PRE-SCHOOL CHILDREN'S IDENTIFICATION OF FAMILIAR SPEAKERS AND THE ROLE OF ACCENT FEATURES

ELLA JEFFRIES

University of York

Abstract

This paper explores pre-school children's abilities in recognising and identifying familiar speakers. Two experiments were run with a group of 22 nursery children from York, in the north of England, aged 2.4-4.9 years. The first experiment investigates the children's ability to identify six known nursery school teachers from short audio stimuli. The results of this experiment showed an improvement in the identification ability of the children correlating with age, greater exposure to the teachers, and a longer stimulus length. The pitch and voice quality of the speakers' voices were also found to play a role in how easily identified the individuals were. The second experiment considers the role of regional accent features in the recognition of one particular speaker from Yorkshire when compared to other, unfamiliar speakers with different regional accents of British English. This experiment presents more mixed results and different results between the sexes. Among the boys, the results showed an age-related improvement. Girls with parents from outside of Yorkshire performed better than those with parents from Yorkshire. Overall, the children were more likely to misidentify an unfamiliar speaker as the familiar speaker if the speakers had a similar regional accent. Furthermore, a disguise in accent by the familiar speaker resulted, as expected, in more children being unable to identify her. The results are discussed in terms of memory models of speaker recognition.

1. Introduction

Studies show that infants 4–5 months old demonstrate a familiarity and preference not just for the voice of their mother (whose speech they are able to process faster; cf. Purhonen et al. 2005), but also more generally for their own accent over other regional accents (Butler et al. 2011). Furthermore, when it comes to word learning, infants have been found to only recognise familiar words in their local accent, before being able to extrapolate this across other regional accents as they grow older (cf. Schmale et al., 2010) whose study shows an improvement in this ability between 9-month-olds and 12-month-olds and Best and Kitamura (2012) who show an improvement between 15-month-olds and 19-month-olds). These studies therefore indicate the role of accent in speech perception and speaker recognition from an early stage in a child's linguistic and cognitive development.

However, beyond the age of infancy (if we take this roughly to mean post-18 months), very little has been done to investigate the further development of both speaker and accent recognition and the link between these perceptual skills in early childhood. While a certain level of familiar speaker recognition amongst pre-school children has been established (Bartholomeus 1973; Spence et al., 2002), an issue of central concern, which these studies do not address, is which particular aspects of the voice the children are using in order to recognise and identify a certain speaker. Both these studies use a large range of speakers (around 20 in each case) and utterances of 12 words or 4 seconds long. Therefore, each child is likely to have heard a varied range of voice qualities, pitches and accent differences (amongst other speaker-specific distinctions) during these experiments. However these differences are not explored in either study and therefore there is no way of knowing whether any or all of these aspects of the voice were important in the children being able to identify the speakers.

The present study focuses on analysing speaker differences in addressing the question of pre-school children's ability to identify familiar speakers and the role of regionally-based phonetic accent features in this process. Two related experiments were carried out in order to investigate the different aspects involved in familiar speaker recognition. First, experiment 1 was carried out in order to establish the level at which the pre-schoolers were able to identify familiar speakers. In this experiment, named the 'Identification experiment', the children were asked to identify familiar speakers in a voice-face match procedure, and the results address the following questions:

1) Are pre-school children from the age of 2.5 years able to identify familiar nursery teachers from short audio stimuli?
2) Does this ability improve with age throughout the pre-school years?
3) Is this ability affected by the amount of exposure the child has had to the speakers?
4) Does this ability improve when the children hear a longer length audio stimulus?
5) Are particular speakers, with more idiosyncratic features of voice such as their pitch or voice quality, more recognisable than others?

With the children's ability to identify familiar speakers established from the results of experiment 1, experiment 2 leads on from this to look at the role of particular phonetic accent features in speaker recognition. In this experiment, named the 'Recognition experiment', children were asked to give yes/no answers as to whether the speaker was recognised as a particular nursery teacher or not. Following on from the Identification experiment, the Recognition experiment addresses the questions:

6) Are the same children able to recognise one particular familiar nursery teacher (with a strong regional accent) from single word stimuli?
7) Is this ability influenced by external social factors, such as the child's age/sex or their exposure to different languages and/or accents at home?
8) Are the children able to distinguish other, unfamiliar speakers from this familiar speaker based on the same single word stimuli?
9) Are some unfamiliar speakers with different regional accents more easily distinguished from the familiar speaker than others?
10) Does a disguised accent result in a reduced ability to recognise the familiar speaker?

2. *Previous work*

Speaker identification relies on both segmental and suprasegmental linguistic cues such as a speaker's pronunciation, pitch, prosody and voice quality (cf. Nolan, 1983). The extent to which these different cues are used depends on factors such as the wider context in which the speech is heard, the familiarity of the listener with the speaker, and the presence of idiosyncrasies in the voice of the speaker. Depending on how familiar the speaker is to the listener, social information associated with particular linguistic cues can be relied upon in the identification process. A speaker may be instantly recognised as male or female, old or young, middle class or from a particular part of the country. This may, in turn, help a listener to identify a particular speaker who they know from that particular social group. Furthermore, this kind of 'personal' information has been found to be intrinsically linked to linguistic information. Nygaard and Pisoni (1998) found that 'talker-specific information' helped listeners in word learning tasks. Listeners performed better in intelligibility tests involving novel words when the words were spoken by speakers they had been familiarised with during the testing procedure. Therefore, there seems to be a strong likelihood that listeners store social information alongside linguistic information in the speaker encoding process. Identification of a more long-term familiar speaker, however, relies more heavily on processes of pattern recognition (Van Lancker & Kreiman, 1985, 1987). This kind of recognition and identification has been shown to activate different regions of the brain compared with unfamiliar speaker discrimination (Van Lancker & Kreiman, 1985, 1987; see Öhman, 2013 for an overview of such studies). Therefore it is important not to conflate findings from studies on familiar speaker recognition and those from speaker identification more generally. Overall, though, more exposure, more attention and/or recent exposure to a particular speaker results in a higher level of accuracy in speaker identification (Watt, 2010). Therefore it is unsurprising that it seems on balance a lot easier to identify speakers we are more familiar with.

In terms of the role that a speaker's accent plays in the identification process, Nolan (1983) suggests that an accent cannot be properly defined outside of the speaker who uses it. He proposes that 'personal' and 'accentual' information in speech are intertwined with one another. Therefore an individual's use of accent features depends on the style of speech they are employing. Thus, the more familiar we are with a speaker, the more likely we are to know their stylistic range. This tendency was investigated by Ladefoged and Ladefoged (1980), who tested one of the authors' ability to identify familiar speakers from a sample of 29 friends, family and acquaintances. Overall he was able to identify 83% of the speakers with 30 second samples of speech but this was reduced to 66% when hearing just a single sentence, and to 31% when the single word 'hello' was heard. One surprising finding in this study was that Ladefoged failed to correctly identify his own mother from both the one-word and one-sentence samples. On reflection, Ladefoged highlights the role of accent here, as he struggled to differentiate his mother's voice from those of the other similarly-accented RP speaking women in the sample. In this case, then, accent had a negative impact on his ability to identify a speaker relative to the other possible candidates in the task. Accent can have the opposite effect, however, and can make a speaker noticeable if they have a different accent to others in a group. In the same study, Ladefoged specifically refers to hearing 'North Country vowels' and then being able to identify the speaker in the sample he knew came from the North. Also, Foulkes and Barron (2000) found that within a group of 10 young male friends of diverse backgrounds, the two most consistently identified were those with the strongest accents (Tyneside and London) relative to the group. These findings suggest that accent is the most salient feature of a speaker's voice in the recognition process.

Sjöström et al. (2009) present further evidence for the important role of accent in their study of a bidialectal speaker who was not recognised across his two Swedish dialects. Participants were familiarised with the speaker in one of his two dialects by listening to a reading passage. They were then asked to identify the speaker from a voice line-up, with four foil voices alongside the target heard saying phrase-length utterances. When the target speaker shifted dialects, listeners found it much more difficult to identify him than when he was heard speaking in the same dialect as the reading passage. Features of a speaker's dialect or accent, therefore, do appear to be playing a role in the identification and recognition process. However, the questions of which specific features and to what extent they contribute to identification have yet to be fully addressed.

