

A RE-ANALYSIS OF f_0 IN ETHNIC VARIETIES OF LONDON ENGLISH USING REAPER

Anita Szakay¹, Eivind Torgersen²

¹Macquarie University, ²NTNU-Norwegian University of Science and Technology
anita.szakay@mq.edu.au, eivind.torgersen@ntnu.no

ABSTRACT

Previous research has investigated voice quality in London English using acoustic measurements such as harmonic structure (H1-H2) and fundamental frequency (f_0) in Praat [1], and found significant interactions with regard to voice quality and pitch across two ethnic groups [11]. However, Praat's pitch tracker has been shown to be less than optimal when dealing with low fundamental frequencies.

In this study we re-analyse the speech of the same 42 participants from [11] using REAPER: Robust Epoch And Pitch Estimator [12]. The REAPER speech processing system defines the local f_0 as the inverse of the time between successive glottal closure instants, and has been demonstrated to return effective f_0 measurements at low pitch ranges [4,6]. Our results differ somewhat from previously reported values by [11] and we suggest that this is due to REAPER's more accurate f_0 measurements within the creaky voice range [6].

Keywords: creaky voice, REAPER, Multicultural London English, fundamental frequency

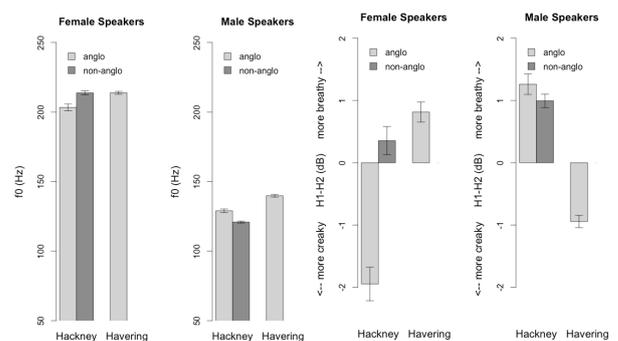
1. INTRODUCTION

This study investigates voice quality in Multicultural London English (MLE) [3] by re-analysing previously recorded speech by 42 speakers in London [11]. The data consists of a total of 46 minutes of continuous speech extracted from sociolinguistic interviews. Unlike the previous study that used Praat [1] to analyse fundamental frequency, we make use of a new tool called REAPER: Robust Epoch And Pitch Estimator [12].

MLE is spoken in multicultural inner-city areas of London that have a high immigrant population. The *Linguistic Innovators: the English of adolescents in London* research project [9] investigated the speech of adolescents of different ethnic backgrounds in Hackney, the traditional East End of inner London and Anglo speakers in Havering in the outskirts of London. It was hypothesised that any innovation first would be observed in inner London in line with [14]'s claim that London is the origin of phonological innovation in English.

MLE speakers are typically adolescents of different ethnic immigrant backgrounds referred to as non-Anglos with a short residence history in the area, but also Anglo adolescents with a long residence history in the area who are part of ethnically diverse friendship networks. The variety includes some traditional London phonological features, some phonological features that have been documented in young people's speech in south-east England in general, and some phonological features that are found in varieties of English outside of the United Kingdom. In addition, phonological innovation has been documented: there is complete H-reinstatement (/h/ in word-initial position in lexical words is pronounced), backing of /k/ in front of non-high vowels and extreme GOOSE fronting [2,7]. The use of near-monophthongal diphthongs in FACE and GOAT [10] is part of the impression of a syllable-timed speech rhythm [13] where the stressed vowels become shorter and schwa becomes longer, and in turn successive vowels have a more equal duration.

Figure 1: Mean f_0 and H1-H2 values by gender and ethnicity as reported in [11].



