

# AN ACOUSTIC ANALYSIS OF VOICE QUALITY IN LONDON ENGLISH: THE EFFECT OF GENDER, ETHNICITY AND F0

Anita Szakay<sup>1</sup>, Eivind Torgersen<sup>2</sup>

<sup>1</sup>Queen Mary University of London, <sup>2</sup>Sør-Trøndelag University College  
a.szakay@qmul.ac.uk, eivind.n.torgersen@hist.no

## ABSTRACT

This study investigates voice quality in London English, using acoustic features such as harmonic structure (H1-H2) and fundamental frequency (f0). We compare the speech of 28 Inner London Hackney speakers of Anglo and non-Anglo background with the speech of 14 Outer London Anglo speakers from Havering. The results reveal that voice quality and pitch significantly differ between the two locations, where Inner London speech is lower in pitch, yet more breathy in voice quality. This suggests that the H1-H2 measure is not necessarily dependent on global f0 values.

**Keywords:** voice quality, MLE, H1-H2, ethnicity, gender, pitch

## 1. INTRODUCTION

There have been dramatic changes in the English of inner-city London over the last thirty years. The traditional working class accent known as ‘Cockney’ has been replaced by a way of speaking teenagers from the area describe as ‘slang’, but it is sometimes, rather negatively, referred to in the media as ‘Jafaikan’ because it supposedly sounds ‘black’. Researchers prefer the term ‘Multicultural London English’ (MLE) [2] because its features are fundamentally multi-ethnic in origin. MLE is typically found in areas of the city which are populated by recent immigrants, such as Hackney, and where there is a fairly equal mix of ethnicities with no dominant immigrant group. The innovative features include near-monophthongal diphthongs [5], a reduction of the indefinite article paradigm (e.g. [3]), levelling of the past tense BE paradigm (e.g. [1]), the lexical pragmatic marker ‘you get me’ [9], the quotative ‘this is + speaker’ (e.g. [2]), and syllable-timed speech rhythm ([10]).

## 2. METHODOLOGY

The *Linguistic Innovators: the English of adolescents in London* research project [4] investigated whether London was the origin of linguistic change that has been found in south-east England and sometimes beyond. The two research sites, Hackney in the traditional East End and Havering in the outskirts of London to the east, were chosen to reflect the source of these changes. The sample of young speakers from the inner city area included in this study consisted of 28 speakers, 18 male and 9 female. 9 of these speakers have an Anglo background with relatively local family roots. The remaining 19 are the children or grandchildren of immigrants, labelled non-Anglo. The population of Havering, by contrast, is predominantly Anglo. The sample for Havering for this study consists of 14 speakers, 7 male and 7 female, all of white British background reflecting the local population.

To determine whether voice quality forms part of the innovative features in inner city London, we use the H1-H2 spectral tilt measure. This measure calculates the difference between the amplitude of the first harmonic and the amplitude of the second harmonic, where larger values indicate more breathy phonation, and lower values indicate more creaky voice quality.

The speech of the 42 speakers was FAVE-aligned [6], and a total of 10,598 vowels were analysed. H1-H2 values were taken at vowel midpoints. Data points over 2.5 SD away from the mean were removed, which resulted in 10,345 remaining observations.

