Voice quality variation in Scottish adolescents: gender versus geography

Janet Beck & Felix Schaeffler
Clinical Audiology, Speech and Language (CASL) Research Centre, Queen Margaret University Edinburgh
jbeck@qmu.ac.uk; fschaeffler@qmu.ac.uk

ABSTRACT

Given the importance of voice quality in signalling personal identity and social group membership, effective control of voice features may become especially important during adolescence, yet this has to be achieved in the context of significant physical changes within the speech production system. Most previous research has focussed on phonation, but this study used Vocal Profile Analysis (VPA) [11] for perceptual analysis of both laryngeal and vocal tract voice settings in Scottish adolescents, in order to identify voice quality markers of gender and geographical background in this age group. VPA analysis was carried out for 76 speakers (31 male; 45 female), drawn from three geographically distinct areas of Scotland. Some of the observed variation in voice quality (especially phonatory settings) may be attributable to physical changes associated with puberty, but other setting adjustments seem more likely to be sociophonetic in origin.

Keywords: voice; Vocal Profile Analysis; adolescence; sociophonetics

1. INTRODUCTION

A person's voice quality has a pervasive effect on social interaction and integration, through the signals it sends (or fails to send) to the listener about a wide range of factors, including psychological and physical state, geographical and social background and communicative intent (e.g [6,9]). In addition to playing an important role in regulating discourse (e.g. [17,18,19]), voice quality is closely tied up with personal identity. Adolescence is a time when young people may be experimenting with and establishing their own personal identity, and exploring membership of new social groupings. Given the importance of voice quality as a social marker, it is likely that effective manipulation of voice quality may be of particular importance for adolescents seeking social inclusion, yet this is a period of rapid physical growth and change [14,20] when we know that at least some vocal parameters are subject to fluctuation and instability [2]. There has been a significant body of research focusing on the relationship between laryngeal output and physical changes during adolescence (e.g. [7]), but there has been less emphasis on other voice quality features, or on the social and psychological factors which may influence voice quality during adolescence. One reason for the paucity of research into voice quality features that result from vocal tract adjustments may be that it is very difficult to derive reliable acoustic measures of these features.

This study therefore relies on auditory perceptual judgments, using a well established tool for voice analysis, the Vocal Profile Analysis Scheme (VPAS) [11]. The VPAS was developed as a tool for research and clinical voice assessment and draws on the detailed phonetic framework for voice quality analysis developed by Laver [9, 10]. The scheme allows a systematic and comprehensive analysis of the long-term features that characterise a person's habitual voice quality. The aim is to identify the particular combination of "settings" [8] that contribute to a speaker's voice quality. A setting, in this context, can be defined as a long term muscular bias or tendency that underlies the moment-to-moment movements involved in segmental articulation. The VPAS allows reliable quantification of 20 settings describing vocal tract configuration and 6 phonatory settings, as well as overall tension settings and range of articulatory movement. These settings have predictable physiological and acoustic correlates, and are compared with a clearly defined neutral baseline.

A substantial body of research using VPA has demonstrated that long term settings of the vocal tract (i.e. lips, jaw, tongue, velopharyngeal system and pharynx) may be just as important as phonation in communicating affect [1], in marking gender and social or geographical group membership [3,13,21] and in the identification and management of clinical disorder [e.g.13,16]. Shorter term use of tongue and other settings has also been shown to be important in signalling momentary shifts in communicative intent, e.g. in mother-baby interactions [15].

