition and naming test. Material and procedure:** In this task (created by Neuner & Schweinberger, 2000), participants were presented (using iTunes) with 56 environmental sounds (e.g. dentist's drill, ducks, applause) and decided whether each sound was produced by an animate or an inanimate object. They were then asked to name the sound.

*Results:* The eight controls correctly categorised 94.87% (S.D. = 2.60) sounds as animate or inanimate, and correctly named 63.17% (S.D. = 11.61) of the sounds. KH correctly classified 94.64% of the sounds as animate or inanimate and named 67.86% of the items, both of which are within the controls' range.

#### 2.2.7. Music perception

In addition to non-environmental sounds, we investigated KH's ability to perceive music. Because some studies have shown that voice recognition impairments and amusia tend to co-occur (Assal et al., 1981; Mazzucchi, Marchini, Budai, & Parma, 1982; Peretz et al., 1994), it could be that KH's voice recognition difficulties reflect an impairment to mechanisms that are both important for voice recognition and certain music processing tasks.

**2.2.7.1. Familiar tunes. Material and procedure:** This task is distributed as part of the Montreal Battery of Evaluation of Amusia (MBEA—Peretz, Champod, & Hyde, 2003). It consists of 20 tunes, half familiar and half unfamiliar. For each tune, participants had to decide whether it was familiar or not. If they reported it was familiar, they were asked to name that tune.

*Results:* Controls correctly classified 99.29% (S.D. = 1.89) tunes as familiar or unfamiliar. They correctly named or provided specific information for 70% (S.D. = 16.33) tunes. KH correctly classified all tunes as familiar or unfamiliar, and correctly identified 70% of tunes, which is within the controls' range.

**2.2.7.2. Timbre discrimination test. Material and procedure:** In this test (created by Jason Warren and Rohani Omar), participants were presented with two different melodic excerpts each played by a single instrument, and their task was to decide whether the excerpts were played by the same instrument or by different instruments. Excerpts were between 1.1 and 6.9 s in duration (mean 4.4 s). Sixteen instruments were included (accordion, bassoon, castanets, cello, clarinet, flute, French horn, guitar, harp, harpsichord, oboe, organ, piano, saxophone, violin, and xylophone), all in common use in Western classical and/or popular music. Excerpts were selected such that the timbre of the instrument was strongly established whereas recognition of the piece was unlikely based on the information included. Within each pair the two excerpts differed in pitch range, to reduce the use of non-timbral cues. Twenty trials were presented (10 same, 10 different pairs).

*Results:* The eight controls correctly discriminated the timbre on 82.14% (S.D. = 12.54) of trials. KH's score of 85% was above the controls' average.

**2.2.7.3. Montreal battery of evaluation of amusia (MBEA). Material and procedure:** To further explore KH's music processing abilities, she was tested on four subtests of the MBEA (Peretz et al., 2003): Scale, Contour, Interval, and Rhythm tests. Each subtest consisted of 30 trials in which a musical phrase was presented, followed

by either the same musical phrase (15 'same' trials), or an altered version of the musical phrase (15 'different' trials). For the first three subtests, the different pairs were created by manipulating a particular note in the musical phrase. In the Scale subtest, this manipulation involved replacing a single note from the original phrase with an out of key note. In the Contour subtest, the manipulation involved replacing the original note with another within-key note that changed the pattern of the ups and downs of the musical phrase. In the Interval subtest, the manipulation involved replacing the original note with a note that was within key and did not change the pattern of ups and downs. The final test, Rhythm, involved different pairs in which the duration of two adjacent notes was manipulated, thereby changing the rhythmic grouping, while leaving metre and total number of sounds unaltered. Participants were asked to verbally report whether the two musical phrases were the same or different.

