“This is what a mechanic sounds like.” Children’s vocal control reveals implicit occupational stereotypes

Article (Accepted Version)

“This is what a mechanic sounds like.” Children’s vocal control reveals implicit occupational stereotypes

Word count: 2,029

“This is what a mechanic sounds like.” Children’s vocal control reveals implicit occupational stereotypes

Abstract

This study explored the use of variation in sex-related cues of the voice to investigate implicit occupational stereotyping in children. Eighty-two children between the ages of 5 and 10 took part in an imitation task where they were provided with descriptions of nine occupations (three traditionally male, three traditionally female and three gender neutral professions) and asked to give voices to them (e.g. “How would a mechanic say...?”). Overall, children adapted their voices to conform to gender-stereotyped expectations by masculinising (lowering voice pitch and resonance) and feminising (raising voice pitch and resonance) their voices for the traditionally male and female occupations, respectively. The magnitude of these shifts increased with age, particularly in boys, and was not mediated by children’s explicit stereotyping of the same occupations. We conclude by proposing a simple tool based on voice pitch for assessing levels of implicit occupational gender stereotyping in children.

Keywords: occupational gender stereotypes, implicit stereotypes, children, voice imitation

Introduction

Occupational gender stereotyping is present from early childhood (Ruble et al., 2006). Developmental research in this area typically uses explicit judgments (e.g., "Can girls be doctors?"; "Who can do this job? Only men? Only women? Both women and men?": Signorella et al., 1993), which can be subject to expectancy and social desirability biases (as they are in adults: White & White, 2006).

However, investigating spontaneous or automated manifestations of occupational stereotypes in children is more challenging. For example, the Implicit Association Task (IAT: Greenwald et al., 1998), which is commonly used with adults, requires that participants maintain two sets of representations simultaneously and have a certain level of reading proficiency (e.g., assigning words in a stereotype-compatible (i.e., boy-truck) or stereotype-incompatible (i.e., girl-truck) manner by pressing different buttons), which can be too cognitively demanding for children younger than eight years old (McKeague et al., 2015). While child-friendly variants of the IAT (e.g. target categories are represented as images, as in AIP: Banse et al., 2010), and other behavioural interference paradigms (e.g. Stroop-like tasks: Most et al., 2007) have been developed, a further limitation of these measures is that they prime children with gender information (e.g. by asking children to assign attributes to boys or girls), and responses might be influenced by the task demands.

In the present study we developed and tested a simple tool to investigate behavioural manifestations of spontaneous gender stereotyping in children which overcomes the above limitations by exploiting the volitional characteristics of the human voice.

Besides being a medium of linguistic communication, the human voice has been shown to signal various aspects of one's identity, including gender, masculinity

and femininity. Indeed, adults routinely and volitionally adjust voice fundamental frequency (F0, or pitch) and the spacing between vocal tract resonances or formants (ΔF), which are sexually dimorphic in adults (being lower in men than in women), to vary the expression of their masculinity and femininity in gendered contexts (e.g. dating scenarios and job interviews: Pisanski et al., 2016). These vocal behaviours are likely to be acquired early in childhood: from around 4 years of age, boys speak with lower formants than girls despite the absence of significant overall anatomical differences in the vocal apparatus between the two sexes (Vorperian & Kent, 2007). More recently, two imitation studies have shown that primary school children spontaneously masculinise (by lowering F0 and ΔF) and feminise (by raising F0 and ΔF) their voices when imitating the opposite gender (Cartei et al., 2014), and when giving voice to stereotypically masculine or feminine child characters of the same gender as them (Cartei et al., 2019). Given that sex-related voice cues (F0 and ΔF) can be flexibly and dynamically controlled from childhood, and that their variation can be quantified, we propose that voice imitation tasks can be applied to implicitly access and monitor changes in children's gender constructs, such as occupational stereotyping.

More specifically, in the present study we use a voice imitation paradigm to explore children's implicit beliefs about gender-typed occupations, where children are asked to give voices to male-typed, female-typed and gender-neutral occupations. We predict that children will raise their F0 and ΔF (thus feminising their voices) for the female-typed occupations, and lower their F0 and ΔF (thus masculinising their voices) for the male-typed occupations, while their F0 and ΔF will be somewhere in the middle for the gender-neutral occupations. We will also evaluate the extent to which these acoustic variations are made independently of children's explicit beliefs about

occupational stereotypes by controlling for children's explicit ratings of occupational stereotypes (given via a questionnaire asking "Who does the job of being…?" on the same occupations). Finally, we will test the extent to which F0 and ΔF manipulations can be operationalised into an "Index of Stereotypicality" which could be used to quantify implicit occupational stereotyping in children.

Methods

Participants

The sample consisted of 82 children, from Years 1-6 in the UK system, 28 5- to 6- year olds (from UK Years 1-2, 16 girls, *M* 5.9, range 5.22-7.1), 27 7- to 8- year olds (from UK Years 3-4, 12 girls, *M* 7.4, 7.4-9*),* and 27 9- to 10-year olds (from UK Years 5-6, 14 girls, *M* 9.2, 9.3-10.8*).* The children were all native British-English speakers, recruited from four primary schools in Sussex, UK. Written consent was obtained from teachers and parents prior to testing, and children also gave their verbal assent on the day of testing. The study received ethics approval from the Sciences & Technology Cross-Schools Research Ethics Committee (C-REC) at the University of Sussex (Certificate ER/VC44/13).

Procedure

Children were individually tested in a quiet room at their school. They first completed an occupation imitation task, and then an occupations questionnaire.