Other suprasegmental features, such as mean fundamental frequency and voice quality have also been shown to differ between MLE and other London varieties. [11] investigated voice quality in London speech, using 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 in the speech signal, where large positive values indicate more breathy phonation, and lower negative values indicate a more creaky voice quality. [11] showed that inner-city Hackney male

speakers (both Anglo and non-Anglo), were significantly more breathy / less creaky than the male Anglo speakers in Havering. Havering speakers showed the creakiest voice quality, yet the highest mean pitch overall. Amongst the female speakers, the Hackney Anglo group had the lowest mean pitch, and also the creakiest voice quality overall. Havering females showed the highest mean pitch, as well as the most breathy voice quality. These results are reproduced as Figure 1 in this paper.

However, accurate H1-H2 measures rely on an accurate pitch tracker. Praat's pitch tracker has been shown to struggle with very low f_0 values, and oftentimes it is unable to track the fundamental frequency [6].

2. METHODOLOGY

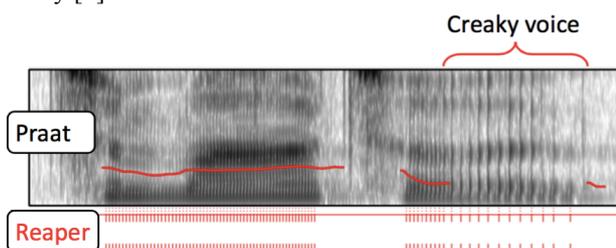
2.1. Participants

For our investigation, the sample of adolescent speakers from Hackney includes 28 speakers, of which 9 have an 'Anglo' background with relatively local family roots, and 19 'non-Anglo' speakers who are the children or grandchildren of immigrants. Havering is predominantly Anglo, with a low immigrant population. We include 14 speakers from Havering, 7 males and 7 females, who all have a white British background.

2.2. Analysis

For our analysis, instead of Praat we rely on REAPER: Robust Epoch And Pitch Estimator [12] using MacReaper [5]. The REAPER speech processing system defines the local f_0 as the inverse of the time between successive glottal closure instants, and has been demonstrated to return effective f_0 measurements at pitch ranges as low as 20 Hz [4,6]. REAPER is able to successfully track pitch even when Praat cannot find the fundamental frequency in the speech signal [6].

Figure 2: Pitch tracking by REAPER vs Praat. Praat's pitch tracker misses most of the creaky vowel but REAPER is able to detect the f_0 . Image by [4].



After running all 42 speakers' speech samples through REAPER, we obtained a total of 393,005

observations of f_0 . Rather than calculating H1-H2 values in this paper we take low f_0 values as indicators of creaky voice [8].

Consequently, for each of our speaker groups we analyse the distribution of all f_0 values obtained by REAPER, and plot it on a density plot. Most of the time this results in a bimodal distribution, such that the peak with the higher fundamental frequency corresponds to the mode of the *modal voice* phonation, while the other peak at lower fundamental frequency values corresponds to the mode of the *creaky voice* phonation. In between these two peaks we find the antinode, which is the f_0 value that represents the division between the creaky voice and modal voice distributions.

We also calculate the percent of creak as the time spent creaking over the total time spent phonating. In addition, for each speaker group we report the mean f_0 for the modal voice phonation alone, as well as the overall mean f_0 , which we compare with previously published results obtained by Praat.

3. RESULTS

3.1 Female Speakers

We first present results for our female speakers. In the inner London Hackney area the multicultural non-Anglo group has an overall mean f_0 of 175 Hz. The antinode between the creaky and modal phonation distributions is at 141 Hz. The mean f_0 for the modal voice phonation is 220 Hz. These speakers on average spent 10.9% of the time producing creaky phonation. These results are presented in Figure 3 (top panel).

The Anglo females living in Hackney only produced 5.32% creak on average. Their overall mean pitch was 179 Hz, comparable to the non-Anglo counterparts, while the modal voice mean 201 Hz. The antinode between the two peaks was at 119 Hz, which is lower than the non-Anglo group's antinode. These results are plotted in Figure 3 (middle panel).

