An acoustic analysis of 'happy-tensing' in the Queen's Christmas broadcasts

Jonathan Harrington*

Institute for Phonetics and Digital Speech Processing (IPDS), University of Kiel, Olshausenstrasse 10, D-24908 Kiel, Germany

Abstract

This paper presents a longitudinal analysis of some vowels from the annual Christmas broadcasts produced by Queen Elizabeth II over a 50-year period in order to investigate whether adults adapt to sound changes taking place in the community. The sound change that was analyzed in this paper, which is sometimes known as happY-tensing, concerns the tensing of the final vowel in words like 'happy' in British English Received Pronunciation over the course of the last 50 years. In the first part of the study, schwa vowels in Christmas broadcasts separated by 40–50 years were analyzed in order to exclude as far as possible any long-term acoustic effects due to vocal tract maturation. The results of this analysis show a large decrease in both the fundamental and F1, F2, and F4 from earlier to later broadcasts. It is then shown that the Queen's 1950s happY vowel is less tense than in a 1980s corpus of four female speakers of Standard Southern British. A subsequent comparison between the 1950s and 1990s Christmas broadcast happY vowels shows a small change towards the tenser position. It is argued that the vowels exemplified by KIT and happY have undergone phonetic raising in RP, with the latter also having fronted. The Queen has participated in the first of these changes and marginally in the second.
© 2005 Elsevier Ltd. All rights reserved.

1. Introduction

The focus of this paper is the change in phonetic quality of the phonologically lax, word-final vowel in words like 'city', 'duty', and 'happy' in the standard accent of England, Received Pronunciation (RP). The paper is also about whether there is evidence for this type of phonetic change in recordings taken from Queen Elizabeth II spanning a 50-year period and how such changes can be separated from any longer-term changes that might occur due to vocal tract maturation.

There is now extensive evidence from studies in sociolinguistics in the last 30–40 years that sound change and its relationship to synchronic variability can be observed in the same community of speakers (Labov, 1994, 2001; Trudgill, 1974). Typically, these analyses are carried out in so-called apparent-time studies in which the pronunciation of younger and older speakers of the same community are compared with each other (Bailey, Wikle, Tillery, & Sand, 1991; Weinreich, Labov, & Herzog, 1968). The assumption that underlies such analyses is that there is no, or scarcely any, shift of a person's accent beyond the critical age for language

*Tel.: +89 431 880 3319; fax. +49 431 880 1578.

E-mail address: jmh@ipds.uni-kiel.de.

acquisition (Lenneberg, 1967); accordingly, any changes observed in apparent time studies should indicate a diachronic change in the accent of the community over a period that is roughly equal to the average age difference between the younger and older speakers being investigated, and there is a good deal of evidence to support this position (Labov, 1994; Sankoff, 2004).

The question of the extent to which an adult's accent shifts in the direction of community changes has been less extensively studied (but see, e.g., Brink & Lund, 1975; Sankoff, in press; Yaeger-Dror, 1994), partly because of the difficulty in obtaining speech data from the same individual over a number of years. According to recent exemplar theories and usage-based models of phonology (Bybee, 2001; Pierrehumbert, 2002, 2003), an adult's accent is expected to show some adaptation in the direction of community changes because phonological categories emerge from remembered word tokens: therefore, if there is a shift in the pronunciation of word tokens in the community that are perceived and remembered by an adult, the relationship between an adult's phonological categories and its articulatory exponents should show a corresponding shift as well. But given the paucity of longitudinal studies, we have a very incomplete model of the relative stability of pronunciation characteristics beyond the critical age.

The present paper is concerned with a longitudinal study that is, as in previous analyses (Harrington, Palethorpe, & Watson, 2000a, b, 2005), based on the annual Christmas broadcasts of Queen Elizabeth II. There are a number of reasons why these speech data are suitable for longitudinal investigations. The most important is that they are probably the only acoustically high-quality recordings from the same person producing speech with more or less the same communicative intent (an annual message to Britain and the Commonwealth) in a very similar reading style every year over a 50-year period. Another is that if we are to observe longitudinal changes, then the accent itself has to be in a state of change: if there is no, or minimal, change to an accent, then the conditions for longitudinal change would obviously not be met. As we know from socio-linguistic studies, social change is often a catalyst for sound change (e.g., Milroy, 1994) and this seems to be the case as far as the standard accent of England, RP is concerned: since the 1950s, there have been dramatic social changes involving a collapsing class structure, especially between the early 1960s and 1980s (Cannadine, 1998), which are likely to be related to a number of changes to RP from below. Thus, Wells (1997) documents various ways in which the lower class London Cockney English has influenced RP in recent times and since the 1980s various researchers have commented on the rise of Estuary English (Coggle, 1993; Crystal, 1995; Rosewarne, 1994), an accent spoken in the South of England which shares characteristics of both RP and Cockney (Wells, 1997).