Imitation task. Children were first recorded as they read out loud four sentences in their own voice until they were familiar with them ("The cat is on the box", "Jane runs up a hill", "The sheep is blue", "Where were you yesterday?"). These sentences were selected from primary school phonics teaching material, as they contained the main vowels in British English, they were gender-neutral in content and

were relatively easy to read. After reading the sentences in their natural voices, the children were told to read the sentences aloud as if they were someone in a specific occupation (9 occupations in total, 3 male-typed: builder, lorry driver, mechanic, 3 female-typed: babysitter, beautician, nurse and 3 gender-neutral: doctor, student, writer (compiled from a questionnaire study on UK primary school children, Miller et al., 1999, and archival data on true gender ratios from the UK Office of National Statistics, 2019). The instructions also explained what each occupation entailed e.g. "Imagine that you are a doctor. A doctor is someone who cures people who are ill. Imagine how a doctor sounds when they talk and read the following sentences**…".**

Recordings were performed with the child sitting in front of a desk and a ZOOM H1 recorder positioned on the desk at approximately 30 cm from the child's mouth. A Marantz Sound Shield Reflection Filter was positioned on the desk around the recorder to minimise environmental noise. The sentences were presented in a child friendly font and size (Comic Sans Serif, font size 20) on a laminated A4 sheet positioned at eye level behind the recorder, so that children could read the sentences without tilting their heads. If a child had difficulty reading the sentences in the practice stage, the researcher played a pre-recorded female voice reading the sentence out loud and asked the children to repeat the sentence after listening to the recording. Sound files were recorded at 44.1 kHz, 16 bits and saved in WAV format.

Occupations Questionnaire. After the imitation task, children were asked to complete a questionnaire. The questionnaire asked children "Who does the job of being…?" and then listed each of the same nine occupations in a random order. To determine explicit sex-stereotyping, children were asked to rate the occupations on the scale from "only women", "mostly women, some men", "both men and women", "mostly men, some women", "only men" (scored 1 to 5). Each occupation had a

picture (an icon related to the occupation, but without gender-related information e.g. a spanner and a car for "Mechanic") and a non-gendered description of the occupation to improve clarity. The researcher read each question and the scale and she marked the child's choice on the answer sheet.

Sound analyses

Recordings were edited to remove noise (e.g. child's short inspiration bouts, hums, researcher's instructions) and long pauses. To extract F0 and ΔF from each recording, we used a dedicated batch-processing script in PRAAT (Boersma & Weenink, 2019). The script extracted mean fundamental frequency (hereafter F0) using the To Pitch (cc), command (pitch range parameters: 60-600Hz). Additionally, the script estimated the centre frequencies of the first four formants (F1 to F4) of each recording (formant parameters as follows: number of formants 6, max formant 7800Hz, dynamic range 30 dB, length of the analysis window 0.03 s.). The difference between any two adjacent formant frequencies, also defined as formant spacing, was then calculated ($\Delta F = Fi_{i+1} - F_i$) following Cartei and colleagues (2012). ΔF is determined by, and inversely correlated to, the length of the vocal tract of the speaker: longer vocal tracts produce narrower formant spacing. This measure is a more accurate estimate of global vocal tract adjustments than individual formant values (Pisanski et al., 2014). The computed values for F0 and individual formants were double-checked by visual inspection of the spectrogram, and analysis parameters were adjusted to correct erroneous estimates. The means and standard deviations for each acoustic parameter are reported in the Supplementary Table S1.

Statistical analyses

All analyses were conducted using the statistical program SPSS version 24. We first tested the effect of age group (5-6, 7-8, and 9-10 year-olds) and occupation

type (stereotypically male, stereotypically female or gender-neutral) on the measured acoustic variables, F0 and ΔF, by running two Linear Mixed Models separately by sex. Age group, occupation type and their interaction were entered in the models as fixed factors, while children's F0 and ΔF from the recordings of the imitated occupations were defined as the outcome variables. LMMs also included participant identity as a subject term, and occupation (within occupation type) as a nested random factor term. Children's explicit ratings on the questionnaire were added to the model as a covariate, given that, on average, children rated the stereotypically male occupations as done mostly by men, the stereotypically female as mostly done by women, and the gender-neutral occupations by both sexes (for statistical analysis and results of the questionnaire ratings see Appendix 1). All pairwise comparisons were Bonferroni corrected.

We also devised an "Index of Stereotypicality" (IoS) for each acoustic variable, and tested its use as a simple tool to quantify occupational stereotyping in children. F0 IoS was obtained by dividing F0 across the feminine occupations by F0 across the masculine occupations per participant. Similarly, ΔF IoS was obtained by dividing mean ΔF across the feminine occupations by ΔF across the masculine occupations.

One-sample t-tests were used to analyse the extent of occupational stereotyping within each age group: absence of stereotyping would be reflected by index values equal to 1 and increasing levels of stereotyping would be reflected by index values above 1. Standard estimates of effect sizes (Cohen's *d*) with values of .2, .5, and .8 represent small, medium, and large effects (Cohen, 1988). Our sample size was based on previous research (Cartei et al., 2012; Cartei et al. 2019), and a power analysis in G*Power (Faul et al., 2007) confirmed that our sample size (n = 82)

allowed us to detect an effect as small as .15 with a power of .80, which was smaller than most effects reported in this paper.

Results

Voice imitation task

Results from the LMMs for F0 and ΔF are presented in Table I.

Table I
LMMs testing the effect of experimental factors on Fundamental Frequency (F0). and Formant spacing (ΔF) of Girls (a) and Boys (b)

		a. F0				b. ΔF			
		df1	*df2*	*F*	*p*	*df1*	*df2*	*F*	*p*
Girls	Intercept	1	167.44	1253.72	<.001	1	73.99	14304.23	<.001**
N=42	Age Group	2	36.473	0.06	.936	2	36.87	19.59	<.001**
	Occupation Type	2	290.94	23.04	<.001	2	288.96	27.51	<.001**
	Occupation Rating	1	298.77	0.02	.869	1	291.45	1.98	0.16
	Age Group * Occupation Type	4	288.04	4.51	.001	4	288.05	2.08	0.082
	Intercept	1	177.16	1489.10	<.001	1	73.33	12715.94	<.001**
Boys	Age Group	2	40.09	1.05	.359	2	39.745	9.28	<.001**
N=40	Occupation Type	2	305.53	4.25	.015	2	303.40	15.52	<.001**
	Occupation Rating	1	309.97	4.34	.038	1	304.64	1.637	0.202
	Age Group * Occupation Type	4	303.92	6.15	<.001	4	302.96	1.273	0.281

* The mean difference is significant at the .05 level.