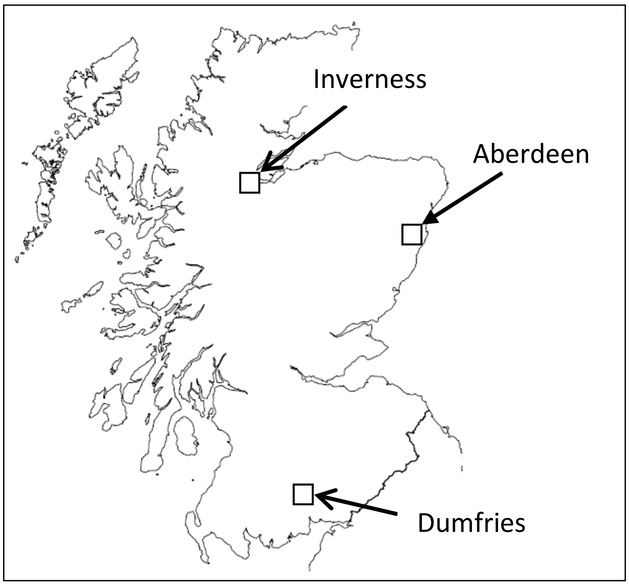
Given the potential impact of rapid changes in the size and proportions of the vocal apparatus during adolescence on the ability to manipulate vocal settings consistently, we have a particular interest in knowing which settings emerge as the strongest indicators of gender and sociolinguistics group in this age group. This will support future investigations of the relative salience of different settings for listeners.

2. METHOD

2.1 Speakers and speech material

Speech samples were selected from the Voice of Young Scots (VOYS) database [4]. This database includes a range of speech data from pupils aged 12-18 years, recorded in schools in 10 geographical locations across Scotland. For the purposes of this study three geographically distant schools were selected, from Dumfries, Aberdeen and Inverness, as shown in Figure 1.

Figure 1: Geographical distribution of speaker groups



All speakers from the selected schools were included in an initial analysis, but a small number were later rejected, where reliable perceptual analysis was impeded by poor recording quality. Table 1 shows the gender distribution and age characteristics of the speakers who were included for full analysis.

The speech data used for this study comprised the read text "The Dog and Duck" [12], providing at least 50 seconds of data, which is sufficient to allow analysis of all VPA settings.

2.2 Voice analysis

A full Vocal Profile Analysis was carried out for all speakers by an experienced VPA judge, using the VPA protocol form shown in Figure 2. Scalar degree judgments were translated into numerical values for statistical analysis. For unipolar scales, such as whisperiness, any scalar deviation from neutral was assigned a positive value. For bipolar scales (e.g. lip rounding/spreading or tongue fronting/backing), scalar deviations from neutral in one direction were arbitrarily assigned a negative value in the spreadsheet, while deviations in the other direction were assigned a positive value.

Table 1: Speaker information

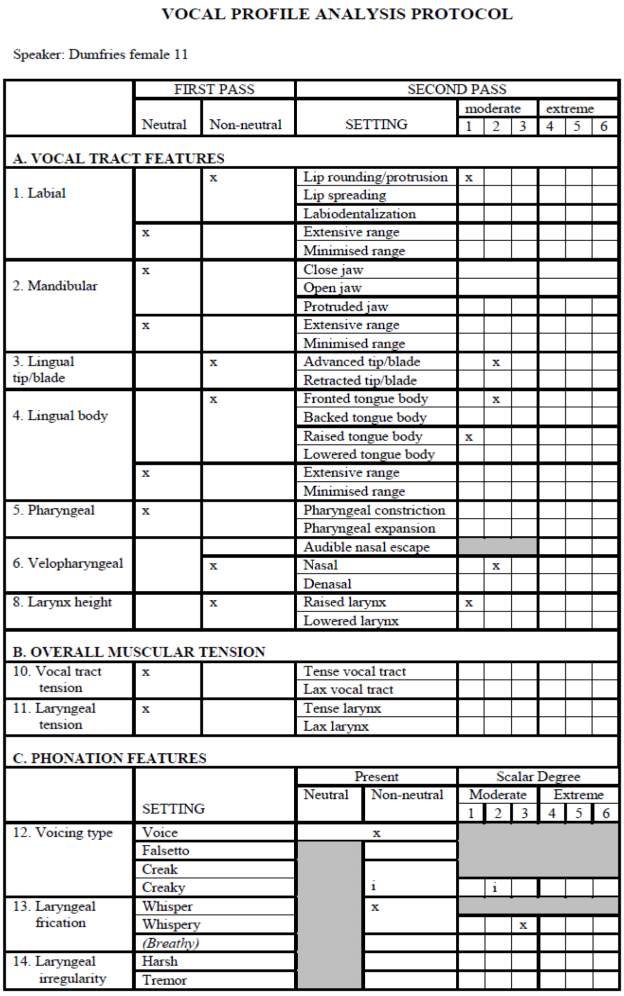
	Male		Female	
	N	Mean age (range)	N	Mean age (range)
Dumfries	11	14.4 (13-18)	12	15.0 (13-18)
Aberdeen	13	15.2 (12-18)	17	15.5 (12-18)
Inverness	7	17.6 (13-18)	16	15.2 (13-18)
All	31	15.5 (12-18)	45	15.3 (12-18)