Yet another reason for studying the Christmas broadcasts longitudinally is that there are many existing impressionistic (e.g., Gimson & Cruttenden, 1994; Laver, 1994; Wells, 1982); and some acoustic (Bauer, 1985; Deterding, 1997; Henton, 1983) studies of RP that can provide a background against which to interpret any changes observed in the Christmas broadcasts as being in the same direction as those that have been observed in the community or not.

The present study extends our earlier research by analyzing changes that might be the result of the physiological effect of vocal tract maturation. In Harrington et al. (2000a, b, 2005), we favored a phonetic interpretation of the observed sound changes, both because the changes to the Christmas broadcast vowels between the 1950s and 1980s were in the direction of those of 1980s speakers of a more widely spoken, and less aristocratic form of RP, but also because the changes did not seem to be compatible with the evidence from a small number of other studies that had investigated changes to formants with increasing age (e.g., Rastatter, McGuire, Kalinowski, & Stuart, 1997). One of the aims of the present study is to investigate age-related formant changes more closely in the Christmas broadcast schwa vowels. Schwas were chosen for this purpose because we can get a better understanding of physiological changes by neutralizing phonetic vowel quality as far as possible. Since schwas communicate a phonetic identity only weakly, perhaps because they may be virtually targetless (van Bergem, 1994), and since there is no evidence for any diachronic change to schwa in RP, schwa is appropriate for assessing the long-term changes due to age.

As far as the effects of aging on the speech signal are concerned, there is extensive evidence for a relationship between increasing age and decreasing fundamental frequency (Helfrich, 1979; Decoster & Debruyne, 2000; Linville, 2001; Linville & Fisher, 1985) and there are several studies of the perceptual correlates of aging (e.g., Jacques & Rastatter, 1990; Huntley, Hollien, & Shipp, 1987; Neiman & Applegate, 1990; Ptacek & Sander, 1966), but there are far fewer analyses of the effect of aging on formant frequencies. Laver and Trudgill (1979)

make the suggestion, based on remarks by Luchsinger and Arnold (1965), that the entire respiratory system and digestive tract are in a lower position with increasing age. Such a lowering would cause the vocal tract to lengthen, resulting in the well-known downwards shift in the formant frequency ranges (Lindblom & Sundberg, 1971). There are a small number of studies that are consistent with this view. These include Xue and Hao (2003) who recently showed a lowering of formant frequencies, particularly of F1, in elderly speakers of both genders, and Whiteside's (2001) reanalysis of the data in Lee, Potamianos, and Narayanan (1999) of 436 children and 56 adults who also showed formant lowering with increasing age. Compatibly, a study by Linville and Rens (2001) of long-term spectra from 80 speakers divided by age and gender showed a lowering of formants, especially of F1, from young adulthood to old age in both men and women. On the other hand, Meurer, Wender, Corleta, and Capp (2004) found less variation, but no lowering, of formants in older women while Rastatter et al. (1997) found consistent lowering only in the first formant frequency: as in other studies (e.g., Whiteside, 2001), they also found that the formant changes between the older and younger speakers were vowel-dependent and that the age-related changes were different for male and female speakers. Finally, in Labov's (1994) analysis of the same person aged 68 and then 85, there was no evidence of any significant long-term change to F1 or F2 over these years. The relationship between aging and formant changes remains, therefore, very unclear, partly because of the difficulty of obtaining controlled data across the different generations. Moreover, what is lacking in most of the very few studies that are available is an analysis over several decades *from the same person*, which would eliminate the confounding influences due to the substantial differences between different speakers' anatomical vocal tracts.