Fundamental frequency (F0). There was a significant main effect of occupation type on F0 for both boys and girls, both *ps* < *.001*. Pairwise comparisons (Table II) revealed that, across age groups, children's F0 followed the expected sex-stereotypical pattern: both sexes significantly raised their F0 for the stereotypically female occupations compared to the other two occupation types, and lowered their F0s for the stereotypically male occupations compared to the other two occupation types, although in boys the difference between F0s in the gender-neutral and stereotypically male occupations was not significant, $p > .05$ (also see Figure 1). Explicit ratings of occupations (via questionnaire) had a significant main effect on boys' F0, $p = .038$. However, a simple linear regression between boys' F0 and explicit

ratings revealed that the ratings accounted for only 3% ($F(1,352) = 11.84$, $p = .001$, $R^2 = .03$) of the variance in boys' F0.

Table II

Pairwise comparisons for F0 among occupation types across age groups: (Bonferroni corrected): stereotypically male (Mas), female (Fem) and gender-neutral (Neu)

	(I) OccType	(J) OccType	Mean Dif (I-J)	SE	df	p	95% CI Lower	95% CI Upper	Cohen's d
Girls n=42	Fem	Mas	32.80	4.87	293.78	<.001**	21.06	44.55	0.932
	Neu	Fem	-19.06	3.93	290.25	<.001**	-28.53	-9.59	0.552
	Neu	Mas	13.74	3.97	290.43	.002*	4.18	23.30	0.399
Boys n=40	Mas	Fem	-12.8	4.78	307.03	.022*	-24.39	-1.38	0.383
	Neu	Fem	-9.92	3.83	305.08	.03*	-19.15	-0.69	0.302
	Fem	Mas	2.96	3.73	305.41	.100	-6.02	11.95	0.091

The mean difference is significant at the ** = $p < .001$, * = $p < .05$ levels.

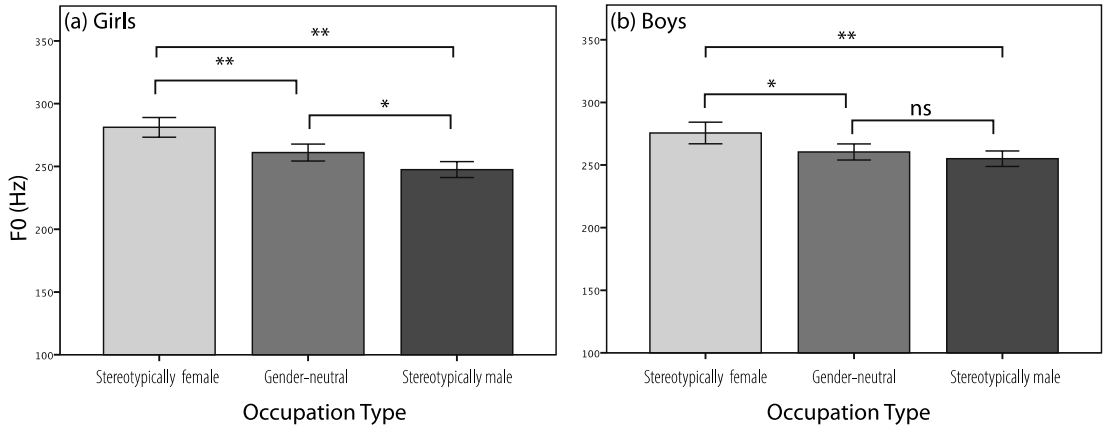

Figure 1. F0 (Hz) of girls (a) and boys (b) according to type of occupation (stereotypically female, gender-neutral, stereotypically male), ** = $p < .001$, * = $p < .05$, *ns* = not significant.

Because the main effect of occupation type on F0 was qualified by a significant interaction effect with age group in both sexes, we ran LMMs by sex and age group. Pairwise comparisons (Table III, Figure 2a) revealed that girls' F0 in the youngest age group (5- to 6-year-olds) was significantly higher in the stereotypically female occupations compared to the stereotypically male occupations. In the older age groups, girls' F0 was significantly higher in stereotypically female, and lower in stereotypically male occupations, compared to the gender-neutral occupations, reaching the highest value by 7-8 years of age. All effect sizes for the significant comparisons increased with age from medium ($d > .5$) to large effects ($ds > .8$).

In boys (Table III, Figure 2b), F0 was not significantly different between occupation types in the youngest age group (5- to 6-year-olds). Boys' F0 in the older age groups was significantly higher for stereotypically female occupations compared to the neutral or stereotypically male occupations ($p < .05$) and this difference increased with age, while no significant differences in F0 between the neutral and stereotypically male occupations were found ($p > .05$).