For some speakers, settings occurred only intermittently during the speech sample, but were none-the-less judged to be characteristic of that speaker's voice. This was particularly true for harshness and creakiness, where patterns of distribution within the speech sample varied widely. The approach used here is as follows:

- Where the phonation setting was consistently present throughout the reading passage, it was relatively simple to assign a scalar degree.
- Where there were only one or two brief episodes of either harshness or creakiness within the passage, it was not judged to be characteristic of the speaker's voice and a zero value was given for the setting.
- Where harshness or creakiness occurred intermittently but frequently and regularly enough to be viewed as characteristic then (a) the scalar degree selected reflected the level of creakiness or harshness during the episodes when it was present, but (b) it was coded as "intermittent" rather than present.

Although this is a phonetically justified approach, it creates a challenge when trying to quantify the data. For example, it did not seem reasonable for the same scalar degree to be assigned to two speakers where one had consistent creakiness at scalar degree 3 and the other had intermittent creakiness at the same scalar degree. For this reason, where a setting was marked as "intermittent" the scalar degree was simply halved for statistical purposes. This is a somewhat crude approach, and in future studies it may be better to apply a more refined calculation, taking into account the proportion of speech affected by a setting.

Figure 2: An example of a completed Vocal Profile Analysis form showing the analysis for a representative female speaker from Dumfries.



VOCAL PROFILE ANALYSIS PROTOCOL

Speaker: Dumfries female 11

	FIRST PASS		SECOND PASS						
				moderate			extreme		
	Neutral	Non-neutral	SETTING	1	2	3	4	5	6
A. VOCAL TRACT FEATURES									
1. Labial		x	Lip rounding/protrusion	x					
			Lip spreading						
			Labiodentalization						
	x		Extensive range						
			Minimised range						
2. Mandibular	x		Close jaw						
			Open jaw						
			Protruded jaw						
	x		Extensive range						
			Minimised range						
3. Lingual tip/blade		x	Advanced tip/blade		x				
			Retracted tip/blade						
4. Lingual body		x	Fronted tongue body		x				
			Backed tongue body						
			Raised tongue body	x					
			Lowered tongue body						
	x		Extensive range						
			Minimised range						
5. Pharyngeal	x		Pharyngeal constriction						
			Pharyngeal expansion						
6. Velopharyngeal			Audible nasal escape						
		x	Nasal		x				
			Denasal						
8. Larynx height		x	Raised larynx	x					
			Lowered larynx						
B. OVERALL MUSCULAR TENSION									
10. Vocal tract tension	x		Tense vocal tract						
			Lax vocal tract						
11. Laryngeal tension	x		Tense larynx						
			Lax larynx						

C. PHONATION FEATURES

	SETTING	Present		Scalar Degree					
		Neutral	Non-neutral	Moderate			Extreme		
				1	2	3	4	5	6
12. Voicing type	Voice	x							
	Falsetto								
	Creak								
	Creaky		i		i				
13. Laryngeal frication	Whisper		x						
	Whispery					x			
	(Breathy)								
14. Laryngeal irregularity	Harsh								
	Tremor								

2.3 Statistical analysis

A two-way ANOVA was used to examine the effect of gender and geographical location on voice quality settings. Post-hoc analysis was carried out using Tukey HSD.

3. RESULTS

3.1 Overall findings

Table 2 shows the mean scalar degree judgments for males and females, divided according to geographical location. Standard deviations (SD) are not presented here, but no obvious between-group differences in SD were noted. It is clear from the table that for some settings (e.g. labiodentalization or range of jaw movement) there is little deviation from the neutral baseline, whereas for others (e.g. nasality) the mean scalar degree reflects the finding that these settings were present in all speakers. Other settings (e.g. larynx position) show more variability between the groups. There were multivariate significant main effects of location [F(34,104) = 1.815, p=.012] and gender [F(17,52)=5.567, p001], but there were no significant interactions.