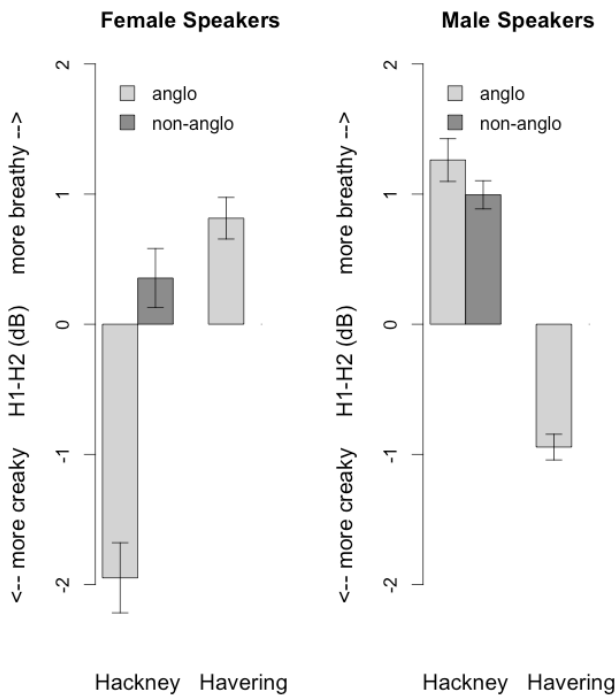
## 3. RESULTS

This section first presents descriptive statistics of the raw data on H1-H2 and f0, followed by results of a mixed effects linear regression model of H1-H2 that was fit to the Hackney data.

### 3.1. H1-H2

Significant differences were found for the H1-H2 measure in terms of speaker ethnicity, gender and location. Overall, inner-city Hackney speakers are significantly more breathy than Havering speakers in outer-London. (Wilcoxon-test,  $W = 14223530$ ,  $p < .0001$ ). However, this difference is driven by the Hackney male speakers (both Anglo and non-Anglo), who are significantly more breathy than the male speakers in Havering ( $W = 5108445$ ,  $p < .0001$ ). In fact, Hackney men are breathier than the Hackney female speakers. In particular, the female Anglo speakers in Hackney exhibit the most creaky phonation of all the groups in our data. The Havering results show a more traditional pattern, with female speech being more breathy, and male speech being more creaky. These results are shown in Fig. 1.

**Figure 1:** Mean H1-H2 and SE values across speaker groups.



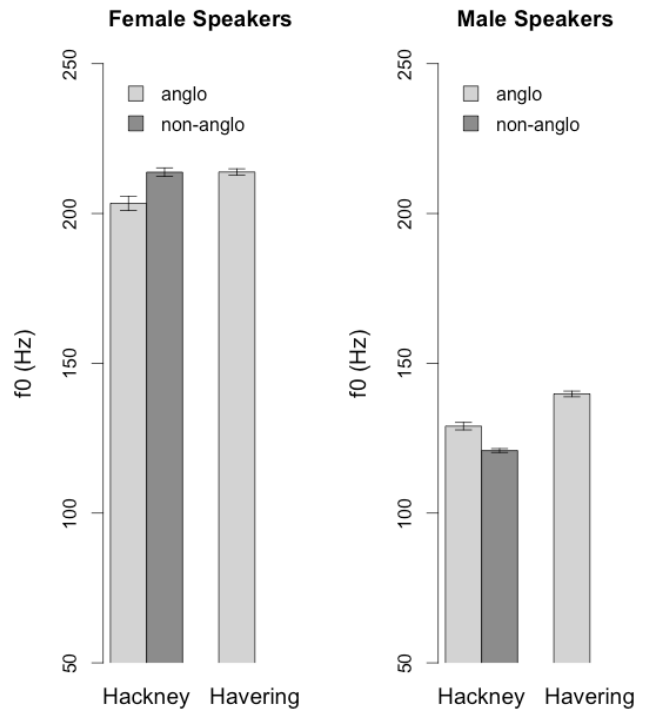
### 3.2. f0

Pitch values also show variation in terms of speaker ethnicity and location, in addition to the expected gender effect. Among the male speakers, Hackney non-Anglo men have the lowest mean pitch. It is

significantly lower than the pitch of Hackney Anglo speakers ( $W = 1006775$ ,  $p < .0001$ ), as well as that of the Havering male speakers ( $W = 3276923$ ,  $p < .0001$ ). A significant difference is also found among the female speakers in Hackney, where Anglo speakers have a lower pitch than non-Anglo speakers ( $W = 350105$ ,  $p < .0001$ ). These findings are shown in Fig. 2.