Table III

Results of pairwise comparisons for F0 among the three occupation types (Bonferroni corrected): stereotypically male (Mas), female (Fem) and gender-neutral (Neu)

	Age group	(I) OccType	(J) OccType	Mean Dif (I-J)	SE	df	p	95% CI		Cohen's d
								Lower	Upper	
Girls	5-6 yr olds	Fem	Mas	13.80	5.04	115.07	.021*	1.56	26.04	0.46
n=42		Fem	Neu	10.01	5.17	115.26	.17	-2.54	22.57	0.33
		Neu	Mas	3.79	5.04	115.07	1.00	-8.45	16.03	-0.13
	7-8 yr olds	Fem	Mas	40.79	7.32	67.06	<.001**	22.81	58.77	0.94
		Fem	Neu	23.86	7.41	67.11	.006*	5.67	42.04	0.55
		Neu	Mas	16.94	7.32	67.06	.07	-1.05	34.92	-0.39
	9-10 yr olds	Fem	Mas	45.27	6.01	106.97	<.001**	30.65	59.89	1.61
		Fem	Neu	23.84	6.06	107.15	<.001**	9.10	38.58	0.85
		Neu	Mas	21.43	6.10	107.21	.002*	6.59	36.27	-0.76
Boys	5-6 yr olds	Fem	Mas	5.82	5.79	83.16	.95	-8.33	19.96	0.17
n=40		Fem	Neu	-9.47	5.71	83.16	.30	-23.42	4.48	-0.28
		Neu	Mas	15.29	5.51	83.13	.02*	1.83	28.74	-0.45
	7-8 yr olds	Fem	Mas	17.91	5.31	122.53	.003**	5.03	30.80	0.61
		Fem	Neu	20.46	5.56	123.37	.001**	6.97	33.95	0.02
		Neu	Mas	-2.55	5.20	122.41	1.00	-15.18	10.09	0.09
	9-10 yr olds	Fem	Mas	35.36	6.49	100.11	<.001**	19.56	51.17	1.13
		Fem	Neu	28.92	6.40	100.03	<.001**	13.34	44.50	0.93
		Neu	Mas	6.44	6.49	100.11	.97	-9.37	22.25	-0.21

The mean difference is significant at the ** = $p < .001$, * = $p < .05$ levels.

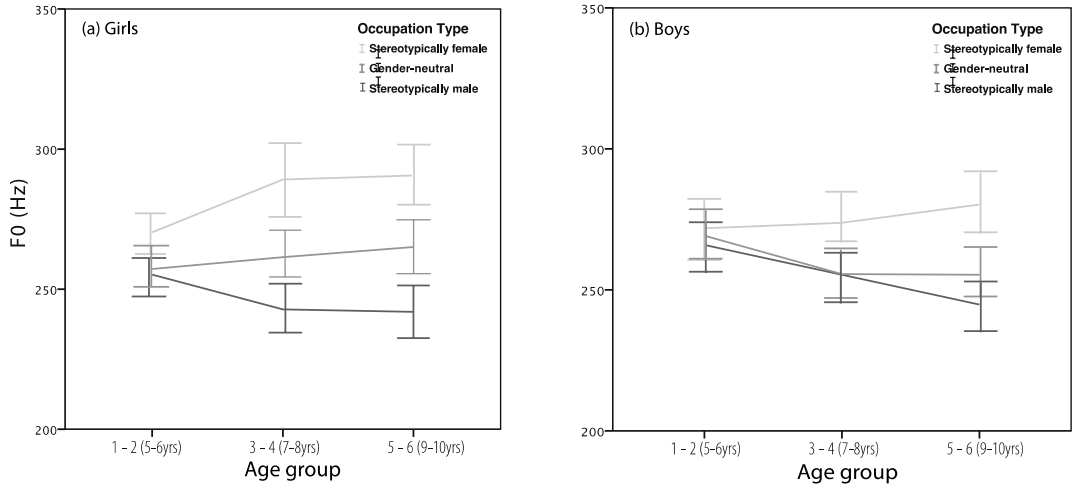

Figure 2. Significant interaction between occupation type and age group on F0 of girls (a) and boys (b). Error bars represent Cousineau (2005)'s Confidence Intervals, CI 95%.

Formant spacing (ΔF). There was a significant main effect of occupation type on ΔF for both boys and girls, *ps* < *.001*. Pairwise comparisons (Table IV) revealed that, across age groups, children's ΔF followed the expected sex-stereotypical pattern: both sexes significantly lowered their ΔF for the stereotypically male occupations compared to the gender-neutral and stereotypically female occupations, and raised their ΔF for the stereotypically female occupations compared to the other two occupation types (also see Figure 3).

Table IV

Pairwise comparisons for ΔF among occupation types across age groups: (Bonferroni corrected): stereotypically male (Mas), female (Fem) and gender-neutral (Neu)

	(I) OccType	(J) OccType	Mean Dif (I-J)	SE	df	p	95% CI		Cohen's d
							Lower	Upper	
Girls	Fem	Mas	40.76	5.51	289.84	<.001**	27.48	54.03	0.613
n=42	Neu	Fem	-18.03	4.43	288.73	<.001**	-28.70	-7.35	0.273
	Neu	Mas	22.73	4.47	288.79	<.001**	11.95	33.51	0.345
	Fem	Mas	31.33	5.68	303.81	<.001**	17.66	45.00	0.447
Boys	Neu	Fem	-13.46	4.54	303.27	.01*	-24.41	-2.51	0.193
n=40	Neu	Mas	17.86	4.43	303.37	<.001**	7.19	28.53	0.258

The mean difference is significant at the ** = $p < .001$, * = $p < .05$ levels.