Few studies have looked at young children's abilities to identify familiar speakers. Bartholomeus (1973) tested 4-year-old nursery children in a voice-face matching task involving other children and staff at the nursery. During the testing procedure the children heard a speech sample and were asked to choose the corresponding picture of the speaker from a choice of 21. Each child took part in 17-19 trials altogether and overall the children scored just under 60% correct. Spence et al. (2002) found that 3-5-year-old children scored above chance in matching a cartoon voice with its picture. The children listened to 20 voice samples and chose from a closed set of 6 pictures after each stimulus. They found an improvement in this ability between 3-year-olds (61% correct overall) and 4-year-olds (81% correct overall). Also, they found that there was a significant improvement in line with the level of familiarity as the children were more likely to match the voice with its picture when the character was more familiar to them. Although these studies show a certain level of familiar speaker identification by young children, they do not further investigate what aspects of a familiar speaker's voice the children are using in the identification/recognition process. Whether the speaker's accent may play a role, like it has been shown to for adults, is still a question for further investigation.

The extent of the awareness of accent differences by children in this age group is a contentious issue. Nathan et al. (1998) found an improvement from a group of 4-year-olds to a group of 7-year-olds (all from London) in a task testing their lexical comprehension of accents. This was based on the children's own reproductions of words they heard in a Glaswegian accent. Conversely, Floccia et al. (2009) found that 5-year-olds were unable to group speakers into a local vs. non-local group based on their accent. Children were asked to listen to sentence length utterances by speakers with an accent from their own region (West Country) and an 'alien' region (Ireland). They were then instructed to divide these speakers into a blue or a red group, depending on whether the speaker was from their home town (Plymouth) or was an alien from somewhere else. The 5-year-olds scored around chance level (just over 56%) in this task. However, some methodological issues with this study means that the ability of 5-year-olds to distinguish and group speakers on accent-based criteria cannot be completely ruled out. The stimuli heard by the children consisted of full-length sentences and therefore there is no way of knowing which particular linguistic cues the children were using to group the speakers. The length of the stimuli meant that the children were listening for content as well as segmental information, making the task cognitively demanding in terms of the processing and storing of information in their short term memory (Beck, 2014).

In an attempt to address these criticisms, Beck (2014) used a different methodology in her study of 5-7-year-olds. She looked at the ability of children in this age group to discriminate between a local and non-local accent. The children took part in an ABX discrimination task in which they heard a token word, either in their native Philadelphian or a General Southern (US) accent. They then heard two more tokens of the same word (one from each accent) and were asked to choose which one sounded most like the first. The children scored above chance, with 64% correct answers overall. Beck's study shows a discriminatory ability of at least some of these young children. How this ability may play a role in identifying familiar speakers is a further question that the present study aims to address. The question is whether children use an ability to differentiate a familiar accent from an unfamiliar one in the process of identifying a familiar speaker, as adults have been shown to do to a certain extent.

3. *Identification experiment*

3.1 Methodology

3.1.1. Participants

22 children aged between 2 and 4 years (2;4 to 4;10) all attendees at the same nursery in York, in the north of England, were tested. Three children were excluded due to them not responding adequately[1] during the experiment, leaving 19 altogether.

3.1.2 Background questionnaire

The children's parents were asked to answer some background questions about the languages and accents their child was exposed to at home, the length of time the parents had lived in the city of York, and how many days/hours a week their child attended the nursery.

3.1.3 Speakers and stimuli

Stimuli for the experiment were taken from all six[2] nursery teachers. The teachers all worked regularly at the nursery and so were in contact with all of the children. They were all females, aged 21-48; five had lived in York all their lives or came from elsewhere in Yorkshire (Knaresborough, Leeds). The exception was one teacher (Jane[3]) who had moved around the country quite a lot, living in the South of England and the Midlands for different periods of time.

All the teachers were recorded in a quiet room at the nursery using a Zoom H4n recorder which was set to record at a 32bit 96kHz sampling rate. All speakers were recorded producing the same stimuli, which consisted of a story passage and 15 short phrases which they were asked to read as naturally as possible in a relaxed, informal style. The story passage consisted of 177 words from the start of the children's book *The Gruffalo* (Donaldson and Scheffler, 1999), a book regularly read to the children at the nursery and written in rhyming couplets such as (1).

(1) A mouse took a stroll through the deep dark wood.
A fox saw the mouse and the mouse looked good.

The 15 phrases that were also recorded were chosen as examples of the kind of phrases the nursery teachers use on a daily basis at the nursery. They ranged from 1 to 14 syllables (see Table 1). This was in order to be able to investigate the potential effect that the length of the stimulus might have on the children's ability to identify the speaker (research question 4).

---

[1] These three children gave no response to the stimuli even after several efforts from the experimenter to help them understand and engage with the task. Therefore their experiments were stopped after attempting the first few stimuli.

[2] Recording of a seventh nursery teacher was also taken but her sample was discarded after pilot testing the experiment. She left the nursery part-way through the experimental design process and children in the pilot study failed to recognise her, even from her picture.

[3] Pseudonyms are used throughout

3.1.4 Experimental design

Audacity software (Audacity Team, 2012) was used to edit the sound files into the stimuli. The Gruffalo passage was divided into seven separate and continuous stimuli, with one stimulus taken from each nursery teacher (apart from Jane who had two stimuli). Also, from the phrases recorded, one phrase was taken from each nursery teacher (apart from Leanne[4]) to be used as stimuli (see Table 1).[5]

Stimulus	Phrase	Nursery teacher
1	*Gruffalo 1*	Jane
2	*Gruffalo 2*	Wynne
3	*Gruffalo 3*	Leanne
4	*Gruffalo 4*	Kristina
5	*Gruffalo 5*	Alice
6	*Gruffalo 6*	Claire
7	*Gruffalo 7*	Jane
8	*Hello, Nursery?*	Leanne
9	*Line up to wash your hands*	Alice
10	*Daddy's here to collect you*	Claire
11	*There's a good girl*	Jane
12	*Where's teddy?*	Wynne
13	*No!*	Alice
14	*Bye!*	Leanne

Table 1: Stimuli and speakers for the Identification experiment

The experiment was constructed using the psychology software Psychopy (Peirce, 2007) and was designed in order to be run on the experimenter's laptop. Each audio stimulus was presented separately while a picture relating to its content was displayed on screen. This was in order to keep the task entertaining for the children, therefore retaining their attention. Also, a picture of all the nursery workers was displayed at the bottom of the screen throughout the experiment (see Figure 1). The experiment was designed to be manually controlled by the experimenter and each new visual stimulus was displayed when a key was pressed on the keyboard, with the audio starting after one second each time. The responses (including any null responses) were logged by the pressing of a corresponding answer key. The stimuli were presented in the same order for each child as they followed a chronological order.

[4] Jane's second Gruffalo passage and Leanne's extra phrase replaced the missing seventh nursery teacher's stimuli
[5]The original list of nursery phrase stimuli was shortened (from 15 to 7) as piloting the experiment revealed that the children found it difficult to concentrate for the full length of time.

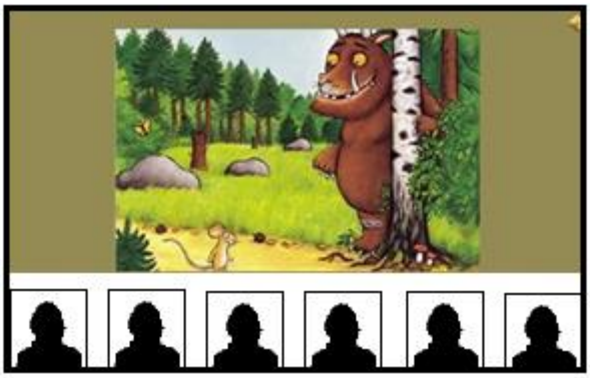

Figure 1: Screen shot from the Identification experiment

3.1.5 Experimental procedure

The children were tested individually in a quiet corner of the nursery, on the experimenter's laptop. Headphones designed for use by children (JVC HA-KD5-Y –E) were worn by the children for the duration of the experiment. The experimenter also wore headphones (SONY MDRXB400W) in order to monitor the experiment. Additionally, the experimenter had a microphone to talk through which was connected to the computer via an amplifier mixer. This allowed the experimenter to be heard by the child in order to give instructions and prompt him or her for a response if necessary. The use of this audio equipment also helped to minimize distraction from the rest of the room.

Each child was asked to sit at the computer with the experimenter and was given the instructions *'You are going to hear your teachers reading a story and I would like you to point at the picture of who you think is talking.'* As the child and the experimenter were wearing headphones, this allowed the child to indicate a response without having to speak. Also, this ensured that s/he was matching voices and faces, a much easier task than having to name the speaker outright (as found by Bartholomeus, 1973).

Before the experiment started, the experimenter checked that the child recognised all of the nursery teachers by pointing at their pictures and asking the child to name them. This ensured that all of the children reached a baseline level of familiarity with the speakers. The experiment was started and the first audio clip played. If the child did not initiate pointing at one of the nursery teachers, the experimenter prompted them: *'Who was that speaking? Can you point to their picture?'*

The child's responses were recorded on to the computer by the experimenter and then the child was asked to do the same again with the rest of the stimuli. At the end of this task, the headphones were removed in order to give further instructions for the Recognition experiment (see section 4 below).