As seen, the two Hackney female groups do not exhibit vastly different patterns. However, the Havering Anglo group shows a strikingly different pattern which can be easily seen by visually investigating the f_0 distribution in Figure 3 (bottom panel). These female Anglo speakers living in the outskirts of London produce creaky phonation 41.9% of the time. Their average antinode is at 139 Hz. The overall mean fundamental frequency is 179 Hz, while the mean f_0 for the modal voice phonation is 221 Hz.

In a previous study [11] found that the Hackney Anglo group had the lowest mean pitch, and also the creakiest voice quality overall. Havering females showed the highest mean pitch, as well as the least creaky voice quality. Our results using REAPER

rather than Praat reveal a different pattern. We find that the Havering Anglo females creak over 40% of the time, and most of the time at very low fundamental frequency values. Praat is known to have difficulties tracking these low f_0 values, and we argue that REAPER presents a more truthful picture of the story due to the fact that the low f_0 values were not picked up by Praat in the previous study. This skewed the results and suggested that Havering females had the highest mean f_0 , as well as the least creaky phonation. However, as REAPER indicates Havering Anglo speakers are in fact creakier overall than both Hackney female groups, and in fact creakier than the three male groups.

3.2 Male Speakers

The inner London multicultural non-Anglo male group from Hackney exhibited creaky phonation 14.9% of the time. Their average f_0 antinode is at 79 Hz. The overall mean f_0 is 110 Hz, while looking at the modal voice phonation alone the mean f_0 is 128 Hz (see Figure 4 top panel).

The Anglo males living in Hackney show a lower overall mean pitch at 106 Hz, while the modal voice mean is at 124 Hz. Their average f_0 antinode is at a low 63 Hz. However, these Hackney Anglo males spend less than half the time creaking compared to their non-Anglo counterparts, with 6.75% creak (Figure 4 middle panel).

The Havering male speakers living in the outskirts of London show the most creak among the three male groups, at 18.6% creak. Their mean f_0 for the modal voice phonation alone is a high 136 Hz, which is pulled down by the relatively high percentage of creaky phonation to a mean overall f_0 of 112 Hz. This group's mean f_0 antinode is at 69.2 Hz (Figure 4 bottom panel).

We see that for both female and male speakers the Anglo groups in the outskirts of London show the highest percentage of creak (females = 41.9%, males = 18.6%). This seems to indicate that creaky voice is not – after all, and contra to [11] – an innovative feature of inner city Multicultural London English, rather a characteristic of Anglo speech in the outskirts of London.