**Results:** KH scored 27 on the Scale Test, 27 on the Contour Test, 24 on the Interval test and 28 on the Rhythm Test. Compared to the results from 160 controls reported in Peretz et al. (2003), KH is within the normal range on all these tasks (controls' results: Scale Test:  $M = 27$ ;  $S.D. = 2.3$ ; Contour Test:  $M = 27$ ;  $S.D. = 2.2$ ; Interval Test:  $M = 26$ ;  $S.D. = 2.4$ ; Rhythm Test:  $M = 27$ ;  $S.D. = 2.1$ ).

### 2.2.8. Face recognition

Finally, voice recognition problems could result from higher-level person recognition deficits (Neuner & Schweinberger, 2000). To examine whether KH has general problems with identity recognition, we tested her face recognition with two tests.

**2.2.8.1. Cambridge face memory test (CFMT). Material and procedure:** The CFMT (Duchaine & Nakayama, 2006b) was designed to identify individuals with prosopagnosia. Participants were introduced to six faces and then were asked to recognize those faces in novel views.

**Results:** Duchaine and Nakayama (2006b) reported that 50 college students scored a mean of 57.9 ( $S.D. = 7.9$ ) out of 72. Control 1 did not perform this task because she did not participate in the last testing session. The other seven controls scored a mean of 56.43 ( $S.D. = 9.71$ ). KH' score of 67 was above controls' average.

**2.2.8.2. Recognition of famous faces. Material and procedure:** In this test (Duchaine & Nakayama, 2005), participants were presented with 60 photographs of famous people for 3 s each, and asked to name or provide specific information about that person. After the test, participants were asked whether they thought they had seen many times the people they failed to identify.

**Results:** Controls 1, 3 and 8 did not perform this task. The other five controls correctly identified an average of 46.40 faces ( $S.D. = 7.92$ ) and reported being exposed to 57.20 ( $S.D. = 2.77$ ) of these 60 celebrities. KH correctly identified 52 faces and said she was exposed to all 60, which are within the controls' range.

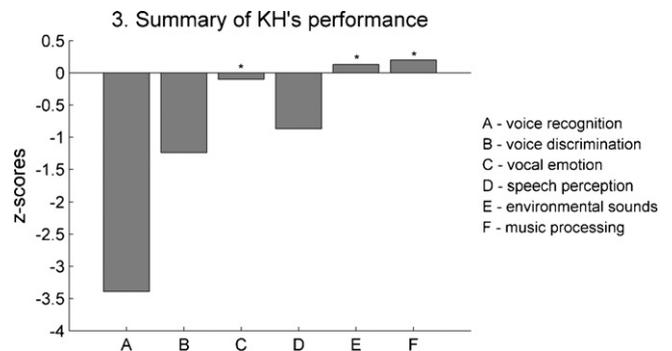
### 2.2.9. Dissociation between speaker recognition and other abilities

When compared to controls, KH's was clearly impaired on voice recognition tasks. Her scores on voice discrimination and speech recognition tasks were not as clear, with some impaired and some unimpaired results, but she performed comparably to controls on all tasks of emotion and gender perception, recognition of faces, environmental sounds and music processing. Crawford and Garthwaite (2005, 2007), however, have pointed out that to be able to say that a patient presents a dissociation between task X and task Y, it is not enough to show that the score in task X is significantly different from controls' mean and that the result in task Y is within normal performance. They argue that it also important to show that the patient's result on task X is significantly poorer than the one on task Y. Therefore, we investigated whether KH's impairment in voice recognition dissociates from the other abilities tested by evaluating whether the difference between KH's performance on voice recognition and each of the other abilities was significantly greater than the one observed for controls.

We created a composite score for each of the abilities tested using standardized scores (see Fig. 3 and Appendix B for details). Then, we compared the difference between KH's standardized scores on voice recognition and each of the other group of tasks with the differences obtained by controls using the Bayesian standardized difference test developed by Crawford and Garthwaite (2007). KH's difference between performance on voice recognition and voice discrimination tasks was not significantly different from the one observed for controls ( $p = 0.142$ ), nor was the difference between voice recognition and speech perception scores ( $p = 0.252$ ). All other results were significant, indicating dissociations between voice recognition and emotion recognition ( $p = 0.016$ ), recognition of environmental sounds ( $p < 0.001$ ), and music perception ( $p < 0.005$ ).