The phonetic changes in this paper are concerned with the apparent tensing in RP in recent years of the final vowel in words like 'happy' (which will be transcribed henceforth with [ɪː]). In the 1950s, [ɪː] was more open, possibly shorter, and in general phonetically closer to [ɪ] in KIT than to [iː] in FLEECE. (As in Wells, 1982, KIT and FLEECE are used to exemplify lexical sets in /ɪ/ (e.g., 'kit', 'hid', 'rid') and in /iː/ (e.g., 'fleece', 'heed', 'reed'), respectively; and happY is used to denote the final vowel in the lexical set of words like 'city', 'duty', 'happy'.) The shift of this lax vowel has taken place in RP in the latter part of the 20th Century, and may, according to Wells (1997), be associated with the rise of Estuary English (see also Trudgill, 2001). As far as the phonology is concerned, /iː/ and /ɪ/ might well show phonetic overlap in this word-final context, because this is an environment in which they are neutralized, i.e., in which the phonological contrast is suspended. Moreover, the final vowel in happY is anomalous because it is the only full (non-schwa) lax vowel that can occur in final open syllables. Perhaps its lengthening has come about by analogy to the generalization that lax vowels (other than schwa in e.g., 'sofa') are phonologically excluded from word-final position in English. Yet another reason for happY to lengthen diachronically is that, being word-final, it often occurs as the last syllable in a prosodic phrase, which is of course a primary context for synchronic lengthening in many languages (e.g., Fletcher (1991) for French; Jun (1995) for Korean; Ueyama (1999) and Umeda (1975) for American English).

In the present paper, some acoustic evidence is sought for a shift of happY towards FLEECE in the last fifty years. Based on what is known about the acoustic differences between tense and lax vowels in English (see e.g., Lehiste & Peterson, 1961; Rakerd & Verbrugge, 1985; Strange, 1989, for American English), a shift towards [iː] is likely to be associated with lengthening and a more peripheral vowel target. Tensing of [ɪː] might also be associated with an increase in F2 curvature towards a maximum F2 at the vowel target, that is, with an increased deviation of F2 from a straight line trajectory between formant onset and offset. Some evidence for this derives from van Bergem (1994) who has shown that less peripheral, short vowels in Dutch tend to have less curvature in their formants than longer, more peripheral ones. Although there seems to be no comparable data for English, it is reasonable to speculate that F2-curvature might also distinguish RP [iː] and [ɪ] in a similar way and that the curvature of [ɪː] might increase, if it has shifted towards [iː].

The results of this paper are presented in three sections. First, a comparison is made between the Christmas broadcast schwas in the 1950s and those from roughly 40 years later: this is done in order to investigate any effect on formant frequencies of maturational changes to the vocal tract that should be factored out in considering whether there has been a diachronic change in phonetic quality. The next two sections are based on the idea, discussed more extensively in Harrington et al. (2000a, b), that there has been a shift between the 1950s and more recent times from the more aristocratic form of RP, sometimes known as U- or upper-crust RP (Wells, 1982), towards mainstream RP or Standard Southern British (SSB) that is characteristic of the professional middle classes. If this is so, then we should expect to observe phonetic differences between the

Queen's 1950s vowels and those of the 1980s SSB speakers—these differences are the subject of the second section. Finally, the Queen's vowels from later broadcasts (1995–2002) should have moved in the direction of the 1980s SSB positions, if an adult's accent is influenced by changes taking place in the community. This prediction is tested in the third section.

2. Method

2.1. Segmentation, labeling, and formant calculation

Segmentation and labeling of [ə], [ɪ] (KIT), [ɪː] (happY), and [iː] (FLEECE) in three corpora were carried out in the IPDS Kiel by two trained transcribers using the same procedures as in Harrington et al. (2000a, b). Two of the corpora were from the annual Christmas broadcasts produced by Queen Elizabeth II, one of these from the years 1952–1960[1] (henceforth the Q50 corpus) and the other from the years 1995–2002 (henceforth the Q90 corpus). The median and mean durations of the broadcasts analyzed here are 5 min 36 s and 5 min 28 s, respectively: the shortest was in 1959 (61 s) and the longest in 1956 (8 min 21 s).

The third corpus (henceforth the SSB corpus) was of four adult female news and current affairs presenters from the 'Machine Readable Spoken English Corpus' (Roach, Knowles, Varadi, & Arnfield, 1994) who were also analyzed in Deterding (1997). These speakers were recorded by the BBC in the 1980s, and they are described by Deterding (1997) as having a 'SSB accent', a style of speech that 'may be familiar…through listening to the BBC World Service'. Some of the digitized data in the Q50 corpus (from 1952, 1954, 1957) and their segmentations and labelings of [ɪ], [iː], and [ɪː] were available from our earlier analyses of the Christmas broadcasts. The remaining Christmas broadcasts were obtained from the BBC, digitized at 16 kHz, and cut up into roughly 1 min files.