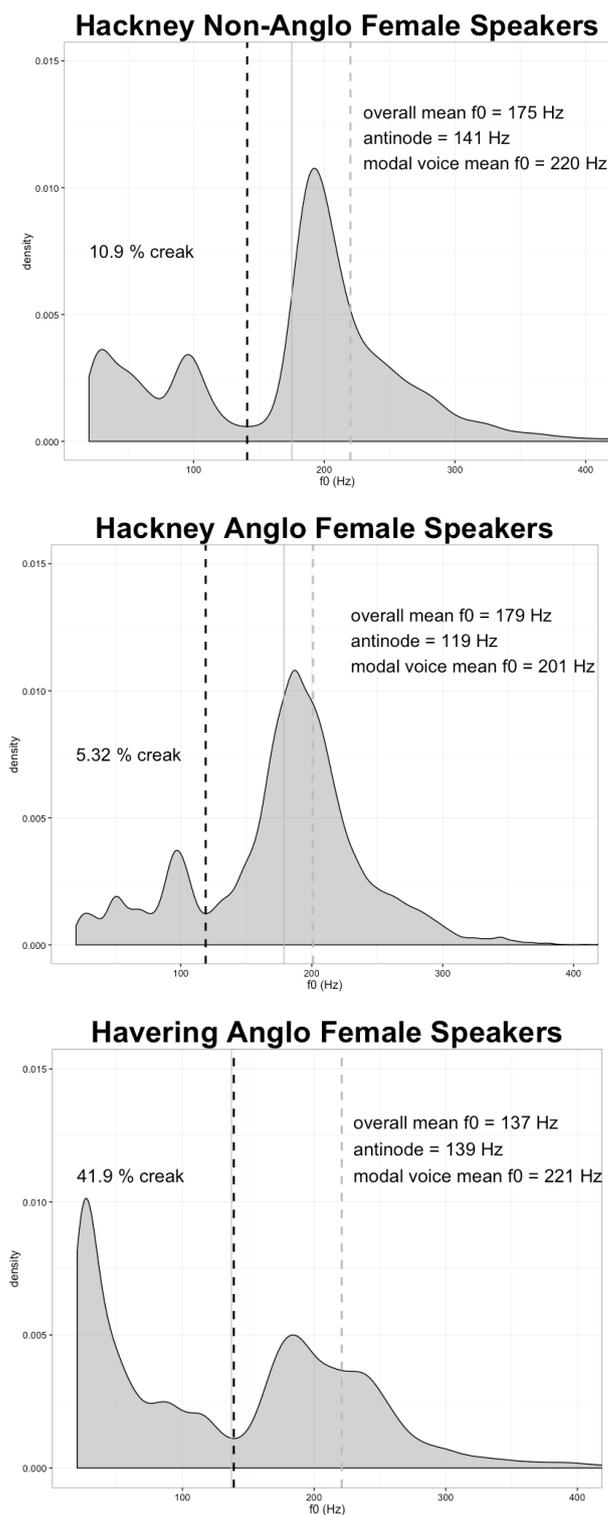
4. SUMMARY

In this paper we demonstrated that by opting to use new innovative tools such as REAPER: Robust Epoch And Pitch Estimator [12] to track fundamental frequency we are able to obtain a more accurate sociophonetic picture of the actual usage of creaky voice in London English.

We advocate the use of REAPER instead of Praat for research studies whose main research question

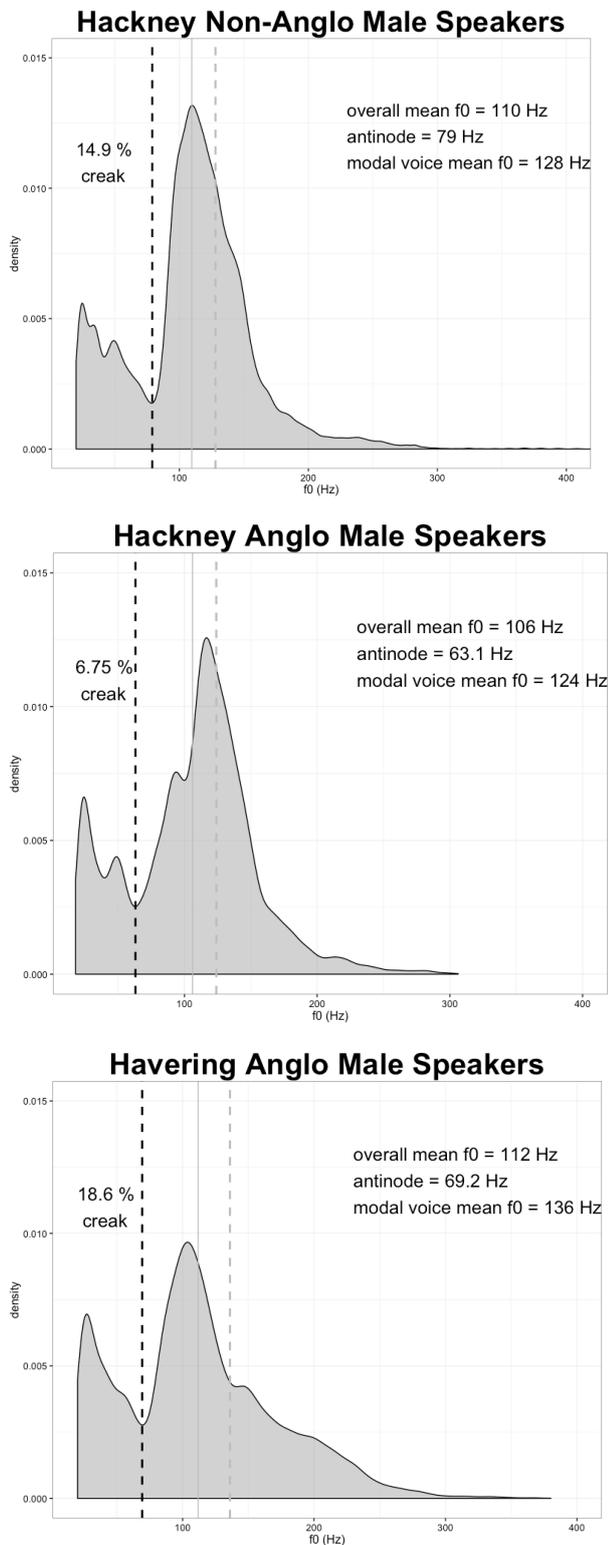
involves the measurement and analysis of low fundamental frequency values, and particularly for studies investigating creaky voice quality.