## 3. Discussion

KH complains of severe voice recognition difficulties in her daily life. She was tested on two series of voice recognition tasks, one requiring the recognition of famous voices and one requiring the learning and recognition of the voices of six unknown speakers.



**Fig. 3.** KH's standardized scores for each group of tasks. We compared the difference between KH's standardized scores on voice recognition and each of the other group of tasks with the differences obtained by controls using the Bayesian standardized difference test developed by Crawford and Garthwaite (2007). Asterisks (\*) indicate that the difference between KH's performance on voice recognition and that specific ability was significantly greater than the one observed for controls.

When compared to matched controls, her performance was clearly impaired on all tasks. KH does not present a neurological lesion that could explain these difficulties. Moreover, she reports that she has always experienced voice recognition problems, therefore we consider her to be a case of developmental phonagnosia. Belin et al. (2004) suggested that this condition might exist, but until now no cases have been documented.

Identification of a case of developmental phonagnosia adds to the list of selective deficits that can result from developmental problems. These include dyslexia (e.g. Ramus et al., 2003), specific language impairment (e.g. van der Lely, 2005), dyscalculia (e.g. Landerl, Bevan, & Butterworth, 2004), prosopagnosia (e.g. Duchaine & Nakayama, 2006a) and amusia (e.g. Ayotte, Peretz, & Hyde, 2002). These developmental conditions have been at the centre of important debates about cognitive development and the specificity of cognitive impairments (e.g. Karmiloff-Smith, 1998; Ramus, 2004). They suggest that specific cognitive mechanisms can fail to develop normally while other abilities are normal, and KH's case indicates that this is also possible for mechanisms involved in voice perception. The study of some developmental conditions has led to the identification of genes that contribute to cognitive development (Fisher, 2006; Galaburda, Lo Turco, Ramus, Fitch, & Rosen, 2006), and has shed light on the neural basis of particular abilities (e.g. Behrmann, Avidan, Gao, & Black, 2007; Hyde, Zatorre, Griffiths, Verch, & Peretz, 2006; Paulesu et al., 2001; Price, Holloway, Räsänen, Lesterinen, & Ansari, 2007; Silani et al., 2005), and we expect that studies of cases similar to KH could lead to similar insights for speaker recognition.

Understanding of the neural correlates of voice recognition has seen considerable development in recent years. Bilateral regions in the upper bank of the superior temporal sulcus (STS) respond selectively to voices (Belin, Zatorre, Lafaille, Ahad, & Pike, 2000), and the identification of speakers seems to be associated with activity in the right anterior STS (Belin & Zatorre, 2003; von Kriegstein, Eger, Kleinschmidt, & Giraud, 2003). Comparable functional regions were found in the macaque monkey, with a region in the anterior superior temporal plane preferentially responsive to the vocalizations of macaques and sensitive to the identity of conspecific individuals (Petkov et al., 2008). Deficits to the regions identified in these studies are likely candidates to explain KH's impairment.

### 3.1. Implications for models of voice and auditory processing

Cases of phonagnosia following brain damage have rarely been reported and their specificity is poorly understood. KH's case allows

examination of how the mechanisms used to recognize vocal identity relate to other voice and auditory abilities. The model of voice recognition proposed by Belin et al. (2004; see also Campanella & Belin, 2007) suggests that the recognition of a speaker's identity could dissociate from the recognition of emotions from the voice as well as from speech perception, but no prior cases exhibit these dissociations. KH, however, did very well on two tasks of emotion recognition, one consisting of verbal stimuli with different emotional intonations and the other consisting of non-verbal emotional vocalizations. Her results therefore provide support for the model's proposal that vocal identity and vocal emotion processing are carried out by separable mechanisms.