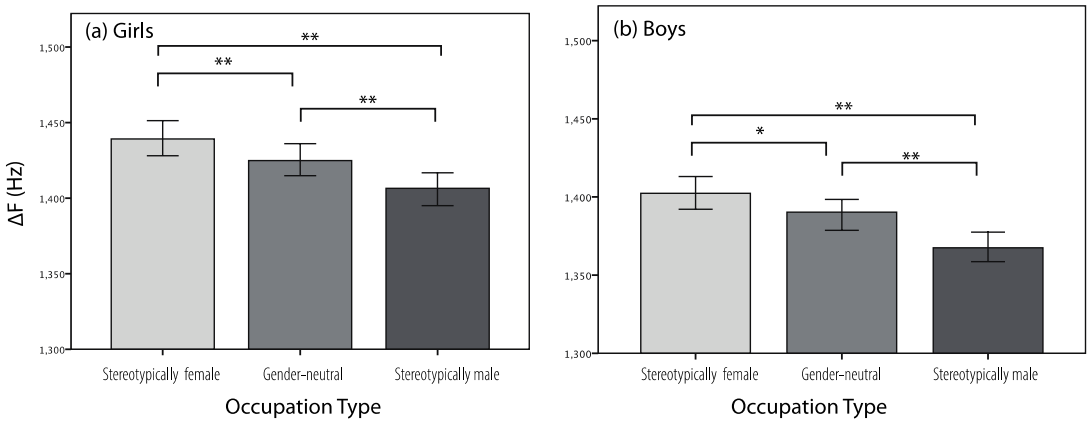

Figure 3. ΔF (Hz) of girls (a) and boys (b) according to type of occupation (stereotypically female, gender-neutral, stereotypically male), ** = $p < .001$, * = $p < .05$.

There was a significant main effect of *age group* on ΔF, *ps* < .001. In line with the age-related growth in anatomical vocal tract length, across occupation types (Table V), 5- to 6- year old children spoke with a wider ΔF compared to the two older age groups, *ps* < .001, and the ΔF of 7- to 8- year old boys was also wider than that of 9- to 10- year old boys, $p = .041$. However, ΔF of 7- to 8- year-old girls was narrower than ΔF of 9- to 10- year old girls, $p = .001$. The interaction effect between age group and occupation type was not significant in either sex, $p > .05$.

Table V
Results of pairwise comparisons for ΔF among the three age groups (Bonferroni corrected).

	Age group (I)	Age group (J)	Mean Diff (I-J)	SE	df	p	95% CI		Cohen's d
							Lower	Upper	
Girls	5-6 yr olds	7-8 yr olds	136.11	24.53	36.96	<.001**	74.61	197.62	2.13
	5-6 yr olds	9-10 yr olds	110.64	22.24	36.78	<.001**	54.85	166.42	1.82
	7-8 yr olds	9-10 yr olds	-25.48	25.18	36.90	.96	-88.63	37.68	0.40
Boys	5-6 yr olds	7-8 yr olds	90.60	25.57	39.84	.003*	26.68	154.52	1.38
	5-6 yr olds	9-10 yr olds	110.15	27.45	39.72	.001**	41.54	178.76	1.61
	7-8 yr olds	9-10 yr olds	19.55	24.95	39.67	.001**	-42.81	81.92	0.30

The mean difference is significant at the ** = $p < .001$, * = $p < .05$ levels.

Indexes of Stereotypicality (IoS). One-sample t-tests were conducted within each sex and year group to investigate whether the Indexes of Stereotypicality, F0 IoS and ΔF IoS, were significantly different from 1, thus reflecting presence of

stereotyping. We found that both girls' and boys' F0 IoS in all year groups was different from 1, *ps* < .05, with large effect sizes increasing with age (all Cohen's *d* >.5), with the exception of 5- to 6- year old boys, *p* > .05 (Table VIa). Mean ΔF IoS was also different from 1 in all year groups for both sexes, *ps* < .05 (Table VIb). To further investigate the main effect of age group on the two Indexes of Stereotypicality, we conducted two analyses of variance (ANOVAs) separately for boys and girls with age group as the independent variable on F0 IoS and ΔF IoS. Pairwise comparisons revealed that children's F0 IoS were significantly lower in the 5- to 6- year-olds than in the 9- to 10- year-olds (girls: *p* = .010, *d* = .15; boys: *p* = .017, *d* =.11). No significant differences were found for ΔF IoS between age groups, *ps* > .05.

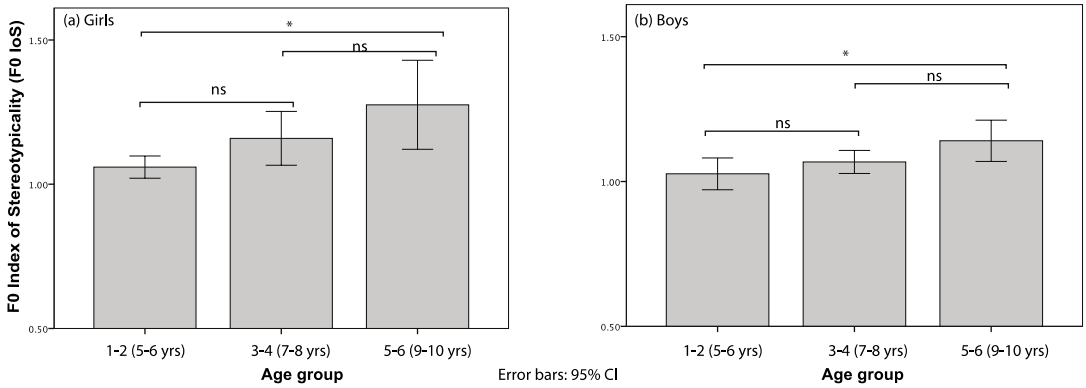

Figure 4. Fundamental Frequency Index of Stereotypicality, F0 IoS, (F0 across the feminine occupations by F0 across the masculine occupations) for girls (a) and boys (b)within each age group. Absence of stereotyping is reflected by index values equal to 1 and increasing levels of stereotyping is reflected by index values above 1.