Figure 3: Density plots showing the distribution of f_0 values for the female groups. The plots indicate the antinode (dashed black line; the division between the creaky and modal voice distributions), the overall mean f_0 (solid grey line), and the mean value for the modal voice distribution (dashed grey line).



7. REFERENCES

Figure 4: Density plots showing the distribution of f0 values for the male groups. The plots indicate the antinode (dashed black line; the division between the creaky and modal voice distributions), the overall mean f0 (solid grey line), and the mean value for the modal voice distribution (dashed grey line).



- [1] Boersma, P., Weenink, D. 2018. Praat: doing phonetics by computer [Computer program]. Version 6.0.43, retrieved 8 September 2018 from <http://www.praat.org/>
- [2] Cheshire, J., Fox, S., Kerswill, P., Torgersen, E. 2008. Ethnicity, friendship network and social practices as the motor of dialect change: Linguistic innovation in London. *Sociolinguistica* 22, 1–23.
- [3] 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(2), 151–196.
- [4] Dallaston, K., Docherty, G. 2017. Sociophonetic variation in the prevalence of creaky phonation in the speech of young adults from two capital cities in Australia. Paper presented at the Language Variation and Change (LVC-A3) Workshop. Sydney, Australia.
- [5] Dallaston, K., West, J. 2018. MacReaper v2 downloaded from <https://kjdallaston.com/projects/>
- [6] Dorreen, K. 2016. Fundamental frequency distributions of bilingual speakers in forensic speaker comparison. Master's thesis, University of Canterbury.
- [7] Fox, S., Torgersen, E. 2018. Language change and innovation in London: Multicultural London English. In N. Braber and S. Jansen (eds.) *Sociolinguistics in England*. Basingstoke: Palgrave Macmillan. 189–213.
- [8] Keating, P., Garellek, M., Kreiman, J. 2015. Acoustic properties of different kinds of creaky voice. In The Scottish Consortium for ICPHS 2015 (ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, UK: The University of Glasgow
- [9] Kerswill, P., Cheshire, J., Fox, S., Torgersen, E. 2007. Linguistic Innovators: The English of Adolescents in London. ESRC Research Project, RES-000-23-0680.
- [10] Kerswill, P., Torgersen, E., Fox, S. 2008. Reversing 'drift': Innovation and diffusion in the London diphthong system. *Language Variation and Change* 20(3), 451–491.
- [11] Szakay, A., Torgersen, E. 2015. An acoustic analysis of voice quality in London English: The effect of gender, ethnicity and f0. In The Scottish Consortium for ICPHS 2015 (ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, UK: The University of Glasgow. ISBN 978-0-85261-941-4. Paper number 0996.
- [12] Talkin, D. 2015. <https://github.com/google/REAPER>
- [13] Torgersen, E., Szakay, A. 2012. An investigation of speech rhythm in London English. *Lingua* 122(7), 822–840.
- [14] Wells, J. 1982. *Accents of English*. Cambridge: Cambridge University Press.