The [ə] vowels in these three corpora were all taken from polysyllabic content words and care was taken not to include those weak vowels that are realized with a weak [ɪ] in RP (see also Flemming (2004) for an acoustic comparison between these weak vowel types in General American): that is, schwas were only taken from weak vowels transcribed in most dictionaries with /ə/ as in 'Rosa's (e.g., 'affected', 'farmers', 'together', 'wisdom') but not from vowels typically transcribed with weak /ɪ/ (e.g., 'before', 'celebration', 'express', 'promises', 'roses', 'united'). Schwas were not included if their duration was so short that no reliable formant frequencies could be calculated. [ɪ] and [iː] were taken from syllables with primary lexical stress and in prosodically accented words. [ɪː] was extracted from all contexts and differentiated according to whether it occurred in a prosodically (intermediate or intonational) phrase-final or phrase-medial position. In all cases, [ɪː] was word-final either in an orthographic '–y' context ('city', 'ceremony', 'duty', 'happy') or in orthographic 'ies' plurals ('countries', 'remedies', etc). The distribution of these four vowel types in these two corpora is shown in Table 1.

The first four formant frequencies were calculated using the EMU speech database analysis system (Cassidy & Harrington, 2001). The parameters for formant calculation were: LPC-order of 10, a pre-emphasis of 0.95, and a 30 ms Blackman window with a frame shift of 5 ms. For the schwa vowels in the Christmas broadcasts, the fundamental frequency was extracted using a pitch-tracker based on the Schaefer-Vincent (1983) periodicity detector (included as part of the EMU package) with a frame shift of 5 ms. The first two formant frequencies of all vowels were checked manually both by inspecting the trajectories on the spectrogram and by identifying outliers from ellipse plots in the $F1 \times F2$ formant plane. All the formant frequencies were converted to Bark using the formula in Zwicker (1961).

2.2. Speaker normalization

A speaker-normalization technique developed by Lobanov (1971) was applied to the (Bark rescaled) formant frequency data of the [iː ɪ ɪː] vowels in the SSB corpus. In this technique, the formants were converted to z-scores by subtracting the formant mean and dividing by the standard deviation. This was done separately

---

[1]The 1953 Christmas broadcast was of a poor audio quality and was not analyzed. The 1955 broadcast was not available at the time of writing.

Table 1
The distribution of the vowel categories in the Q50 (left), Q90 (middle), and SSB (right) corpora

Year	iː	ɪ	iːF	ɪː	ə	Σ	Year	iː	ɪ	iːF	ɪː	ə	Σ	Speaker	iː	ɪ	iːF	ɪː	ə	Σ
Q50							Q90							SSB						
1952	22	44	15	11	53	145	1995	29	32	6	18	85	170	a	15	20	4	19	79	137
1954	16	34	14	16	76	156	1996	11	39	10	30	58	148	d	28	80	19	47	285	459
1956	24	47	16	26	121	234	1997	20	35	13	25	118	211	e	13	15	8	3	65	104
1957	38	52	12	23	116	241	1998	36	39	10	33	110	228	g	45	56	29	35	181	346
1958	0	25	8	17	75	125	1999	20	29	13	21	143	226	Σ	101	171	60	104	610	1046
1959	3	8	0	9	13	33	2000	20	28	8	16	102	174							
1960	11	23	8	18	90	150	2001	17	20	5	22	114	178							
Σ	114	233	73	120	544	1084	2002	12	20	7	21	77	137							
							Σ	165	242	72	186	807	1472							

The rows of Q50 and Q90 are of years in the Christmas broadcasts. The rows of SSB are of the speakers. The vowel types in row 1 are exemplified by the following words respectively: iː ('fleece', 'heed'), ɪ ('kit', 'hid'), iːF ('cit*y*', 'happ*y*' in phrase-final position), ɪː ('cit*y*', 'happ*y*' in non-phrase-final position) ə ('*a*ffected', 'farm*ers*', 't*o*geth*er*', 'wisd*o*m').

for each speaker and for each formant. The mean and standard deviation were calculated across all the speaker's formant frames (from vowel onset to vowel offset) only for the vowels [iː ɪ ɪː] analyzed in this study. The effect of this technique, then, is to rescale linearly the formant space for each speaker at a number of standard deviations away from a mean of zero. In order to scale the normalized data to Hz values more typical of F1 and F2, another linear rescaling was carried out, $G_n = F_n s_n + m_n$, where F_n is a normalized formant (normalized F1 or F2), s_n is the standard deviation of F1 or F2 across all speakers, m_n is the mean of F1 or F2 across all speakers, and G_n is the resulting rescaled, normalized formant (see Disner, 1980; Harrington & Cassidy, 1999, pp. 76–78, for further details). The same type of *z*-score normalization and rescaling were applied separately to the SSB [ə] (for this second normalization, the mean and standard deviation were calculated across all the SSB speaker's [ə] tokens following the same procedure as above).