Table VI
One-Sample Tests for Indexes of Stereotypicality (test value = 1)

		t	df	p	Mean Diff	95% CI		Cohen's d
						Lower	Upper	
a. F0 IoS								
Girls	5-6yr	2.69	15.00	0.017*	.06	.01	.11	.67
	7-8yr	3.12	9.00	0.012*	.16	.04	.27	.99
	9-10yr	4.31	13.00	0.001*	.19	.10	.29	1.15
Boys	5-6yr	0.38	11.00	0.71	.01	-.05	.07	.11
	7-8yr	2.94	16.00	0.01*	.07	.02	.12	.71
	9-10yr	3.43	12.00	0.005*	.14	.05	.23	.95
b. ΔF IoS								
Girls	5-6yr	3.95	15.00	0.001*	.02	.01	.04	.99
	7-8yr	4.12	9.00	0.003*	.02	.01	.03	1.30
	9-10yr	7.09	13.00	<.001**	.04	.02	.05	1.89
Boys	5-6yr	3.47	11.00	0.005*	.02	.01	.04	1.00
	7-8yr	5.55	16.00	<.001**	.03	.02	.04	1.35
	9-10yr	2.10	12.00	0.047*	.02	.00	.04	.58

Significance at the ** = $p < .001$, * = $p < .05$ levels.

Discussion

In line with our hypotheses, children spontaneously altered their voices according to their sex-stereotyped expectations of adult occupations: they feminised their voices (by raising F0 and ΔF), when imitating individuals in stereotypically female occupations, and masculinising them (by lowering F0 and ΔF), when imitating individuals in stereotypically male occupations. The observed voice shifts were significant even when accounting for children's explicit ratings of the same occupations from the questionnaire, thus ruling out the possibility that children were simply thinking of explicit stereotypes. The only study (White & White, 2006) conducted so far comparing implicit (IAT) and explicit stereotypes (questionnaire) in relation to occupations, although in adults, also reported that correlations between these explicit and implicit measures were modest. Our findings add to those of White and White, by showing that the imitation priming task used in this study could be used to assess relatively automatic mental associations in relation to occupational stereotyping beyond the effects reported by explicit self-report measures.

While overall fundamental frequency (F0) and formant spacing (ΔF) shifts followed the expected stereotypical patterns, we also observed sex- and age- related differences in the manipulation of these voice cues.

In relation to ΔF, we found that the youngest children (5- to 6- year-olds) displayed the highest ΔF values, and that boys' ΔF continued to linearly decrease with age, as expected because of the age-related lengthening of the vocal tract (Kent & Vorperian, 2018). Contrary to the growth in the underlying vocal tract anatomy (Fitch & Giedd, 1999), but in line with acoustic data (Cartei et al., 2019), the ΔF of girls showed a slight, but significant, increase after age 8. Given the presence of sex differences in formant values reported in the literature, and the absence of underlying sex differences in vocal tract length (Kent & Vorperian, 2018 for a review), our results support the hypothesis that sex differences in ΔF before puberty are largely due to sex-specific differences in vocal behaviour. Specifically, girls may dynamically shorten their tract by speaking with spread lips (and/or a raised larynx), which would raise the formants and widen their spacing. Interestingly, one study on facial expressions reports that girls start to smile more than boys from about age 9, and this behaviour continues throughout adulthood (Dodd et al., 1999). It is also possible that boys' age-related decline in ΔF values may be further accentuated by boys rounding their lips, as suggested in previous studies (Cartei et al., 2019; Perry et al., 2001). Robust, non-invasive, methods for lip tracking have now been developed (e.g. Optotrack motion tracking system: NDI, 2019) which would make it possible to test these hypotheses e.g. by replicating the present study while measuring the amount of lip spreading and rounding during children's imitations of occupations.

In relation to F0, the shifts between the three occupation types did not reach significance in the youngest group (5- and 6- year-olds). One possibility for this

pattern of results is that the youngest children held stereotypical beliefs regarding adult occupations to a lesser degree than the older children. This is unlikely, however, given the absence of an age group effect in the questionnaire ratings, indicating that 5- and 6- year old children exhibited the same level of occupational stereotyping as the older children, at least explicitly. Moreover, previous studies reveal that before seven years of age children apply gender stereotypes more rigidly than older children, until they begin to acquire more complex categorisation skills (Bigler & Liben 1992). A more likely explanation relates to restrictions in vocal physiology and behaviour. For example, in studies of children aged five, seven, nine and eleven years, some authors (Busby & Plant, 1995; Cappellari & Cielo, 2008) have reported that the F0 for males aged five years was significantly higher compared to all the other age groups, with no statistical difference between the ages seven, nine and eleven years. This would suggest that the youngest children had a narrower F0 range at their disposal, compared to the older children. It is also possible that the youngest children were less able to finely control F0, given that the neuromotor structures involved with phonation develop with age (Behlau et al., 2001).

We also found that while girls' F0 values for the feminine and masculine occupations reached a high and low plateau, respectively, after age 7, boys' F0 continued to increase for the feminine occupations and decrease for the masculine occupations with age. This pattern suggests that children have differential evaluations of males and females engaging in stereotypic and counter-stereotypic occupations, and that boys in particular are sensitive to gender role boundaries and norm violations, as also found in previous studies of occupational stereotyping (Blakemore, 2003; Wilbourn & Kee, 2010). To determine the origin and development of these gender differences, researchers could deploy the voice imitation paradigm presented here in

conjunction with explicit measures, such as children's beliefs about competency (e.g. whether both genders "can" perform various jobs and social roles) and gendered personality attributions (e.g. whether female- or male- dominated occupations may require stereotypically feminine or masculine traits, respectively), as well as children's own gender role flexibility.