3.2 Identification experiment results

3.2.1 Results overall

Table 2 displays each child's overall number of correct answers for the Identification experiment. Background information pertaining to the amount of time the children have spent at the nursery as well as the languages spoken at home is also shown in the table.

Overall, the children's mean score was 8.53 out of 14 (SD = 3.64), equating to 60.9% accuracy (chance = 16.7%).

Child	**Age (years)**	**Correct answers (/14)**	**Correct answers (%)**	**Years at nursery**	**Hours a week at nursery**	**Languages spoken at home**
F9	3.8	14	100.0	3	45	2
M7	4	13	92.9	2	12	1
F10	4.9	12	85.7	2	15	2
F12	3.8	12	85.7	3	36	2
F8	4.8	12	85.7	4	17	1
F11	2.4	10	71.4	1.5	30	1
F13	2.8	10	71.4	2	16	1
F2	3.8	10	71.4	4	24	1
F6	4.6	9	64.3	2	10	1
M1	3.6	9	64.3	3	32	2
M3	3.6	9	64.3	2	20	1
M4	4.3	9	71.4	2	9	1
F3	3	8	57.1	2	22	1
M5	4.3	7	50.0	3	15	1
F4	4.5	6	42.9	2.5	5	1
M8	4	6	42.9	2	24	1
M6	2.7	3	21.4	2	17	1
M2	2.5	2	14.3	1.5	24	1
F7	3.3	1	7.1	1.5	27	2

Table 2: Identification experiment results (scores in rank order) and background information for each child: F=Female, M=Male

3.2.2 Statistical analysis of Identification experiment results

A stepwise backward regression method was used in a binary, mixed effects logistic model, run in R using the lme4 library (R Core Team, 2013). A logistic model is used when there are two possible outcomes, in this case a correct or incorrect answer, in order to show the likelihood that the predictor variables entered into the model predict a particular outcome (Baayen, 2008). A mixed effects model includes random variables as well as fixed predictor variables in order to account for individual variation. In this model, each individual's responses are represented by their own coefficient, ensuring that the potential for one particular individual's results to warp the overall results is greatly reduced (Drager, 2011).

The children’s responses, correct or incorrect, were turned into a factor so that each response from each individual was used in the model analysis. The dependent variable was therefore a binomial factor distinguishing between correct and incorrect answers, with the default set to correct answers. This means that the model presents the log odds of a correct answer. Independent, fixed predictor variables of stimulus length, years at nursery and hours at nursery were included in the model. All of these were measured and entered as continuous variables. Additionally, child was treated as a random factor in the model. Two-way interactions between all the fixed predictors were also included. Age was not included as a fixed variable as it was found to correlate highly with the children’s years spent at nursery. Therefore the effects of these predictors cancelled each other out when entered into the model together. As the children’s years spent at nursery was a more explanatory predictor in pre-testing (as judged by both a lower AIC and BIC which tests the likelihood of a model being closer to the truth based on both goodness-of fit and the complexity of the model), this was included in the model at the expense of age.

Only the interaction between stimuli length and hours spent at the nursery did not significantly affect the predictive power of the model. Therefore this was the only interaction removed from the final model (see Table 3).

	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**	**Sig.**
(Intercept)	5.810	3.294	1.764	0.078 .	.
STIMULUS LENGTH	-0.494	0.265	-1.865	0.062 .	.
YEARS SPENT AT NURSERY	-2.709	1.341	-2.020	0.043 *	*
HOURS A WEEK AT NURSERY	-0.269	0.144	-1.862	0.063 .	.
STIMULUS LENGTH*YEARS	0.334	0.124	2.706	0.007 **	**
YEARS*HOURS	0.116	0.057	2.049	0.041 *	*

Table 3: Mixed effects logistic regression model fit to the data for the Identification experiment (Significance level: ‘.’ = 0.1, ‘*’= 0.05, ‘**’ = 0.01)

Table 3 shows the effect of the dependent variables on the log odds of the children giving a correct answer. The variables of stimulus length and years at nursery enter into a significant interaction, as do years at nursery and hours spent at nursery every week. Figure 2 plots the significant interaction between the length of the stimulus and the number of years that the children have been attending the nursery. As the majority of the children had attended the nursery for 2, 3 or 4 years, these are the years plotted. This shows that while both a longer stimulus and more years at nursery result in a higher probability of a correct answer from the children, these are not independent of one another. As the number of years the children have been attending the nursery increases, the length of the stimulus has a stronger effect on their predicted ability. In other words, for the lowest performing younger children who have been attending the nursery for fewer years, the length of the stimulus has some effect. This effect is strengthened for the older children, who have been attending the nursery for more years and who perform better overall.

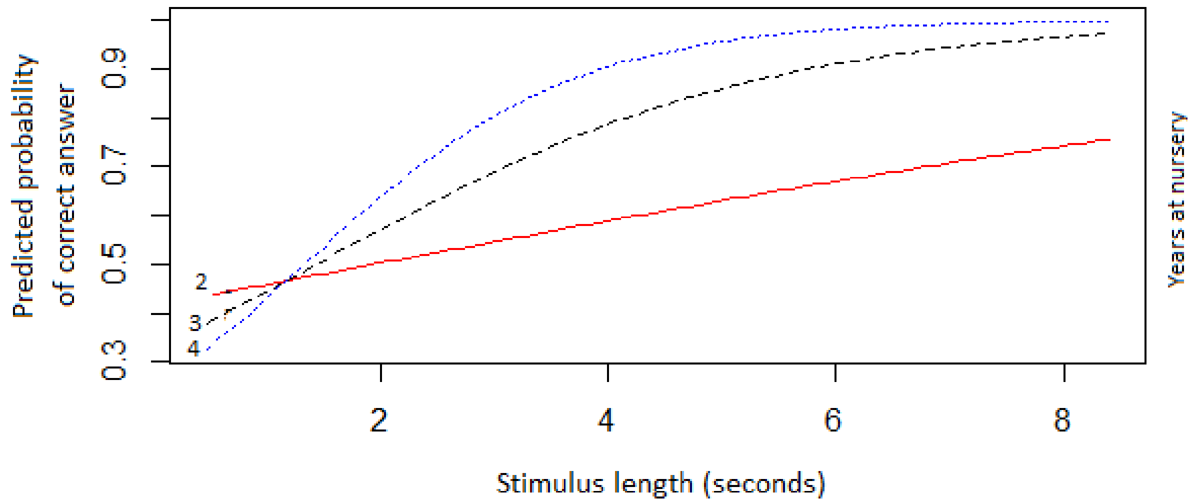

Figure 2: The interaction between stimulus length and the years the child has been attending the nursery.

Figure 3 plots the significant interaction between the number of years that the children have been attending the nursery and the number of hours a week that they attend the nursery. The line plots the predicted probability of a correct answer. For example, 0 means an incorrect answer is predicted and 1 that a correct answer is predicted, dependent on the variables of the child's years at nursery and the hours a week they spend at nursery. Those who have attended the nursery for longer are more likely to give a correct answer overall. Of the children who have been attending the nursery for 3 or 4 years, those who spend more hours per week at the nursery have a higher probability of achieving a correct answer in this task. The children who have been attending the nursery for 2 years however are not showing the same pattern of improvement, possibly due to a couple of low-scoring individuals driving this effect (cf. M2 and M6 in Table 2).

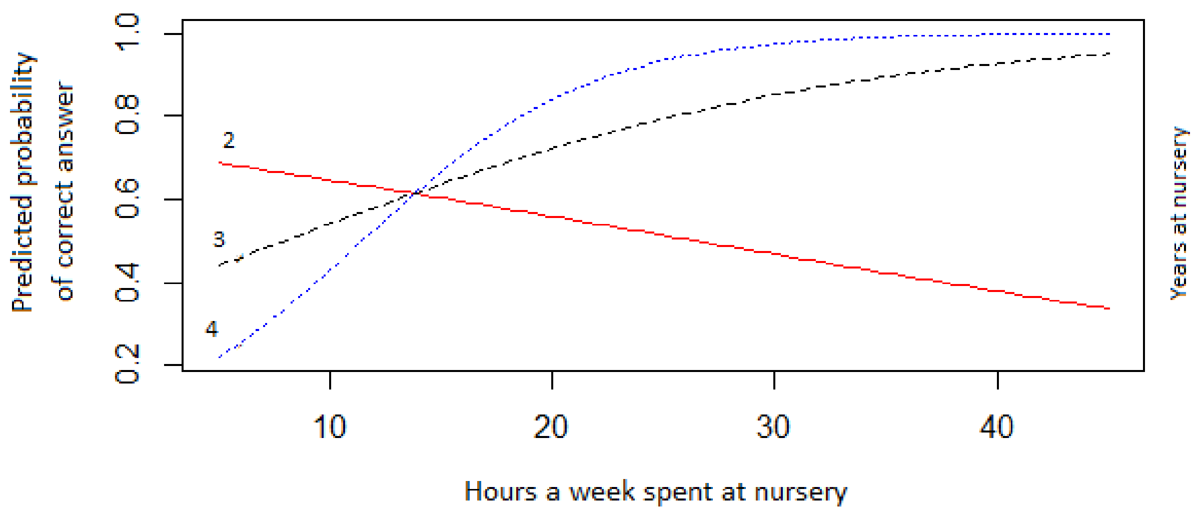

Figure 3: The interaction between the number of hours the children attend the nursery a week and the number of years they have been attending the nursery altogether.