Table 1: mean scalar degree judgments for all settings speaker groups

	Female			Male		
	Aber	Inv	Dum	Aber	Inv	Dum
lipS+/R-	0.53	0.25	0.58	-0.85	-0.57	-1
labio	0.12	0.13	0.17	0	0	0.09
lip ra	-0.12	-0.13	0	-0.15	0	-0.55
J O+/C-	-0.12	0.19	0	0.08	-0.29	0.09
jaw min	0	0.06	0	0	0	0
tipA+/R-	1.24	0.38	1.5	0.23	0.29	0.64
TBF+/B-	1.18	0.06	1.08	-1	-1	-0.45
TBR+/L-	0.82	0.19	0.92	0.54	-0.29	0.82
TB min	-0.12	0	-0.25	0	-0.29	-0.27
nasal	2	1.94	1.33	1.54	1.57	2.05
phar	0.53	0.06	0.25	0.85	0	0.09
VT ten	0.18	0.06	0.17	0.08	0	0
lar ten	0.82	0.25	0.08	0.77	0.14	0.55
larR+/L-	0.35	0	-0.17	-0.62	-0.57	-1.18
whispery	2	2.06	2.21	1.81	2.29	2.36
creaky	1	0.56	0.96	1.69	1.43	1.09
harsh	0.38	0.13	0.21	0.68	0.5	0.73

Key to abbreviations:

Aber =Aberdeen; Inv =Inverness; Dum =Dumfries

lipS+/R-	lip spreading (+)/rounding (-)
labio	labiodental
J O+/C-	open jaw (+)/close jaw (-)
jaw min	minimised jaw range
tipA+/R-	tongue tip advanced(+)/retracted(-)
TBF+/B-	tongue body fronted (+)/backed(-)
TBR+/L-	tongue body raised(+)/lowered(-)
TB min	minimised tongue body range
phar	pharyngeal constriction
VT ten	tense vocal tract
lar ten	tense larynx
larR+/L-	larynx raised(+)/lowered(-)

3.2 Gender differences in voice quality

Univariate analyses showed significant effects for:

- lip position F(1,68)=27.555, p<0.001;
- tongue tip/blade F(1,68)=12.378, p=0.001;
- tongue body front/back F(1,68)=40.877, p<0.001;
- larynx position F(1,68)=12.722, p=0.001;
- creakiness F (1,68) = 11.380, p=0.001;
- harshness F (1,68) = 8.429, p=0.005.

The setting differences are shown in Figure3.

Figure 3: Gender differences in mean scalar degrees for voice quality settings. Abbreviations are as in Table 2.