We proposed two voiced-based Indexes of Stereotypicality, based on the ratios of F0 (F0 IoS) or ΔF (ΔF IoS) across the stereotypically female and male occupations, that can easily be used to implicitly access and monitor stereotypes - even in young children. This unique and novel contribution to the measurement of gender stereotypes not only has value in research settings but can also be implemented in applied contexts (e.g. where monitoring stereotyping and changes in stereotyping supports the reduction or elimination of biases). We found that F0 IoS was a more sensitive proxy than ΔF IoS in the case of the voice adjustments performed by the children for the different occupation types in this study. Indeed, the Index based on F0, but not ΔF, significantly increased with age, reflecting the age-related shifts in children's F0 along the expected gender-stereotypical lines. We would argue that F0 IoS also offers a more direct measure of implicit stereotyping compared to ΔF IoS, given that the latter measure can interact with age-related variation in the articulatory gestures (e.g. lip/laryngeal movements) affecting vocal tract length, as children start expressing their voice gender behaviourally (Cartei et al., 2014). Children's F0 can also be measured much more easily by researchers who have undertaken basic training in bioacoustics, and many open source audio software programs today (e.g. PRAAT: Boersma & Weenink, 2019) can robustly track children's F0 from relatively quiet audio recordings. On the other hand, children's formants are more difficult to measure automatically than adults', because of the widely spaced harmonic

components generated by children's high-pitched voices (Lieberman & Blumstein, 1988). We thus propose that the Index based on F0 (F0 IoS) could be particularly useful as a simple tool for assessing and monitoring stereotypes in children across multiple gender-typed occupations, as in the present study, as well as in relation to individual occupations (e.g. by pairing single gender-typed occupational terms with gender-neutral ones). Given the lack of observed differences in the pitch-based Index for 5- and 6- year olds, we recommend using this measure for children aged 7 onwards. We also propose that, in order to isolate effects of maturation in the vocal apparatus and motor control from actual changes in stereotypes, F0 IoS would be best used in experiments with children of similar age, or as a pre-post measure following stereotype reducing interventions (e.g. a reduction in the F0 IoS signifying greater efficacy of interventions). This type of measure would also be useful to control for extraneous influences on F0 such as those connected to specific phonological features (e.g. pre-voicing and aspiration, Cho et al. 2019).

In summary, following recent work demonstrating children's ability to modulate voice gender in imitating same age peers (Cartei et al., 2014, 2019), our study shows that children also deploy this ability to impersonate adult voices. As well as providing further evidence for the relatively automatic activation of gender associations, even when gender is not explicitly a salient feature of a task, our findings suggest that voice-based imitation tasks can be a relatively simple and direct tool to implicitly access children's occupational stereotypes. Furthermore, our imitation paradigm may not only be used as a stand-alone measure, but it may allow researchers to verify or challenge the results of studies that have used the IAT or priming studies with children, as well as with adolescents and adults.

Acknowledgements.

Funder: The Leverhulme Trust [RPG-2016-396]. Prof Reby was also supported by the University of Lyon IDEXLYON project (ANR-16-IDEX-0005).

Open Practices Statement

The dataset supporting this article has been uploaded to the Sussex Research Online (SRO) repository: http://sro.sussex.ac.uk/id/eprint/87017

References

Banse, R., Gawronski, B., Rebetez, C., Gutt, H., & Bruce Morton, J. (2010). The development of spontaneous gender stereotyping in childhood: Relations to stereotype knowledge and stereotype flexibility. *Developmental Science*, *13*(2), 298-306.

Behlau, M., Madazio, G., Feijó, D., & Pontes, P. (2001). Avaliação de voz. *Voz: o livro do especialista*, *1*, 85-245.

Bigler, R. S., & Liben, L. S. (1992). Cognitive mechanisms in children's gender stereotyping: Theoretical and educational implications of a cognitive-based intervention. *Child development*, *63*(6), 1351-1363.

Blakemore, J. E. O. (2003). Children's beliefs about violating gender norms: Boys shouldn't look like girls, and girls shouldn't act like boys. *Sex roles*, *48*(9-10), 411-419.

Boersma, Paul & Weenink, David (2019). Praat: doing phonetics by computer [Computer program]. Version 6.0.55, retrieved 13 June 2019 from http://www.praat.org/

Busby, P. A., & Plant, G. L. (1995). Formant frequency values of vowels produced by preadolescent boys and girls. *The Journal of the Acoustical Society of America*, *97*(4), 2603-2606.

Cappellari, V. M., & Cielo, C. A. (2008). Características vocais acústicas de crianças pré-escolares. *Brazilian Journal of Otorhinolaryngology*, *74*(2), 265-272.

Cartei, V., Garnham, A., Oakhill, J., Banerjee, R., Roberts, L., & Reby, D. (2019). Children can control the expression of masculinity and femininity through the voice. *Royal Society open science*, *6*(7), 190656.

Cartei, V., Cowles, W., Banerjee, R., & Reby, D. (2014). Control of voice gender in pre-pubertal children. *British Journal of Developmental Psychology*, *32*(1), 100-106.

Cartei, V., Cowles, W., & Reby, D. (2012). Spontaneous Voice Gender Imitation Abilities in Adult Speakers. *PloS one*, *7*(2), e31353.

Cho, T., Whalen, D. H., & Docherty, G. (2019). Voice onset time and beyond: Exploring laryngeal contrast in 19 languages. *Journal of phonetics*, *72*, 52-65.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2'EU.). *HilIsUale, NJ: Lawrence Eribaum Associates*.

Dodd, D. K., Russell, B. L., & Jenkins, C. (1999). Smiling in school yearbook photos: Gender differences from kindergarten to adulthood. *The Psychological Record*, *49*(4), 543-553.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175-191

Fitch, W. T., & Giedd, J. (1999). Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *The Journal of the Acoustical Society of America*, *106*(3), 1511-1522.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of personality and social psychology*, *74*(6), 1464.

Kent, R. D., & Vorperian, H. K. (2018). Static measurements of vowel formant frequencies and bandwidths: A review. *Journal of communication disorders*, *74*, 74-97.

McKeague, L., O'Driscoll, C., Hennessy, E., & Heary, C. (2015). Using implicit measures to explore children's intergroup attitudes: methodological and practical considerations for researchers. *International Journal of Social Research Methodology*, *18*(1), 1-13.