3.2.3 Speaker effects

The rate of identification of each individual speaker was further analysed. This was in order to address research question (5).

Certain voices were more readily recognised by the children than others. Table 4 shows the number of children who correctly identified each speaker for each of their stimuli.

Stimulus	Number of children correct (/19)	Speaker
1	16	Jane
2	16	Jane
3	15	Kristina
4	15	Alice
5	14	Leanne
6	14	Claire
7	13	Claire
8	12	Jane
9	12	Leanne
10	11	Alice
11	10	Leanne
12	8	Wynne
13	5	Wynne
14	2	Alice

Table 4: Number of children with correct answers for each speaker

Excluding Kristina, as she had only one stimulus token as opposed to two or three from all the others, Jane was the most often recognised with an average of 14.67 correct responses. Wynne was the least well recognised with an average of 6.5 correct responses. As Jane's voice stood out as being much lower in pitch than the others, this was taken as the basis for further investigation. The teachers' fundamental frequencies were measured to explore the effect of this feature in more detail across the whole group of speakers (cf. Foulkes and Barron, 2000 and Compton, 1963 for previous work showing f0 to be one of the most readily distinguishable characteristics of a voice). F0 was measured by calculating each speaker's mean fundamental frequency across all of their stimuli used in the experiment.

The box plot in Figure 4 shows each speaker's average f0 (in Hz). From this figure it is clear that Jane's average f0 is conspicuous. Whereas Jane's f0 falls near the typical male average of 120Hz, the other speakers are around the female average of 225Hz (cf. Fry, 1979: 68). Although Leanne's pitch range is also noticeably different to the others, as it is above rather than below the average pitch for a woman, it is arguably less conspicuous as her voice is obviously female sounding.

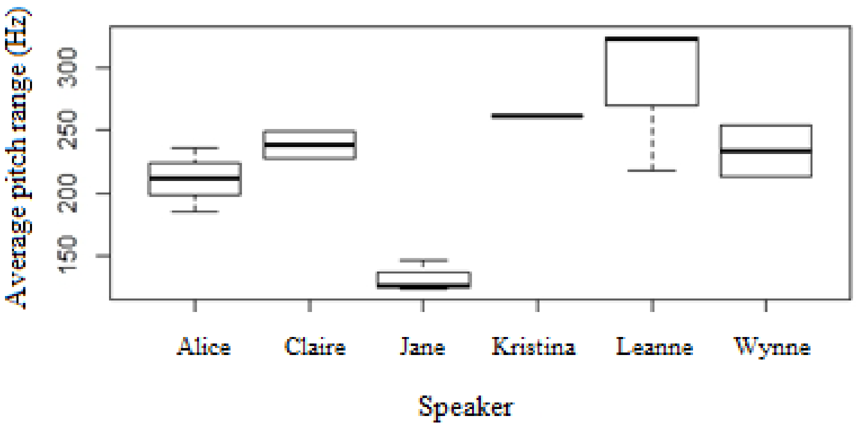

Figure 4: Each nursery teacher's average fo (pitch) range over the stimuli used in the Identification experiment

Jane also has a very distinctive voice quality; which under Laver's VPA protocol might be classified as a deep, harsh, whispery, creaky voice (Laver, 1968: 48). Therefore altogether it appears that Jane has a relatively idiosyncratic f0 and voice quality, which can both be predicted to mark her out as easier to identify than the other speakers.

3.2.4 Discussion of Identification experiment results

In addressing the research questions (1)–(5), results from the Identification experiment show that pre-school children are able to identify six individual familiar speakers with a high rate of accuracy. Overall, they performed well above chance level (with only two children below chance) and at a very similar level to the children in Spence et al.'s (2002) study. Also, similarly to Spence et al., there appears to be a developmental effect, as the older children outperform the younger children. However, this improvement seems to be mostly affected by the amount of exposure the children have had to the speakers. This exposure is both in terms of the years that they have been attending the nursery combined with their regular attendance at the nursery on a weekly basis. Not surprisingly, both of these exposure criteria have an effect, in line with the obvious prediction that more exposure to a speaker results in better recognition rates (e.g. Watt, 2010). Spence et al. showed this to be true for children in terms of how familiar they appeared to be with cartoon characters (e.g. could they name the characters without prompting). The present experiment, however, shows this a little more rigorously though the measurement of exposure.

The children's ability to identify the correct familiar speaker improves with a longer audio stimulus. This confirms earlier findings from adults (Ladefoged and Ladefoged, 1980; Schweinberger et al., 1997). In turn, this finding suggests that better identification relies on a number of different cues, as a longer stimulus is also likely to provide more acoustic indications pertaining to the identity of the speaker. In terms of these cues, f0 and voice quality were found to play a role in the success of the best identified teacher. Jane's overall f0 and voice quality were very distinct from the other five nursery teachers, demonstrating the importance of idiosyncrasies in a speaker's voice in the identification process.

The next question this research addresses is whether accent is also a factor in recognising a familiar speaker from shorter stimuli.

4. Recognition experiment

4.1 Methodology

4.1.1 Participants

The same children from the Identification experiment were tested. The youngest two children were excluded due to not responding to this second experiment. This left 17 children (aged 2.7-4.9 years) altogether.

4.1.2 Stimuli and speakers

Stimuli used for the experiment were taken from recordings of one of the nursery teachers (Alice) and nine other speakers, unknown to the child participants. These nine speakers were a mixture of females (7) and males (2), aged 20-29 with different regional accents. Alice was chosen as she was the teacher judged impressionistically to have the strongest Yorkshire accent. For example, she had a consistently monophthongal vowel quality in the GOAT and FACE lexical sets (Wells, 1982). Therefore, being the most distinctly 'Yorkshire', her accent provided the clearest benchmark for comparison with the other broad accents of the unfamiliar speakers. Also, she was of a comparable age (21 years) to the unfamiliar speakers recorded. The speakers were all recorded reading the same story passage and phrases as those in the Identification experiment.

4.1.3 Experimental design

Eight words were chosen from the recordings to use as stimuli. These words captured differences between the Yorkshire accent and the accents of the other speakers recorded. Seven of these were based on vowel quality differences and one based on the rhotic/non-rhotic accent distinction. These words were chosen based on the main differences between the regional accents of the unfamiliar speakers compared to Alice. The decision of which words to use was also largely based on the quality and clarity of recordings once the individual words were extracted and heard in isolation.

For each word, three tokens were extracted as stimuli: one token from Alice and one each from two unfamiliar speakers with different accents. Seven of the other speakers had distinctly different regional accents compared with Alice (two Standard Southern British English, one North American (Californian), one North East of England, two Northern Irish, one Scottish) while two of the speakers had similar accents to Alice's. These two were another Yorkshire speaker and a speaker from Lancashire, which, as another area of the central north, shares many phonetic based accent features with Yorkshire (Hughes et al., 2012). Stimuli taken from these two speakers were therefore used in order to investigate whether the similarity in their pronunciation to Alice's would cause the children to misidentify these speakers as Alice.

Additionally, Alice was recorded saying two of the stimuli for a second time (*gruffalo* and *coat*). She was asked to disguise her accent by pronouncing these words with an accent from the South of England, using the diphthong [əʊ] as opposed to the monophthong [o:]. These recordings were taken in order to investigate whether Alice would be miscategorised as an unfamiliar speaker when using the phonetic realisation of a different accent from her own. Table 5 shows the full list of stimuli and each speaker's phonetic realisation of the vowel quality/consonant distinction under consideration. Alice’s disguised tokens are indicated with *.