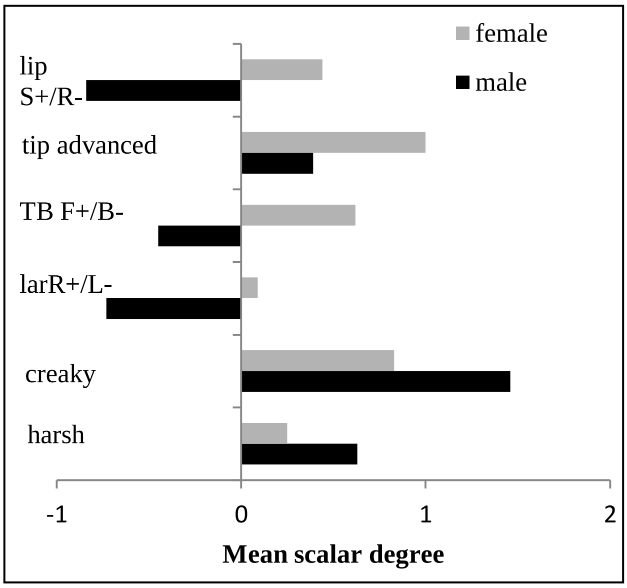
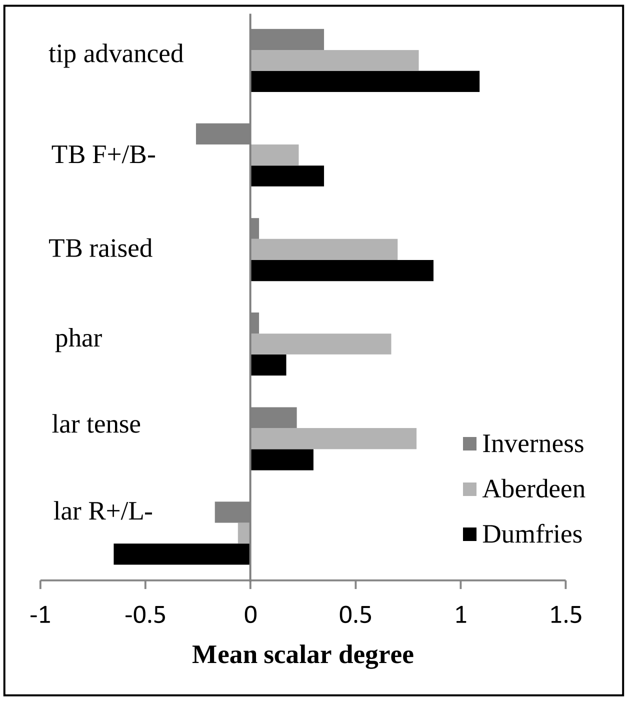


Figure 4: Geographical differences in mean scalar degrees for voice quality settings. Abbreviations are as in Table 2.



3.3 Geographical differences in voice quality

Univariate analysis showed significant effects for:

- tongue tip/blade $F(2,68)=4.452$, $p=0.015$;
- tongue body raised $F(2,68)=10.049$, $p<0.001$;
- pharynx constriction $F(2,68)=13.905$, $p<0.001$;
- larynx tension $F(2,68)=4.202$, $p=0.019$.

Differences for these settings are shown in Figure 4. Post-hoc analysis showed the following significant findings:

- tongue tip/blade: Dumfries-Inverness (p=0.008);
- tongue body raised: Aberdeen-Inverness (p=0.004) and Inverness-Dumfries, (p<0.001);
- pharyngeal constriction: Aberdeen-Inverness (p<0.001) and Aberdeen-Dumfries (p=0.001);
- larynx tension: Aberdeen-Inverness (p=0.029).

Aberdeen voices seem to be characterised by higher levels of tension and pharyngeal constriction, which may link with the non-significant trend in larynx position. Inverness voices deviate least from the neutral baseline overall, and have a less fronted tongue position.

4. DISCUSSION

Gender differences found in this study are largely consistent with previously reported VPA results for Scottish adults [13,21], and they also show a striking tendency for settings to exaggerate the consequences of sexual dimorphism. For example, the combination of lowered larynx and lip rounding/protrusion seen in males will lengthen the vocal tract, thus exaggerating the anatomical lengthening of the vocal tract that occurs in males during puberty. Similarly, tongue backing may enhance the effect of the larger male oral cavity. Even the youngest male speakers showed these male-type voice settings, so it seems that pre-pubertal children adopt vocal behaviour that mimics the voices of adult males. The greater levels of harshness and creakiness, on the other hand, may be the simple consequence of greater physical fluctuations affecting the larynx in males.

Regarding location, we found no significant geographical differences for phonation in these adolescents; only for vocal tract settings. This was somewhat surprising, as previous studies of voice quality in adults have found that phonation is an important sociophonetic marker [3,5,21]. One possible explanation could lie in the reduced phonatory control that may accompany hormonal fluctuations and rapid laryngeal growth during puberty [2,7]. Physical changes affecting the articulatory system and resonating cavities may be more gradual, so fine control of vocal tract settings during puberty could be more reliable than control of phonation. If it is the case that vocal tract adjustments are more dependable as social markers during adolescence, they could be especially salient for young people. For this reason it would be interesting to explore the attributions made by adolescent and adult listeners, as well as undertaking more extensive comparisons of sociophonetic vocal markers in adolescents and adults.

5. REFERENCES

[1] Bezooijen, R. van. 1984. *Characterization and Recognizability of Vocal Expression of Emotion.* Dordrecht: Foris Publications.