2.3. F2-curvature

The comparisons between [iː ɪː ɪ] were based not only on F1 × F2 data extracted at the formant midpoint, but also on the extent of curvature in F2. The curvature parameter was applied to F2 in Hz, since the Bark scale, being quasi-logarithmic, was found to flatten the F2-trajectory to too great an extent, rendering the distinction between [iː] and [ɪ] negligible as far as the trajectory's curvature was concerned.

The measure of the F2-curvature was obtained from the second coefficient of the (discrete) cosine transformation that was applied to the F2-trajectories between the vowel onset and offset. The basis of this transformation is the following: any positive-valued, digital signal (such as a formant trajectory) of *N* points can be decomposed into *N* separate, harmonically related cosine waves between zero and *N*−1 radians/frame such that, when these *N* cosines are summed, the original signal (second formant frequency trajectory in this case) is exactly reconstructed. The cosine coefficients are the amplitudes of these cosines into which the formant trajectory has been decomposed (since these are cosines, there is by definition no phase). The 0th coefficient (i.e., the amplitude of the cosine at a frequency of zero, sometimes known as the d.c. offset) is the mean of the trajectory between its onset and offset; the 1st coefficient is proportional to the linear slope between the formant's onset and offset; and the 2nd coefficient is proportional to the formant's curvature. That is, the closer the 2nd coefficient is to zero, the more the trajectory tends towards a straight line. If the 2nd coefficient is positive, then the trajectory is U-shaped, and if it is negative, it is in the shape of an inverted U. Further details and formulae are given in Watson and Harrington (1999).

2.4. Euclidean distance calculations

Euclidean distance calculations were made of the relative proximity of the tokens of [ɪː] to the categories [iː] and to [ɪ] in the three corpora. This was done by calculating two sets of Euclidean distances: E_1, from the [ɪː]

tokens to the centroid, or to the mean, of [ɪ]; and E_2, from the [ɪː] tokens to the mean of [iː]. The logarithmic ratio between these two sets of Euclidean distances, $\log(E_1/E_2) = \log(E_1) - \log(E_2)$, is then a measure of the relative proximity of [ɪː] to these two vowel categories: if $\log(E_1) - \log(E_2) = 0$, then [ɪː] is as close to [ɪ] as it is to [iː]; if it is negative, [ɪː] is closer to [ɪ]; and if it is positive, then [ɪː] is closer to [iː]. This log ratio parameter was used to compare the relative proximity of [ɪː] to [ɪ] and to [iː] in the three corpora.

In the results to be presented below, all statistical tests were carried out on Bark values (even if the values are reported in Hz), unless otherwise stated. Before applying any *t*-tests, the Shapiro–Wilk test (Royston, 1982) for normality was applied to all the data (using the Shapiro test function in the R programming language) to verify that the data were normally distributed. If they were not, the nonparametric Wilcoxon rank sum test with continuity correction was applied instead.

3. Results

3.1. Schwa vowels in Q50 and Q90

3.1.1. Results

Fig. 1 shows the mean change in the fundamental frequency and in the first four formant frequencies between the Q50 and Q90 corpora. All these data were extracted at the temporal midpoint of the schwa vowels and then averaged in the two sets of Christmas broadcast corpora. Two sets of data are presented in Fig. 1: the changes to all schwas irrespective of context (black bars); and the changes to schwas in the second syllable of 'Christmas' (gray bars) which was the (polysyllabic, content) word in which schwa occurred most frequently in the analyzed broadcasts (for Q50, $n = 32$; for Q90, $n = 30$). The purpose of analyzing these word-specific data was to check that any observed differences in the schwas between the 2-year groups were not an artefact of context.

Considering firstly the data from schwas irrespective of context (Fig. 1, black bars), a comparison of schwas in the Q50 with the Q90 corpora on the acoustic parameters at the schwa midpoint shows the following results. (The statistics were calculated on Bark values; the means are also reported in Hz below, rounded to the nearest integer.)