Word	Associated lexical set/ feature of accent	Alice: Phonetic realisation	Unfamiliar speaker 1: Accent and phonetic realisation	Unfamiliar speaker 2: Accent and phonetic realisation
gruffalo	GOAT	[oː]	SSBE M [əʊ]	Lancashire F [oː]
never mind	rhoticity	no /r/	Northern Irish F /r/ realisation	American F /r/ realisation
know	GOAT	[oː]	SSBE F [əʊ]	North East F [ɔː]
fox	LOT	[ɒ]	Scottish F [ɔː]	American F [ɑː]
mouse	MOUTH	[aʊ]	Northern Irish F [aʉ]	Scottish F [au]
food	GOOSE	[ʉː]	Northern Irish F [ʉ]	Scottish F [ʉ]
coat	GOAT	[əʊ] *	SSBE F [əʊ]	Yorkshire F [oː]
good	FOOT	[ʊ]	Northern Irish F [ʉ]	Scottish F [ʉ]
gruffalo	GOAT	[əʊ] *	SSBE F [əʊ]	Yorkshire F [oː]

Table 5: Stimuli used in the Recognition experiment, along with each speaker’s phonetic realisation of the vowel/consonant distinction. SSBE= Standard Southern British English, M = male, F= female

4.1.4 Experimental procedure

After finishing the Identification experiment and in preparation for the Recognition experiment, the children were told, *‘Now you are going to hear some more people speaking. It might be Alice that you hear, or it might be someone that you don’t know. Each time you hear a voice, I would like you to tell me whether you think it is Alice talking or not. If you think it is Alice, say ‘Alice’ but if you think it is someone else that you don’t know, say ‘No’, okay? Let’s have a go.’* The headphones were then again placed over the child’s head.

A picture of Alice (with the headline ‘Is it Alice?’) appeared on screen to remind the children who they were listening out for. Then the screen was changed to a picture of a teddy bear and the child heard the first stimulus. They were asked *‘Was that Alice speaking?’* and their response was logged by the experimenter.[6] Next, a different teddy bear appeared on screen and a different speaker was heard saying the same stimulus. Again, the children were asked if this was Alice speaking. The process was repeated a third time with a third teddy bear and a third speaker (see Figure 5). This whole process was then repeated for each set of three stimuli. Two versions of the experiment were created, with different orders of the speakers in each set of three; 8 children took version 1 of the experiment and 9 different children took version 2. Children were rewarded with stickers for playing the game.

[6] No trial runs were included in the experiment as piloting the experiment found that the children were able to understand the task but struggled to concentrate if the task was any longer than the final version used for the experiment.

Figure 5: Screens showing the procedure for the Recognition experiment

4.2 Recognition experiment results

4.2.1. Results overall

Table 6 presents each child's results along with relevant background information. Overall, the children's mean score was 21.53 out of 33 (SD = 6.06), equating to 65.1% accuracy (chance = 50%).

Child	Age (years)	Correct answers (/33)	Correct answers (%)	Years at nursery	Hours a week at nursery	Languages spoken at home	Yorkshire parents
F9	3.8	30	90.9	3	45	2	0
F2	3.8	29	87.9	4	24	1	0
F13	2.8	28	84.8	2	16	1	1
F10	4.9	27	81.8	2	15	2	0
F3	3.0	27	81.8	2	22	1	1
M7	4.0	27	81.8	2	12	1	1
F8	4.8	26	78.8	4	17	1	0
F12	3.8	24	72.7	3	36	2	0
F6	4.6	21	63.6	2	10	1	0
M5	4.3	20	60.6	3	15	1	1
M1	3.6	17	51.5	3	32	2	1
M8	4.0	17	51.5	2	24	1	0
F4	4.5	16	48.5	2.5	5	1	2
M3	3.6	16	48.5	2	20	1	1
M4	4.3	16	48.5	2	9	1	1
F7	3.3	14	42.4	1.5	27	2	1
M6	2.7	11	33.3	2	17	1	0

Table 6: recognition experiment results and background information for each child

A post hoc comparison showed no difference between the two versions of the experiment, therefore the results from both versions were combined for the analysis. A significant positive correlation was found between the children's scores for the Identification experiment and those for the Recognition experiment ($r = 0.641$, $p<0.01$), indicating that those who performed well in the first experiment were more likely to score well in the second.

In addressing research question (6), the overall results show that the children's performance was above chance and therefore to some degree they were able to recognise Alice and distinguish the other unfamiliar speakers from Alice. Questions (8)-(10) will be further investigated in section 4.2.3 which will take a closer look at the cases in which Alice was missed and unfamiliar speakers were misidentified as Alice. First, question (7) (repeated below) is addressed in the next section.

(7) Is the children's ability to recognise one familiar nursery teacher influenced by external social factors, such as their age/sex or their exposure to different languages/accents at home?

4.2.2 Statistical analysis

As with the Identification experiment, statistical analyses were carried out in R. A stepwise backward regression method was used in a binary, mixed effects model. Again, the dependent variable was a binomial factor distinguishing between correct and incorrect answers, with the default set to correct answers. Main effects predictors of age, sex and Yorkshire parentage were included in the model[7]. Age was measured and entered into the model as a continuous variable. For the Yorkshire parent predictor, the default was set to no Yorkshire parents and for the sex predictor, the default was set to female. Therefore the coefficient in the 'estimate' column shows the effect of each of the default dependent variables on the log odds of the children scoring a correct answer. Two-way interactions between all the main effect predictors were also included. Only the interaction between age and Yorkshire parents was not found to significantly alter the predicting power of the model. Therefore this was the only interaction removed from the final model (see Table 7 below). As the interactions between the predictor variables are found to be significant, the main effects are interpreted as part of these interactions rather than individually.

	Estimate	**Std. Error**	**z value**	**Pr(>|z|)**	**Sig.**
(Intercept)	4.41	1.31	3.37	0.0007 ***	***
AGE(Years)	-0.70	0.30	-2.37	0.02 *	*
SEX(Male)	-6.65	2.02	-3.3	0.00097***	***
YORKSHIREPARENT(True)	-1.39	0.43	-3.22	0.0013**	**
AGE(Years):SEX(Male)	1.27	0.53	2.4	0.02 *	*
SEX(Male):YORKSHIREPARENT(True)	1.71	0.66	2.61	0.009 **	**

Table 7: Mixed effects logistic regression model fit to the data for the Recognition experiment (Significance level: '*'= 0.05, '**' = 0.01, '***' = 0.001)

Table 7 shows that the final model finds two significant interactions. Figure 6 plots the significant interaction between the children's sex and whether they have any parents from Yorkshire. The model shows that girls generally outperform boys but this is also dependent on whether they have parents from Yorkshire. The model predicts that for girls if they have at

[7] Due to the limited number of participants in the experiment, only three predictors were included in the model. The predictor pertaining to languages spoken at home was left out of the model as this did not show a strong or significant correlation with the children's correct answers. Additionally the children who came from bilingual homes varied in terms of their exposure and fluency with the other languages themselves.

least one parent from Yorkshire, the likelihood of getting a correct answer decreases. For boys the opposite is true but this is over a much smaller range and their probability of getting a correct answer stays below 0.6.

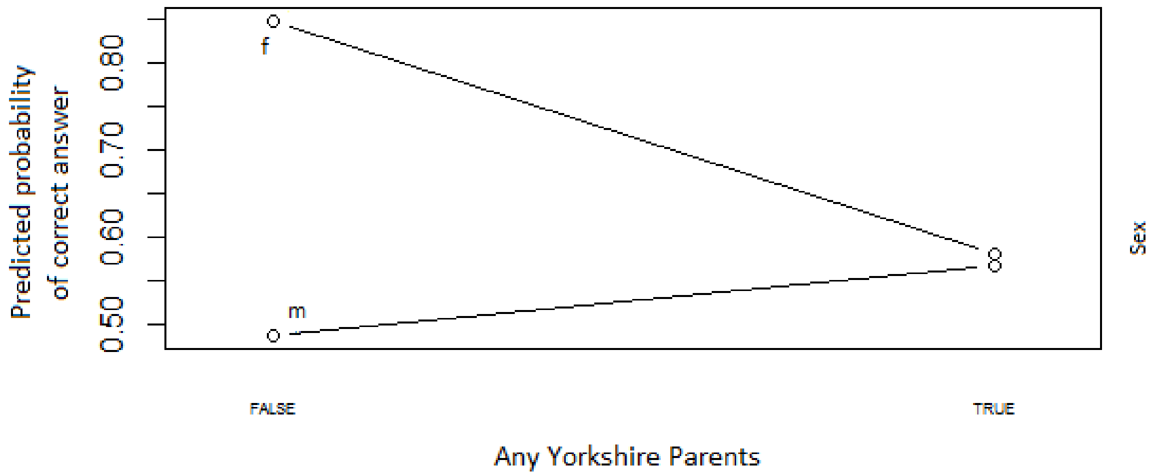

Figure 6: Interaction of sex and Yorkshire parents

Figure 7 shows the interaction of age and sex. Older boys are predicted by the model to perform better than younger boys. A higher performance is maintained for the girls, although there is a slight drop over the ages. This is probably caused by a few individuals as there is a particularly high scoring 2-year-old and 3-year-old (F13 and F3) and a low scoring 4-year-old (F4) (see Table 6). Interpreting Figures 6 and 7 together it is clear that while sex is an important predictor it is not something that can be analysed independently of other factors. With a small sample size it is impossible to be conclusive but it appears that girls are more advanced generally and while age affects the boys' results, exposure to other accents has more of an effect on the girls' results.