It is interesting to note that amongst the female speakers, the Hackney Anglo group has the lowest mean pitch, and also the creakiest voice quality overall. However, the men show a different pattern, where the Havering speakers show the creakiest voice quality, yet they have the highest mean pitch overall.

**Figure 2:** Mean f0 and SE values across speaker groups.



### 3.3. Mixed effects model of H1-H2

To further examine linguistic and sociolinguistic effects on voice quality, a mixed effects model was fit to the Hackney data. Havering speakers were excluded from this analysis, so the effect of ethnicity could be examined. (Recall that ethnicity is constant in the Havering data, all speakers being of Anglo background.) Speaker and word were entered as

random effects, while the fixed effects included the social variables gender and ethnicity, as well as the linguistic variables  $f_0$ , intensity, F1 and vowel duration. Continuous variables were entered into the model as log values to ensure a comparable scale amongst them. The fixed effects are shown in Table 1.

**Table 1:** Fixed effects from the linear mixed model fit to the Hackney data.

	Est	SE	df	t	p
(Intercept)	61.84	8.33	4707	7.43	< .001
$f_0$	1.05	0.73	4842	1.44	0.15
anglo	-2.32	6.01	1891	-0.39	0.69
male	-25.56	5.00	1567	-5.11	< .001
intensity	-16.27	1.85	4862	-8.79	< .001
F1	1.03	0.35	2804	2.93	< 0.01
duration	0.03	0.24	4771	0.16	0.87
$f_0$ :anglo	0.57	1.20	4787	0.48	0.63
$f_0$ :female	4.21	0.94	4834	4.49	< .001
anglo:fem	-33.86	8.40	1244	-4.03	< .001
dur:male	0.64	0.30	4829	2.14	0.032
$f_0$ :angl:fem	6.03	1.58	4812	3.81	< .001

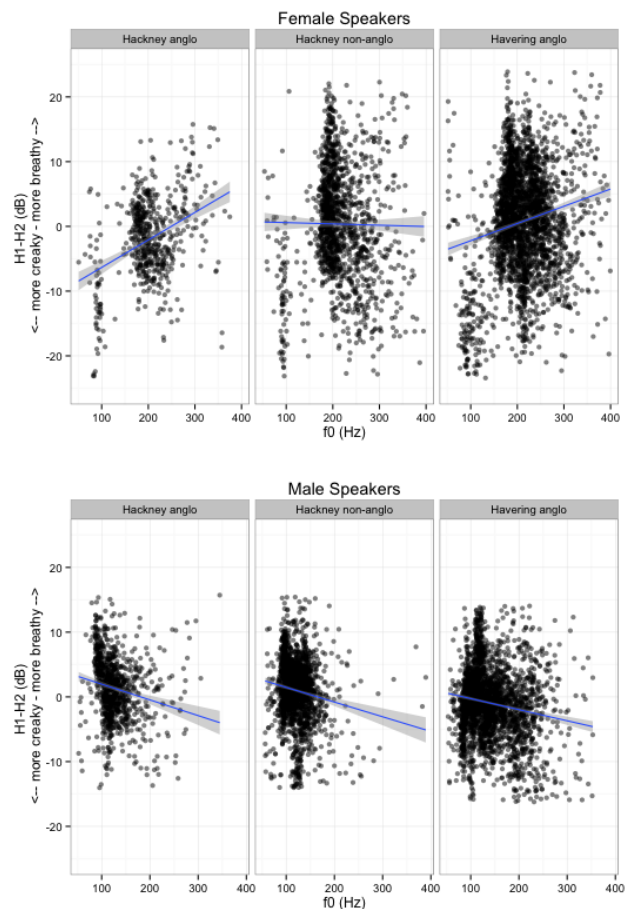
Intensity shows an effect on voice quality, where more prominent vowels are more creaky. F1 also has an effect, such that lower vowels (i.e., higher F1 values) exhibit slightly more breathy phonation. Vowel duration shows up in a significant interaction with speaker gender. Longer vowels are somewhat more breathy, but only for male speakers.