[2] Boltezar, I.H., Burger, Z.R., Zargi, M. 1997. Instability of voice in adolescence: pathologic condition or normal developmental variation? *Journal of Pediatrics*, 130(2), 185-90.

[3] Coadou Toscano, M., Audibert, N. 2009. Voice quality and the teaching of English as a foreign language: a pilot study. *Proc. 3rd Advanced Voice Function Assessment International Workshop* Madrid.

[4] Dickie, C., Schaeffler, F., Draxler, C. & Jänsch, K. 2009. Speech recordings via the internet: an overview of the VOYS project in Scotland. *Proc. INTERSPEECH 2009*, 1807-1810

[5] Esling, J.H. 1978 *Voice quality in Edinburgh: a sociolinguistic and phonetic study.* Unpublished doctoral dissertation, University of Edinburgh.

[6] Gobl, C., Ni Chasaidhe, A. 2003. The role of voice quality in expressing emotion, mood and attitude. *Speech Communication,* 40, 189-212.

[7] Hollien, H. 2012 On pubescent voice change in males. *Journal of Voice* 26(2),29-40.

[8] Honikman (1964). Articulatory settings. In: Abercrombie, D., Fry, P., MacCarthy, N., Trim, J. (eds), *In honour of Daniel Jones*. London: Longman, 73-84.

[9] Laver, J. 1980. *The Phonetic Description of Voice Quality.* Cambridge: Cambridge University Press.

[10] Laver, J. 2000 Phonetic evaluation of voice quality. In: M.J. Ball, M.J., Kent, R.D. (eds), *Voice Quality Measurement.* San Diego: Singular Publishing Group, 37-48.

[11] Laver, J., Wirz, S., Mackenzie, J., Hiller, S.M. 1991 A perceptual protocol for the analysis of vocal profiles. In: Laver, J. (ed) *The Gift of Speech.* Edinburgh: Edinburgh University Press, 265-280.

[12] Lowit-Leuschel A, Docherty, G. J. 2001. Prosodic variation across sampling tasks in normal and dysarthric speakers. *Logopedics, Phoniatrics, Vocology* 26(4), 151-164.

[13] Mackenzie Beck, J. 1994. Perceptual analysis of voice quality: the place of Vocal Profile Analysis. In: Hardcastle, W., Mackenzie Beck, J. (eds) *A Figure of Speech: a Festschrift for John Laver*. New Jersey: Lawrence Erlbaum Associates, 285-322.

[14] Mackenzie Beck, J. 2010. Organic variation of the vocal apparatus. In: Hardcastle, W.J., Laver, J., Gibbon, F (eds) *Handbook of Phonetic Sciences*. 2nd Edition. New Jersey: Wiley-Blackwells, 155-201.

[15] Marwick, H., Mackenzie, J., Laver, J., Trevarthen, C. 1984. Voice quality as an expressive system in mother-to-infant communication: a case study. *Edinburgh University Department of Linguistics Work in Progress*, 17, 85-97

[16] Mathieson, L. 2001. *Greene and Mathieson's The Voice and its Disorders*. (2nd ed). London: Whurr.

[17] Ogden. R. 2004. Non-modal voice quality and turn-taking in Finnish. In: Couper-Kuhlen, E., Ford, C. (eds) *Sound Patterns in Interaction: Cross-linguistic Studies from Conversation.* Amsterdam: John Benjamin Publishing, 29-62.

[18] Pittam, J. 1994. *Voice in Social Interaction: an Interdisciplinary Approach.* USA: Sage Publications.

[19] Reed, B.S. 2011. *Analysing Conversation: An Introduction to Prosody*. Basingstoke: Palgrave Macmillan.

[20] Stanley J., Ulijaszek, F., Johnston, M. ,Preece, A. (eds) 2000. *The Cambridge Encyclopedia of Human Growth and Development*. Cambridge: Cambridge University Press.

[21] Stuart-Smith, J. 1999 Glasgow: accent and voice quality. In: Docherty, G., Foulkes. P. (eds), *Urban Voices: Accent Study in the British Isles*. London: Arnold, 203- 222.