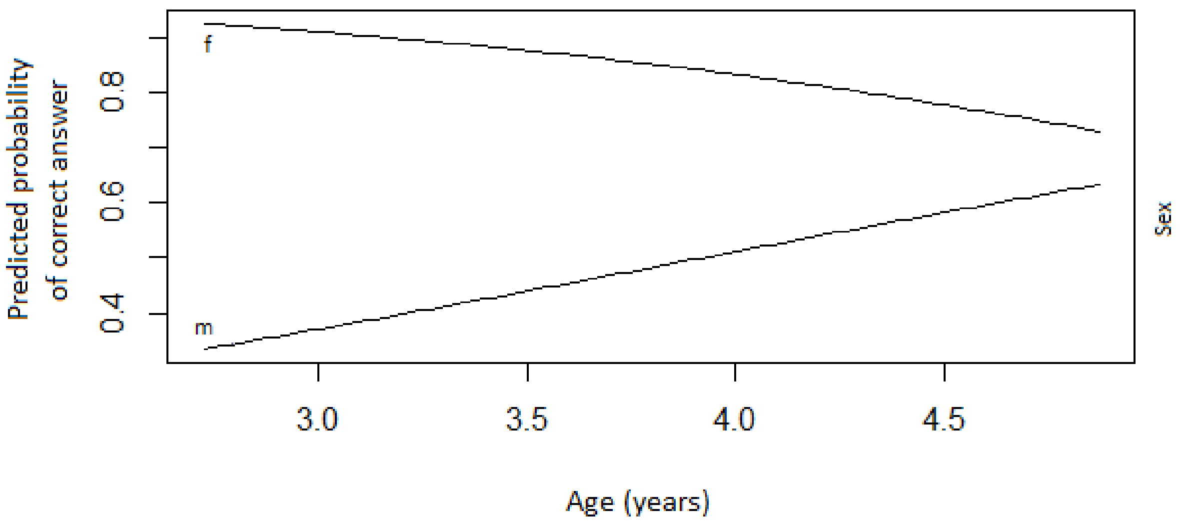

Figure 7: Interaction of age and sex

4.2.3 Misses vs. false alarms

In further analysing the results from the Recognition experiment, the seven children who scored below chance level (50%) were excluded[8]. This leaves the results of ten children to examine further.

As Table 8 shows, children were more likely to miss the familiar speaker (miss) than wrongly identify an unfamiliar speaker (false alarm).

Error type	Number of tokens	Percentage of tokens
Miss	44/110	40%
False alarm	28/220	12.7%

Table 8: Number and percentage of error types for the Recognition experiment

4.2.3.1 False alarms

Table 8 shows that the overall false alarm rate is low. In answering research question (8) this indicates that, in general, the children are able to distinguish the unfamiliar speakers from Alice. However, in addressing question (9), it is noticeable that the speakers with the highest number of false alarms are the Yorkshire and Lancashire female speakers.

Lexical set	Word	Unfamiliar speaker 1: phonetic realisation	No. of false alarms	Unfamiliar speaker 2: phonetic realisation	No. of false alarms
GOAT	*gruffalo*	SSBE M [əʊ]	1	Lancs F [əʊ]	3
rhoticity	*never mind*	N Irish F [ɹ]	2	American F [ɹ]	1
GOAT	*know*	SSBE F [əʊ]	1	North East F [oː]	0
LOT	*fox*	Scottish F [ɔː]	1	American F [ɑː]	1
MOUTH	*mouse*	N Irish F [aʉ]	1	Scottish F [au]	1
GOOSE	*food*	N Irish F [ʉ]	2	Scottish F [ʉ]	0
GOAT	*coat*	SSBE F [əʊ]	2	Yorkshire F [ɵː]	3
FOOT	*good*	N Irish F [ʉ]	2	Scottish F [ʉ]	1
GOAT	*gruffalo*	Yorks F [ɵː]	3	SSBE F [əʊ]	2

Table 9: False alarms for each unfamiliar speaker

In relation to question (7), this suggests that the children were more likely to misidentify an unfamiliar speaker as Alice if they heard a female speaker with a similar regional accent and

[8] As this part of the analysis was based on individual responses and due to the relatively low number of children who took part in the Recognition experiment, this was a decision made by the experimenter, after careful consideration of each child's performance and whether their contribution could justifiably be seen as showing an understanding of the task.

vowel pronunciation. Although the Lancashire female pronounced the GOAT vowel with a diphthong rather than a monophthong, this diphthong has less formant movement in the offglide than the diphthongs pronounced by the other unfamiliar speakers. This is evident by comparing the Euclidean distance between the F1 and F2 formants of all the stimuli featuring a GOAT vowel. The Euclidean distance between the nucleus value and the offglide value was calculated using the equation in (11):

$$11)\ \sqrt{(F1nucleus - F1offglide)^2 + (F2nucleus - F2offglide)^2}$$

The vowels were hand measured at the 20% and 80% vowel duration points in Praat (see Figure 8). Figure 8 shows that the SSBE speakers had a high level of offglide movement, suggestive of highly diphthongal [əʊ] vowels. The familiar speaker had a much lower level of movement, indicating a monophthongal vowel. As well as monphthongal, this vowel is a fronted GOAT vowel, sounding more like [ɵ:] than its backed equivalent [o:] (see Watt & Tillotson, 2001 and Haddican et al., 2013 for accounts of GOAT fronting in Yorkshire English) . The unfamiliar Yorkshire speaker's GOAT vowels were similarly monophthongal and fronted and while the Lancashire speaker's GOAT vowel showed more movement, this was still less than that produced by the SSBE speakers. The North East speaker also produced a monophthong with little movement, but this was not a fronted GOAT vowel, rather the back [o:] vowel. Therefore, this distinction meant that she was not mistaken for the familiar speaker.[9] The high amount of diphthongal movement of the familiar speaker's disguised GOAT tokens most likely accounts for why she is not correctly identified from these tokens by most of the children (this is further explored below).

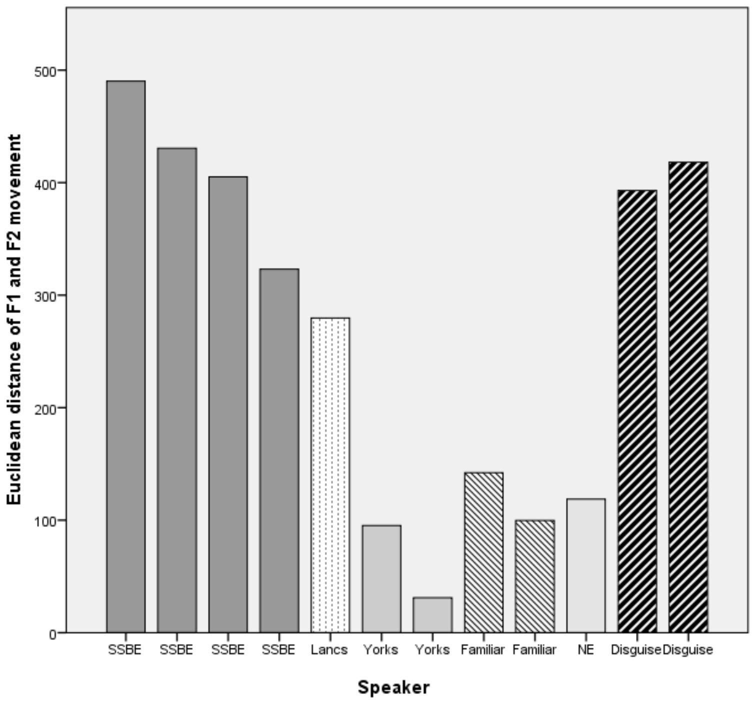

Figure 8: Euclidean distance between the F1 and F2 nucleus and offglide values of GOAT vowels

[9] See Haddican et al. (2013) for a discussion of the social awareness of GOAT fronting in York

4.2.3.2 Misses

Table 10 shows the number of misses for each word spoken by the familiar speaker. The words for which a disguised accent was used are indicated with *. Apart from *good*,[10] these words result in the most number of misses. Whereas none of the children miss Alice when she used a monophthongal GOAT vowel (in *gruffalo*), the diphthongal quality of her GOAT vowel in both the disguised words resulted in many misses. Therefore, an accent disguise appears, as predicted, to inhibit recognition for most of these children.