We did not find clear evidence for a dissociation between recognition of speaker and perception of speech. However, the failure to find a dissociation between these abilities was primarily due to KH's impaired performance on one of the four speech perception tasks with which she was tested. We believe her poor performance could be due to fatigue. Moreover, KH's performance was normal for the three other tasks, including a test that required the identification of vowels in noise, in which it was not possible to rely on the context of a sentence to perform accurately. In any case, it will be important to repeat these tasks in the future. Future testing of KH's speech perception should take into account results that show that speech perception can be improved by increased familiarity with the talker (Nygaard, Sommers, & Pisoni, 1994). All tasks of speech perception we described used the same speaker for all trials, therefore controls' performance might have benefited from the increased exposure to the same speaker.

To investigate the processing of other auditory stimuli, KH was tested with environmental sounds and music tests. Neuner and Schweinberger (2000) described a few cases of brain-damaged patients with voice recognition problems who could still recognize environmental sounds. These authors argued that this demonstrated that those impairments could not be explained by a general auditory agnosia but were restricted to person recognition from the voice. Consistent with their finding, KH performed as well as controls on two tasks requiring the recognition and naming of environmental sounds. However, these results, like past studies, should be interpreted cautiously because it could be that these tasks are easier than the voice recognition tasks. In tasks with environmental sounds, KH was asked to identify sounds from different categories, whereas in the voice recognition tasks, she was asked to identify specific individuals from the same category, which were highly similar. This mirrors an important issue in the literature on the neuropsychology of face processing (Damasio, Damasio, & Van Hoesen, 1982). Early tests measuring face recognition and object recognition were not equivalent in that the face tests required recognition at the individual level whereas the object tests only required recognition at the basic level. Double dissociations between face and object tests requiring recognition at the individual level have resolved this issue (Duchaine, Yovel, Butterworth, & Nakayama, 2006; McNeil & Warrington, 1993; Moscovitch, Winocur, & Behrmann, 1997; Sergent & Signoret, 1992), but this remains unexamined for voice recognition. Variable expertise with different stimulus classes has also been a key concern in the face literature (Diamond & Carey, 1986; Robbins & McKone, 2007), and issues of expertise should be explored in voice processing as well.

Music perception tasks were also used to test KH's ability to process non-vocal auditory stimuli. Results from some brain-damaged patients suggest that impairments in voice processing and certain music tasks tend to co-occur (Assal et al., 1981; Mazzucchi et al., 1982; Peretz et al., 1994). However, Ayotte et al. (2002) demonstrated that a group of congenital amusics, with impairments in processing pitch as well as memory for music, could still recognize and name famous voices as well as controls. The opposite pattern

was found in KH, who did as well as controls on all tasks of music processing we used. KH also reports that she enjoys listening to music and that she usually recognizes the songs she has listened to before, though interestingly she says she usually cannot recognize specific singers. The dissociation between her impaired speaker recognition and preserved musical listening abilities suggest that these abilities rely, at least in part, on different processes.

### 3.2. KH's deficit in processing speaker's identity

KH shows a clear impairment in memory for voices. This impairment does not seem to extend to other auditory material, nor can it be explained as a general deficit in person recognition. She did very well on face recognition tasks, either requiring the learning and recognition of new faces or the recognition of famous people. With respect to the voice abilities, it is unclear, though, whether KH can discriminate between two voices. Van Lancker and colleagues have reported a double dissociation between voice discrimination and familiar voice recognition (Van Lancker & Kreiman, 1987; Van Lancker et al., 1988). From our results, however, there is no clear evidence for a dissociation in performance between the tasks requiring memory for voices and the tasks requiring perceptual discrimination of two voices (see Fig. 3). Nevertheless, KH's performance was certainly worse for the former, and she had some unimpaired scores on voice discrimination tasks, though this differential performance could result from the simpler designs in the discrimination tasks which had only two response options. When tasks were made more difficult by distorting the stimuli, KH had more difficulty with them and future testing should use other challenging tasks to assess matching.