Pitch does not show up as a significant effect by itself, but it is part of a significant three-way interaction, where ethnicity, gender and  $f_0$  together have an effect on H1-H2 values. Traditionally it is assumed that higher pitch values correspond to higher – more breathy – H1-H2 values. In this data this pattern is only observed for female Anglo speakers, both in Hackney and in Havering. The non-Anglo female group in Hackney does not show any  $f_0$  effect on voice quality, while all three male groups show the opposite effect. For these male speakers, as pitch increases, voice quality become more creaky. This interaction is shown in Fig. 3.

#### 4. DISCUSSION

Our results indicate that voice quality is another innovative feature of inner-city London English speech, where non-Anglo speakers have breathier phonation than Anglo speakers in the outskirts of London. A significant interaction was found between ethnicity and gender. Hackney men have a

**Figure 3:** The effect of  $f_0$  on H1-H2 by speaker groups.



more breathy phonation than Havering men. The opposite is true for females, where Hackney women are more creaky than Havering women. In fact, they are more creaky than the Hackney men. This is interesting on two counts. First, traditionally women are believed to exhibit a more breathy phonation than men due to vocal fold physiology. Second, [7] argues that the H1-H2 measure inflates breathiness values for females. The fact that we see higher H1-H2 values for Hackney men, and lower H1-H2 values for Hackney women suggests that the inner city London speakers use voice quality for sociolinguistic purposes. In Outer London we find a more traditional pattern, with Havering females sounding more breathy, and Havering males sounding more creaky. In terms of  $f_0$ , despite being more breathy, Hackney men were found to have a significantly lower mean pitch than Havering men. [8] found similar results – but in the opposite direction – in New Zealand, where Maori male speakers have a higher

mean pitch than Anglo speakers, yet they exhibit a more creaky voice quality. These results indicate that the H1-H2 measure is not necessarily dependent on global f0 values. Taken together, the results show that voice quality and pitch do form part of the innovative features of inner city London speech.

## 5. REFERENCES

- [1] Cheshire, J., Fox, S. 2009. Was/were variation: a perspective from London. *Language Variation and Change* 21, 1–38.
  - [2] Cheshire, J., Kerswill, P., Fox, S., Torgersen, E. 2011. Contact, the feature pool and the speech community: the emergence of Multicultural London English. *Journal of Sociolinguistics* 15, 151–196.
  - [3] Gabrielatos, C., Torgersen, E., Hoffmann, S., Fox, S. 2010. A corpus-based sociolinguistic study of indefinite article forms in London English. *Journal of English Linguistics* 38, 297–334.
  - [4] Kerswill, P. 2007. Linguistic Innovators: The English of Adolescents in London: Full Research Report. ESRC End of Award Report, RES-000-23-0680.
  - [5] Kerswill, P., Torgersen, E., Fox, S. 2008. Reversing ‘drift’: innovation and diffusion in the London diphthong system. *Language Variation and Change* 20, 451–491.
  - [6] Rosenfelder, I., Fruehwald, J., Evanini, K., Yuan, J. 2011. FAVE (Forced Alignment and Vowel Extraction) Program Suite.
  - [7] Simpson, A. P. 2012. The first and second harmonics should not be used to measure breathiness in male and female voices. *Journal of Phonetics* 40, 477–490.
  - [8] Szakay, A. 2012. Voice quality as a marker of ethnicity in New Zealand: From acoustics to perception. *Journal of Sociolinguistics* 16(3), 382–397.
  - [9] Torgersen, E., Gabrielatos, C., Hoffmann, S., Fox, S. 2011. A corpus-based study of pragmatic markers in London English. *Corpus Linguistics and Linguistic Theory* 7, 93–118.
  - [10] Torgersen, E., Szakay, A. 2012. An investigation of speech rhythm in London English. *Lingua* 122, 822–840.
-