Word	Feature	Pronunciation	Misses (/10)
good	FOOT	[ʊ]	9
coat	**GOAT**	**[əʊ]**	**8**
gruffalo	**GOAT**	**[əʊ]**	**7**
food	GOOSE	[uː]	6
know	GOAT	[oː]	4
mouse	MOUTH	[aʊ]	3
fox	LOT	[ɒ]	1
never mind	RHOTIC	No /r/	1
gruffalo	GOAT	[oː]	0

Table 10: Misses for each word from the familiar speaker

4.2.4 Discussion of Recognition experiment results

In addressing research questions (6)-(10), results from the Recognition experiment show that pre-school children are mixed in their abilities in what was expected to be a difficult and unusual task for them. Some children were able to recognise the familiar speaker with a strong Yorkshire accent from single word stimuli while others were not. This ability is somewhat affected by external social factors. The girls generally perform better than the boys but the sexes are affected differently by external factors. The boys improve with age, indicating an important developmental effect. This development could be in terms of the understanding of the task, their level of concentration and/or their amount of exposure to the familiar speaker. Girls who have neither parent from the region performed better than those with at least one Yorkshire parent. This suggests that those who have had more exposure to different accents are better able to distinguish different pronunciations from each other. Both these findings between the sexes and whether these abilities develop differently would need to be tested with a larger dataset consisting of a greater range of ages.

The children who perform above chance level overall are good at distinguishing unfamiliar speakers from the familiar speaker, as shown by the low false alarm rate. The false alarms that do occur are mainly from the Yorkshire and Lancashire speakers who use a monophthongal pronunciation of the GOAT vowel, similar to that of the familiar speaker. An accent disguise confuses the children to the extent that many are unable to recognise the familiar speaker. Although this is similar to findings from Sjöström et al. (2009), the present study has a narrower

[10] Further work is needed to investigate the FOOT vowel and its salience in the recognition and discrimination process. Wells (1982) remarks that unlike the GOAT vowel, the FOOT vowel does not vary markedly in its use by English speakers. However this observation needs to be updated with more current production data.

focus; one-word accent-based phonetic realisations as opposed to whole phrases. Therefore this limits the range of features in the voice that the children could be using to recognise the speaker.

Wider implications of the findings from this experiment relate to how speakers' voices are stored cognitively and the kind of social information that may be stored alongside the linguistic information. This is discussed in the next section.

5. *General discussion*

Results from these two experiments show that, overall, age and exposure play a key role in children's ability to recognise and identify familiar speakers. Although this corresponds to intuitions about the nature of speaker recognition, the findings can be explained if we interpret them through cognitive models of recognition. Recognition involves long-term memory retrieval and this can manifest itself in two quite different ways. On the one hand, specific information or events may be recalled so that an instance of a particular familiar item is brought to the forefront of the mind. Alternatively, there may just be a strong sense of familiarity about something but nothing more specific than this is retrieved from memory (Reisberg, 2010). 'Dual process' theories of recognition specifically claim that these two types of memory actually involve separate memory processes. Diana et al. (2006) explain these as 'recollection' and 'familiarity' processes. The familiarity process is primarily based on the activation of conceptual information which has been stored alongside previous encounters with the familiar item. The recollection process goes further and accesses stored episodic information as well as conceptual information. This results in the retrieval of specific instances of the remembered item.

This framework can be applied to the present study, to explain the processes involved in identifying a familiar speaker. In being able to name a particular speaker, the children are recalling other stored instances of the speaker from their long-term memory and therefore accessing both conceptual and episodic information. Children's conceptual abilities are qualitatively not all that different from those of adults. For example pre-school children have been shown to rely on basic-level categories and use hierarchical categorisation just as adults do (see Murphy, 2002 for an overview of such studies). Therefore any improvement of conceptual ability with age is due to general developmental advances that come with increased experience and exposure to members of categories, a better understanding and knowledge of the world, and increased processing speed that develops with maturation of the brain (Murphy, 2002). This helps to explain the results of the Identification experiment in which the older children who have had more exposure to, and therefore experience with, the nursery teachers perform better than the younger children. In the Recognition experiment, the girls who had parents from outside of Yorkshire performed better than those with at least one parent from Yorkshire. Therefore it appears that their increased exposure to members of accent categories outside of Yorkshire increases their conceptual ability to differentiate between individuals producing regionally-based phonetic accent features.

The improvement in the Identification experiment throughout the pre-school years can also be explained by the dramatic change in episodic memory during this time (cf. Newcombe et al., 2000). The remembering of particular episodes uses explicit memory which, unlike implicit memory, is strongly age dependent. For example, Drummey and Newcombe (1995) found that 5-year-olds outperformed 3-year-olds in remembering having seen particular animal pictures three months previously. However there was no age-related improvement in the implicit memory task which involved identifying the animal from a blurry picture becoming increasingly focused.

Exemplar-theoretic accounts which rely upon the storing and accessing of episodic memories are now fairly commonplace in the speech perception and sociolinguistic literature (cf. Pisoni, 1997; Foulkes and Docherty, 2006). These accounts suppose that linguistic encounters, in the form of words, are stored in memory alongside other non-linguistic information about the encounters. This includes social information about the speaker, for example their voice quality or the particular segmental features used (Foulkes and Docherty, 2006). Therefore when particular linguistic exemplars are accessed through episodic memories, this social information is accessed as well. Which particular words are stored in this way depends on the listener's attention and the importance of the encounter to the listener at the time (cf. Foulkes and Hay in press). This account would explain the inaccuracies of the children in the Recognition experiment. When Alice's voice is disguised, the segmental phonetic information does not match closely that of the children's stored exemplars of Alice. Therefore there is an impairment in the children's ability to use the phonetic information to access the social information regarding who the speaker is. Furthermore, an unfamiliar speaker may be mis-identified as Alice if phonetic information is similar enough to activate stored exemplars of Alice, which in turn activates the social information regarding the speaker's identity. This mistaking of an unfamiliar speaker for someone familiar has been found previously in studies with adults (cf. Ladefoged & Ladefoged, 1980).

Whether exemplars of speakers are stored and accessed in this way is a matter of considerable contention. Exemplar theory acknowledges that a listener might not always rely on detailed representations of instances stored in memory but some abstraction of them. However, as Foulkes and Hay (in press) highlight, the exact nature of the relationship between exemplar-specific instances and abstractions has yet to be fully worked out. Another possibility is that listeners primarily depend on prototypes. While still reliant on experience with particular speakers, prototypes form a more abstract cognitive representation of speakers. Öhman (2013) describes Papçun et al.'s (1989) memory model of voices which suggests that an average 'prototype' of voices is used as a means of comparison with any unfamiliar voices which are heard. For familiar voices, deviations from the average prototype can be stored through experience, and it is these idiosyncracies that are searched for and accessed in the recognition process. Therefore, this theory makes a prediction that the more a particular voice deviates from the prototype, the easier it is to recognise. This could explain the results from the Identification experiment in which the speaker whose voice deviates most from the others in terms of pitch and voice quality (Jane) is the most readily recognised.[11]

Whether the memory retrieval process used in recognition is based on exemplars or prototypes is a matter for conjecture as this is very hard to test empirically. Perhaps both processes are used, as Reisberg (2010) suggests. Both models rely on the accessing of a memory and a similarity judgement. Also, both models advocate the storing of features, whether this is social information associated with particular instances or a series of deviations more generally. Therefore it is possible that prototypes are used in some cases of recognition and exemplars in others. This is likely to depend on the experience of the individual and the strength of familiarity they have for a particular memory they are accessing.

[11] Something to further investigate would be the effect of including another deep female voice in the speaker line-up. Another speaker who similiarly deviates from an average prototype would potentially be confused for Jane. Perhaps in this situation the listener would rely on other, less salient cues in order to identify the speaker as Jane over another, similar sounding speaker

Whichever model we use to interpret them, the findings from the present study indicate that linguistic information pertaining to a familiar speaker's regional background is stored in memory and somehow accessed during the recognition process. Further work is needed to investigate, as suggested by Foulkes (2010), how children advance from the storing of accent information alongside familiar speakers' exemplars to the forming of speaker groups based on social differentiations such as regional accent.

6. Conclusion

This paper has investigated pre-school children's ability to identify familiar speakers and in particular, which aspects of a speaker's voice seem to play an important role in the recognition process. The Identification experiment revealed that the familiar teacher with the most distinctive voice pitch and voice quality was most readily recognised by the children. This indicates the importance of these voice attributes in identifying an individual who has a particularly idiosyncratic voice pitch/quality. The Recognition experiment found that regionally-based phonetic accent features also appear to contribute to speaker recognition. The children were more likely to misidentify an unfamiliar speaker as their familiar teacher if they used the same phonetic accent features. Additionally, the children generally failed to recognise their teacher when she disguised her accent and used a phonetic realisation different to her own.