It would also be valuable to characterize which aspects of vocal information KH fails to process normally. It has been shown that listeners can extract information from the voice about the speaker's age, gender, body size or pronunciation (e.g. Krauss, Freyberg, & Moresella, 2002; Rendall, Vokey, & Nemeth, 2007). From KH's results, it seems she can identify vocal gender without difficulties, but it would be interesting to investigate how she perceives age, height, and different dialects. These are important aspects of the identity of a speaker, but it is usually thought that the recognition of a specific speaker goes beyond the perception of these characteristics. However, little is understood about which properties of the voice are important for speaker recognition, and it is possible that different acoustic features contribute to the recognition of different familiar voices (Van Lancker, Kreiman, & Emmorey, 1985), even though it is frequently argued that the acoustic properties of timbre play a particularly important role in identifying a speaker. Warren, Scott, Price, & Griffiths (2006) focused on the spectral envelope as a dimension of timbre that is especially important for the analysis of speaker identity. Remez, Fellowes, & Rubin (1997), on the other hand, showed that phonetic properties of speech can also serve to identify the speaker. Careful manipulation of these properties with KH may allow a more specific description of the nature of her voice recognition deficit and a better understanding of the properties that are important in normal speaker recognition.

## 4. Summary

To conclude, this study identified the first reported case of developmental phonagnosia. KH's impairment seems to be highly specific to a single cognitive domain—voice recognition. Despite her serious impairments in recognizing voices, KH performed normally with tasks measuring face recognition, speech perception, vocal affect recognition, music perception and recognition of environmental sounds. Her results provide support for modular models

of voice and high-level auditory processing. Other selective developmental conditions have shed light on the cognitive, neural, developmental and genetic basis of particular abilities, and we expect that developmental phonagnosia will provide a means to address these issues for voice processing.

### Acknowledgments

We are very grateful to KH and the controls for their time and enthusiastic participation. We thank Ayse Saygin and Ljubica Damjanovic for making their tasks available to us, all the speakers whose voices we recorded, Laura Germiné for help with testing, and Matthew Longo for comments in an earlier version of the manuscript. This research was supported by a studentship from the Portuguese Foundation for Science and Technology (SFRH/BD/22580/2005) to L.G., a Wellcome Intermediate Clinical Fellowship to J.W. and an ESRC (RES-061-23-0400) grant to B.D.

### Appendix A

Stimuli consisted of 18 different vowels in English. These fell into three sets of six each: short monophthongs (bad (æ), bed (e), bid (ɪ), bod (ɔ), bood (ʊ), bud (ʌ)), long monophthongs (bard (ɑː), bared (ɛː), bead (iː), bird (ɜː), booted (uː), bored (ɔː)) and diphthongs (bade (eɪ), beard (ɪə), bide (aɪ), bode (əʊ), bowed (aʊ), boyed (ɔɪ)).

Five tokens of each vowel were recorded, in the lexical context given above, in an anechoic chamber. The vowel portion was extracted from each of these tokens. Each of the sets of isolated vowels was then normalized (within-set) for duration and root-mean-square sounds pressure. Finally, each vowel was spliced back into an identical/b/-/d/frame, taken from one of the tokens of 'boyed'. All normalization and replacing procedures were performed in PRAAT. Duration normalization was performed using PSOLA (based on an algorithm by Charpentier & Stella, 1986).

### Appendix B

KH's and the eight matched controls' results for each task were converted to z scores based on means and standard deviations of controls. We calculated individual composite scores for each ability by calculating the mean of z scores of the tasks that were used to test that ability. For the voice recognition composite, we did not include the results from task 2.2.1.2.1 (Learning phase), given that it appears to be highly reliant on the discrimination of voices. The results from tasks 2.2.5.4 (Identification of vowels) and 2.2.7.3 (MBEA) were not included because only KH, and not the matched controls, was tested on these tasks. Furthermore, we did not include a composite score for the perception of vocal gender or face recognition because one control was not tested with these tasks. But note that KH's performance was better than controls' average for both tasks. We finally note that not all controls completed all other tasks (see main text), but we calculated the composite scores with the scores available for each ability.

The Bayesian standardized difference test calculates the probability (two-tailed) that the difference between a case's two scores is an observation from the control population.