Miller, L., & Budd, J. (1999). The development of occupational sex-role stereotypes, occupational preferences and academic subject preferences in children at ages 8, 12 and 16. *Educational Psychology*, *19*(1), 17-35.

Lieberman, P., & Blumstein, S. E. (1988). Speech physiology, speech perception, and acoustic phonetics. New York: Cambridge University Press.

Most, S. B., Sorber, A. V., & Cunningham, J. G. (2007). Auditory Stroop reveals implicit gender associations in adults and children. *Journal of Experimental Social Psychology*, *43*(2), 287-294.

NDI (2019). *Optotrak Certus.* Retrieved from: https://www.ndigital.com/msci/products/optotrak-certus/

Office of National Statistics (2019). Employment by occupation by sex. Retrieved from: https://www.nomisweb.co.uk/datasets/aps168/reports/employment-by-occupation?compare=K02000001

Perry, T. L., Ohde, R. N., & Ashmead, D. H. (2001). The acoustic bases for gender identification from children's voices. *The Journal of the Acoustical Society of America*, *109*(6), 2988-2998.

Pisanski, K., Cartei, V., McGettigan, C., Raine, J., & Reby, D. (2016). Voice modulation: a window into the origins of human vocal control?. *Trends in cognitive sciences*, *20*(4), 304-318.

Pisanski, K., Fraccaro, P. J., Tigue, C. C., O'Connor, J. J., Röder, S., Andrews, P. W., ... & Feinberg, D. R. (2014). Vocal indicators of body size in men and women: a meta-analysis. *Animal Behaviour*, *95*, 89-99.

Ruble, D. N., Martin, C. L., & Berenbaum, S. (2006). Gender development. In N. Eisenberg (Ed.), *Handbook of child psychology: Vol. 3, personality and social development (6th Edn.)*. New York: Wiley.

Signorella, M. L., Bigler, R. S., & Liben, L. S. (1993). Developmental differences in children′ s gender schemata about others: A meta-analytic review. *Developmental review*, *13*(2), 147-183.

Vorperian, H. K., & Kent, R. D. (2007). Vowel acoustic space development in children: A synthesis of acoustic and anatomic data. *Journal of Speech, Language, and Hearing Research*.

White, M. J., & White, G. B. (2006). Implicit and explicit occupational gender stereotypes. *Sex roles*, *55*(3-4), 259-266.

Wilbourn, M. P., & Kee, D. W. (2010). Henry the nurse is a doctor too: Implicitly examining children's gender stereotypes for male and female occupational roles. *Sex Roles*, *62*(9-10), 670-683.

Table S1

Means and standard deviations for the acoustic parameters (fundamental frequency, F0, and formant spacing, ΔF) in children's natural voices (by sex and age group)

	Age group		
	5 to 6 yrs	7 to 8 yrs	9 to 10 yrs
Girls			
F0 (SD)	243.72 (26.8)	239.67 (37.84)	250.22 (17.56)
ΔF (SD)	1486.69(68.47)	1379.17 (68.92)	1388.33(56.60)
Boys			
F0 (SD)	247.88(29.51)	241.28(22.53)	245.94(26.05)
ΔF (SD)	1448.74(102.56)	1381.13(47.34)	1329.06(69.79)

Appendix 1

The effect of occupation type on children's explicit ratings ("Who does the job of being…?") of the nine occupations (Table S2) was investigated via a Linear Mixed Model (outcome variable: occupation rating score; fixed factors: sex, age group (5-6, 7-8, 9-10 year olds), occupation type (stereotypically male, stereotypically female and gender-neutral) and their interactions; random factor: occupation within occupation type. Occupation type on children's explicit ratings was the only significant effect, $F(2,594.963) = 269.79$, $p < .001$: the stereotypically male occupations were preferentially assigned to men, the stereotypically female occupations to women, and the gender-neutral occupations to both sexes (Figure S1).

Table S2. *List of occupations used in the study*

	Occupation Type	
Stereotypically male	*Stereotypically female*	*Gender-neutral*
Mechanic	Babysitter	Writer
Lorry driver	Beautician	Student
Builder	Nurse	Doctor

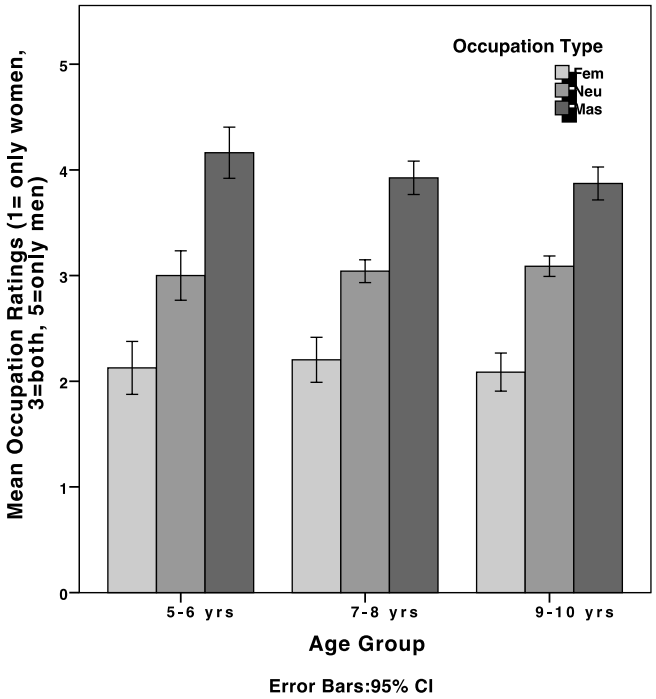

Figure S1. Children's ratings of 9 occupations (3 stereotypically male, 3 stereotypically female and 3 gender-neutral) via a questionnaire ("Who does the job of being…?").