This study also looked at the effect of external influencing factors on the abilities of the children in both tasks. Generally, there was an improvement in both tasks throughout the pre-school years as the oldest children outperformed the youngest children. However, this was not found to be a straightforward relationship and must be considered along with other factors. In the Identification experiment, the amount of exposure the children had to the nursery teachers interacted with their years spent at nursery. This meant that overall, older children who spent more time at the nursery on a weekly basis were better able to identify their teachers. In the Recognition experiment, the children's sex seemed to have more of an impact on their ability. For the boys, there was a significant improvement in the task throughout the pre-school years, showing an age-related development. The girls, who generally outperformed the boys, did not show an improvement with age but with an exposure to different regional accents at home. This suggests a more advanced conceptual ability to differentiate speakers with different accents amongst those who have increased exposure to members of accent categories outside of their home region. Further work to investigate these findings and their implications is needed, with a larger data set across a greater range of ages.

Acknowledgements

This research is funded by the Arts and Humanities Research Council. I would like to thank Paul Foulkes, Carmen Llamas and Tamar Keren-Portnoy for feedback on earlier versions of this paper. Thanks also to Márton Sóskuthy for his help with statistics in R. I am very grateful to all those involved in the experiments; both the adults who were recorded and the children who participated. Finally, thanks to two anonymous reviwevers for their comments and advice.

References

AUDACITY TEAM (2012). Audacity(R) Version 2.0.3.

BAAYEN, R. H. (2008). *Analyzing linguistic data: A Practical Introduction to Statistics using R.* Cambridge: Cambridge University Press.

BARTHOLOMEUS, B. (1973). Voice identification by nursery school children. *Canadian Journal of Psycholinguistics, 27*(4): 464-472.

BECK, E.L. (2014). *The Role of Socio-indexical Information in Regional Accent Perception by Five to Seven Year Old Children.* Unpublished PhD dissertation: University of Michigan.

BEST, C. AND KITAMURA, C. (2012). Accent on language development : using dialects to trace how children come to recognise spoken words. In L.Burton (ed.) *Psychology*. Australia: John Wiley & Sons.

BUTLER, J., FLOCCIA, C., GOSLIN, J., & PANNETON, R. (2011). Infants' Discrimination of Familiar and Unfamiliar Accents in Speech. *Infancy*, 16(4): 392-417.

COMPTON, A. J. (1963). Effects of Filtering and Vocal Duration upon the Identification of Speakers, Aurally. *The Journal of the Acoustical Society of America* 35(11): 1748-1752.

DIANA, R. A., REDER, L. M., ARNDT, J., & PARK, H. (2006). Models of recognition: A review of arguments in favour of a dual-process account. *Psychonomic Bulletin & Review*, 13(1): 1-21.

DONALDSON, J. & SCHEFFLER, A. (1999). *The Gruffalo*. Macmillan.

DRAGER, K. (2011). Speaker age and vowel perception. *Language and Speech*, 54(1): 99-121.

DRUMMEY, A.B., & NEWCOMBE, N. (1995). Remembering versus Knowing the Past: Children′ s Explicit and Implicit Memories for Pictures. *Journal of Experimental Child Psychology, 59*(3): 549-565.

FLOCCIA, C., BUTLER, J., GIRARD, F., & GOSLIN, J. (2009). Categorization of regional and foreign accent in 5-to 7-year-old British children. *International Journal of Behavioral Development*, 33(4): 366-375.

FOULKES, P. (2010). Exploring social-indexical knowledge: A long past but a short history. *Laboratory Phonology, 1*(1): 5-39.

FOULKES, P. & BARRON, A. (2000). Telephone speaker recognition amongst members of a close social network. *International Journal of Speech Language and the Law, 7*(2): 180-198.

FOULKES, P., & DOCHERTY, G. (2006). The social life of phonetics and phonology. *Journal of Phonetics*, 34(4): 409-438.

FOULKES, P. & HAY, J. (in press, due 2015) The emergence of sociophonetic structure. To appear in B. MacWhinney & W. O'Grady (eds.) *The Handbook of Language Emergence*. Oxford: Blackwell. pp. 292-313.

HADDICAN, B., HADDICAN, B., FOULKES, P., HUGHES, V., & RICHARDS, H. (2013). Interaction of social and linguistic constraints on two vowel changes in northern England. *Language Variation and Change*, 25(3): 371-403.

HUGHES, A., TRUDGILL, P., & WATT, D. (2012). *English accents and dialects: An introduction to social and regional varieties of English in the British Isles.* 5th ed. Routledge.

LADEFOGED, P., & LADEFOGED, J. (1980). The Ability of Listeners to Identify Voices. *University of California Working Papers in Phonetics*, 49: 43-51.

LAVER, J. D. (1968). Voice quality and indexical information. *International Journal of Language & Communication Disorders*, 3(1): 43-54.

MURPHY, G.L. (2002). *The big book of concepts*. Cambridge, Mass.; London: MIT Press.

NATHAN, WELLS AND DONLAN (1998). Children's comprehension of unfamiliar regional accents: a preliminary investigation'. *Journal of Child Language,* 25:345-365.

NEWCOMBE, N. S., DRUMMEY, A. B., FOX, N. A., LIE, E., & OTTINGER-ALBERTS, W. (2000). Remembering Early Childhood How Much, How, and Why (or Why Not). *Current directions in psychological science,* 9(2): 55-58.

NOLAN, F. (1983). *The phonetic bases of speaker recognition.* Cambridge: Cambridge University Press.

NYGAARD, L. C., & PISONI, D. B. (1998). Talker-specific learning in speech perception. *Attention, Perception, & Psychophysics,* 60(3): 355-376.

ÖHMAN, L. (2013). *All Ears: Adults and Children's Earwitness Testimony*. Unpublished Doctoral Thesis. University of Gothenburg.

PAPÇUN, G., KREIMAN, J., & DAVIS, A. (1989). Long-term memory for unfamiliar voices. *The Journal of the Acoustical Society of America,* 85(2): 913-925.

PEIRCE, JW (2007). PsychoPy - Psychophysics software in Python. *Journal of Neuroscience Methods* 162(1-2): 8-13

PISONI, D. B. (1997). Some thoughts on "normalization" in speech perception. *Talker variability in speech processing*: 9-32.

PURHONEN, M., KILPELÄINEN-LEES, R., VALKONEN-KORHONEN, M., KARHU, J., & LEHTONEN, J. (2005). Four-month-old infants process own mother's voice faster than unfamiliar voices—Electrical signs of sensitization in infant brain. *Cognitive Brain Research,* 24(3): 627-633.

R CORE TEAM (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

REISBERG, D. (2010). *Cognition: exploring the science of the mind.* 4th Ed. New York; London: W.W.Norton.

SCHMALE, R., CRISTIÀ, A., SEIDL, A., & JOHNSON, E. K. (2010). Developmental Changes in Infants' Ability to Cope with Dialect Variation in Word Recognition. *Infancy,* 15(6): 650-662.

SCHWEINBERGER, S. R., HERHOLZ, A., & SOMMER, W. (1997). Recognizing Famous Voices: Influence of Stimulus Duration and Different Types of Retrieval Cues. *Journal of Speech, Language, and Hearing Research,* 40(2): 453-463.

SJÖSTRÖM, M., ERIKSSON, E. J., ZETTERHOLM, E., & SULLIVAN, K. P. (2009). A switch of dialect as disguise. *Lund Working Papers in Linguistics,* 52: 113–116.

SPENCE, M.J., P.R. ROLLINS AND S. JERGER. (2002). Children's recognition of cartoon voices. *Journal of Speech, Language and Hearing Research,* 45: 214-222.

VAN LANCKER, D., & KREIMAN, J. (1985). Unfamiliar Voice Discrimination and Familiar Voice Recognition Are Independent and Unordered Abilities. *University of California Working Papers in Phonetics*, 62: 50-60.

VAN LANCKER, D., & KREIMAN, J. (1987). Voice discrimination and recognition are separate abilities. *Neuropsychologica,* 25: 829–834.

WATT, D., TILLOTSON, J.(2001). A spectrographic analysis of vowel fronting in Bradford English. *English World-Wide,* 22(2): 269-303.

WATT, D. (2010). The Identification of the Individual Through Speech. In C. Llamas, D. Watt (eds.), *Language and Identities* (pp. 76-85). Edinburgh: Edinburgh University Press.

WELLS, J. C. (1982). *Accents of English 1: An Introduction* (Vol. 1). Cambridge: Cambridge University Press.

Ella Jeffries
Department of Language and Linguistic Science
University of York
Heslington
York
email: erwj500@york.ac.uk