### References

Assal, G., Aubert, C., & Buttet, J. (1981). Asymétrie cérébrale et reconnaissance de la voix. *Revue Neurologique*, 137, 255–268.

Ayotte, J., Peretz, I., & Hyde, K. (2002). Congenital amusia. A group study of adults afflicted with a music specific disorder. *Brain*, 125, 238–251.

Behrmann, M., & Avidan, G. (2005). Congenital prosopagnosia: Face-blind from birth. *Trends in Cognitive Sciences*, 9, 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, 2354–2363.

Belin, P., Fecteau, S., & Bédard, C. (2004). Thinking the voice: Neural correlates of voice perception. *Trends in Cognitive Sciences*, 8, 129–135.

Belin, P., & Zatorre, R. (2003). Adaptation to speaker's voice in right anterior temporal lobe. *Neuroreport*, 14, 2105–2109.

Belin, P., Zatorre, R., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice selective areas in the human auditory cortex. *Nature*, 403, 312–319.

Bench, J., Kowal, A., & Bamford, J. (1979). The BKB (Bamford–Kowal–Bench) sentence lists for partially hearing children. *British Journal of Audiology*, 13, 108–112.

Boersma, P., & Weenink, D. (2005). Praat: Doing phonetics by computer. [Computer programme]. <<http://www.praat.org/>>

Bornstein, B. (1963). Prosopagnosia. In L. Harpen (Ed.), *Problems in dynamic neurology* (pp. 283–318). Jerusalem: Hadassah Medical School.

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, 77, 305–327.

Burton, A. M., Bruce, V., & Johnston, R. A. (1990). Understanding face recognition with an interactive activation model. *British Journal of Psychology*, 81, 361–380.

Campanella, S., & Belin, P. (2007). Integrating face and voice in person perception. *Trends in Cognitive Sciences*, 11, 535–543.

Charpentier, F., & Stella, M. (1986). Diphone synthesis using an overlap-add technique for speech waveforms concatenation. In *Paper presented at the IEEE International Conference ASSP*.

Cohen, J., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments and Computers*, 25, 257–271.

Crawford, J., & Garthwaite, P. (2005). Testing for suspected impairments and dissociations in single-case studies in neuropsychology: Evaluation of alternatives using Monte Carlo simulations and revised tests for dissociations. *Neuropsychology*, 19, 318–331.

Crawford, J., & Garthwaite, P. (2007). Comparison of a single case to a control or normative sample in neuropsychology: Development of a Bayesian approach. *Cognitive Neuropsychology*, 24, 343–372.

Crawford, J. R., & Howell, D. C. (1998). Comparing an individual's test score against norms derived from small samples. *The Clinical Neuropsychologist*, 12, 482–486.

Damasio, A., Damasio, H., & Van Hoesen, G. (1982). Prosopagnosia: Anatomic basis and behavioral mechanisms. *Neurology*, 32, 331–341.

Damjanovic, L., & Hanley, R. (2007). Recalling episodic and semantic information about famous faces and voices. *Memory and Cognition*, 35, 1205–1210.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, 115, 107–117.

Duchaine, B., & Nakayama, K. (2005). Dissociations of face and object recognition in developmental prosopagnosia. *Journal of Cognitive Neuroscience*, 17, 249–261.

Duchaine, B., & Nakayama, K. (2006a). Developmental prosopagnosia: A window to content-specific processing. *Current Opinion in Neurobiology*, 16, 166–173.

Duchaine, B., & Nakayama, K. (2006b). The Cambridge Face Memory Test: Results from neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic individuals. *Neuropsychologia*, 44, 576–585.

Duchaine, B., Yovel, G., Butterworth, E., & Nakayama, K. (2006). Prosopagnosia as an impairment to face specific mechanisms: Elimination of the alternative hypotheses in a developmental case. *Cognitive Neuropsychology*, 23, 714–747.

Ellis, H., Jones, D., & Mosdell, N. (1997). Intra- and inter-modal repetition priming of familiar faces and voices. *British Journal of Psychology*, 88, 143–156.

Fisher, S. (2006). Tangled webs: Tracing the connections between genes and cognition. *Cognition*, 101, 270–296.

Galaburda, A., Lo Turco, J., Ramus, F., Fitch, R., & Rosen, G. (2006). From genes to behaviour in developmental dyslexia. *Nature Neuroscience*, 9, 1213–1217.

Gopal, K., Briley, K., Goodale, E., & Hendea, O. (2005). Selective serotonin reuptake inhibitors treatment effects on auditory measures in depressed female patients. *European Journal of Pharmacology*, 520, 59–69.

Hyde, K., Zatorre, R., Griffiths, T., Lerch, J., & Peretz, I. (2006). Morphometry of the amusic brain: A two-site study. *Brain*, 129, 2562–2570.

Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences*, 2, 389–398.

Krauss, R., Freyberg, R., & Moresella, E. (2002). Inferring speaker's physical attributes from their voices. *Journal of Experimental Social Psychology*, 38, 618–625.

Kress, T., & Daum, I. (2003). Developmental prosopagnosia: A review. *Behavioural Neurology*, 14, 109–121.

Landerl, K., Bevan, A., & Butterworth, B. (2004). Developmental dyscalculia and basic numerical abilities: A study of 8–9-year-old students. *Cognition*, 93, 99–125.

Lezak, M. (1995). *Neuropsychological assessment* (3rd ed). New York: Oxford University Press.

Mazzuchi, A., Marchini, C., Budai, R., & Parma, M. (1982). A case of receptive amusia with prominent timbre perception defect. *Journal of Neurology Neurosurgery and Psychiatry*, 45, 644–647.

McConachie, H. (1976). Developmental prosopagnosia A single case report. *Cortex*, 12, 76–82.

McMillan, N., & Creelman, C. (1990). Response bias: Characteristics of detection theory, threshold theory and 'nonparametric' indexes. *Psychological Bulletin*, 107, 401–413.

- McNeil, J., & Warrington, E. (1993). Prosopagnosia: A face-specific disorder. *The Quarterly Journal of Experimental Psychology A*, 46, 1–10.
- Moscovitch, M., Winocur, G., & Behrmann, M. (1997). What is special about face recognition? Nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition. *Journal of Cognitive Neuroscience*, 9, 555–604.
- Neuner, F., & Schweinberger, S. R. (2000). Neuropsychological impairments in the recognition of faces, voices, and personal names. *Brain and Cognition*, 44, 342–366.
- Nygaard, L., Sommers, M., & Pisoni, D. (1994). Speech perception as a talker-contingent process. *Psychological Science*, 5, 42–46.
- Paulesu, E., Démonet, J., Fazio, F., McCrory, E., Chanoine, V., Brunswick, N., et al. (2001). Dyslexia: Cultural diversity and biological unity. *Science*, 291, 2165–2167.
- Peretz, I., Champod, A., & Hyde, K. (2003). Varieties of musical disorders—The Montreal battery of evaluation of amusia. *Annals of the New York Academy of Sciences*, 999, 58–75.
- Peretz, I., Kolinsky, R., Tramo, M., Labrecque, R., Hublet, C., Demeurisse, G., et al. (1994). Functional dissociations following bilateral lesions of auditory cortex. *Brain*, 117, 1283–1301.
- Petkov, C., Kayser, C., Studel, T., Whittingstall, K., Augath, M., & Logothetis, N. (2008). A voice region in the monkey brain. *Nature Neuroscience*, 11, 367–374.
- Price, G., Holloway, I., Räsänen, P., Vestnerinen, M., & Ansari, D. (2007). Impaired parietal magnitude processing in developmental dyscalculia. *Current Biology*, 17, 1042–1043.
- Ramus, F. (2004). Neurobiology of dyslexia: A reinterpretation of the data. *Trends in Neurosciences*, 27, 720–726.
- Ramus, F., Rosen, S., Dakin, S., Day, B., Castellote, J., White, S., et al. (2003). Theories of developmental dyslexia: Insights from a multiple case study of dyslexic adults. *Brain*, 126, 841–865.
- Raven, J. C., Court, J., & Raven, J. (1981). *Advanced progressive matrices set I*. London: H.K. Lewis.
- Remez, R., Fellowes, J., & Rubin, P. (1997). Talker identification based on phonetic information. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 651–666.
- Rendall, D., Vokey, J., & Nemeth, C. (2007). Lifting the curtain on the wizard of Oz: Biased voice-based impressions of speaker size. *Journal of Experimental Psychology: Human Perception and Performance*, 33, 1208–1219.
- Rey, A. (1964). *L'examen clinique en psychologie*. Paris: Presses Universitaires de France.
- Robbins, R., & McKone, E. (2007). No face-like processing for objects-of-expertise in three behavioural tasks. *Cognition*, 103, 34–79.
- Sauter, D. (2006). *An investigation into vocal expressions of emotions: the roles of valence, culture, and acoustic factors*. Doctoral Thesis. University College London.
- Sauter, D., & Scott, S. (2007). More than one kind of happiness: Can we recognize vocal expression of different positive states? *Motivation and Emotion*, 31, 192–199.
- Saygin, A., Dick, F., & Bates, E. (2005). An online task for contrasting auditory processing in verbal and nonverbal domains and norms for younger and older adults. *Behavior Research Methods*, 37, 99–110.
- Saygin, A., Dick, F., Wilson, S., Dronkers, N., & Bates, E. (2003). Neural resources for processing language and environmental sounds—Evidence from aphasia. *Brain*, 126, 928–945.
- Sergent, J., & Signoret, M. (1992). Varieties of functional deficits in prosopagnosia. *Cerebral Cortex*, 2, 375–388.
- Shannon, R., Zeng, F., Kamath, V., Wygonski, J., & Ekelind, M. (1995). Speech recognition with primarily temporal cues. *Science*, 270, 303–304.
- Silani, G., Frith, U., Demonet, J., Fazio, F., Perani, D., Price, C., et al. (2005). Brain abnormalities underlying altered activation in dyslexia: a voxel based morphometry study. *Brain*, 128, 2453–2461.
- Strik, J., Honig, A., Klinkenberg, E., Dijkstra, J., & Jolles, J. (2006). Cognitive performance following fluoxetine treatment in depressed patients post myocardial infarction. *Acta Neuropsychiatrica*, 18(1), 1–6.
- van der Lely, H. K. (2005). Domain-specific cognitive systems: Insight from Grammatical-SLI. *Trends in Cognitive Sciences*, 9, 53–59.
- Van Lancker, D., & Canter, G. (1982). Impairment of voice and face recognition in patients with hemispheric damage. *Brain and Cognition*, 1, 185–195.
- Van Lancker, D., Cummings, J., Kreiman, J., & Dobkin, B. (1988). Phonagnosia: A dissociation between familiar and unfamiliar voices. *Cortex*, 24, 195–209.
- Van Lancker, D., & Kreiman, J. (1987). Voice discrimination and recognition are separate abilities. *Neuropsychologia*, 25, 829–834.
- Van Lancker, D., Kreiman, J., & Cummings, J. (1989). Voice perception deficits: Neuroanatomical correlates of phonagnosia. *Journal of Clinical and Experimental Neuropsychology*, 11, 665–674.
- Van Lancker, D., Kreiman, J., & Emmorey, K. (1985). Familiar voice recognition: Patterns and parameters. Part I. *Journal of Phonetics*, 13, 19–38.
- von Kriegstein, K., Eger, E., Kleinschmidt, A., & Giraud, A. (2003). Modulation of neural responses to speech by directing attention to voices or verbal content. *Cognitive Brain Research*, 17, 48–55.
- Warren, J., Scott, S., Price, C., & Griffiths, T. (2006). Human brain mechanisms for the early analysis of voices. *NeuroImage*, 31, 1389–1397.