For the fundamental frequency, there was a significant decrease ($t = 13.1$, df $= 1166.3$, $p < 0.001$) of 0.61 Bark from a mean of 253 Hz in Q50 to a mean of 189 Hz in Q90.
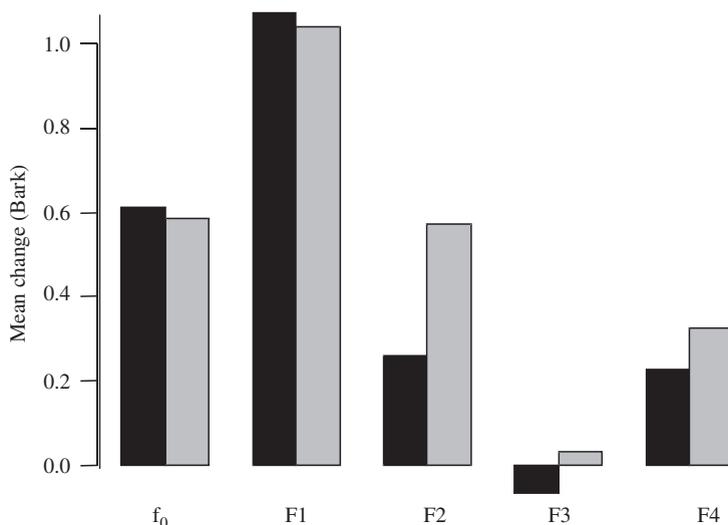

Fig. 1. The average difference between schwa vowels in the Q50 and in the Q90 corpora for the parameters f_0 and the first four formant frequencies extracted at the vowel temporal midpoint. Positive values mean that the values in the Q90 corpus are less than those in the Q50 corpus. Black bars: schwa in all contexts. Gray bars: schwa in the second syllable of 'Christmas'.

For F1, there was a significant decrease ($t = 21.0$, df $= 1213.9$, $p < 0.001$) of 1.07 Bark from a mean of 551 Hz in Q50 to a mean of 428 Hz in Q90.

For F2, there was a significant decrease ($t = 4.8$, df $= 999.2$, $p < 0.001$) of 0.26 Bark from a mean of 1696 Hz in Q50 to a mean of 1624 Hz Q90.

The change in F3 was much less than in F1 or F2 and is in the opposite direction. There was a significant increase ($t = 2.06$, df $= 1130.0$, $p < 0.05$) of 0.07 Bark from a mean of 2834 Hz in Q50 to a mean of 2864 Hz in Q90.

For F4, there was a significant decrease ($t = 7.2$, df $= 1312.6$, $p < 0.001$) of 0.23 Bark from a mean of 4113 Hz in Q50 to a mean of 3962 Hz in Q90.

There was a significant decrease in duration of the schwa vowels ($t = 2.42$, df $= 1040.6$, $p < 0.01$) although the change from Q50 ($m = 63.1$ ms, S.D. $= 35.7$ ms) to Q90 ($m = 58.6$ ms, S.D. $= 30.6$ ms) at 4.5 ms was very small.

Fig. 1 also shows that there is a very similar pattern of results for schwas that occur in the word 'Christmas'. As for schwa in all contexts, there was a significant change and decrease from Q50 to Q90 in F1 (Wilcoxon test, $W = 754.5$, $p < 0.0001$), in F2 (Wilcoxon test, $W = 794$, $p < 0.0001$), and in F4 (Wilcoxon test, $W = 617$, $p < 0.05$). There was no significant change between the year groups in F3 for schwa in 'Christmas' (whereas there was a (just) significant increase in F3 from Q50 to Q90 for all schwas). As far as f_0 is concerned, there was a decrease from Q50 to Q90 for schwa in 'Christmas' which is in the same direction and of a very similar magnitude as for schwa in all contexts, although this decrease was not quite significant (Wilcoxon test, $W = 574$, $p = 0.06$). There was a decrease in duration of the schwa in 'Christmas' from Q50 to Q90, which is in the same direction as for all schwas, although this difference was not significant.

In general, the pattern of changes for schwa in 'Christmas' is very similar to that observed for all the schwas regardless of context.

3.1.2. Discussion

There is a long-term change in the Christmas broadcast schwas between 1952–1960 and 1995–2002 with significant decreases in the fundamental frequency, F1, F2, and F4, and a much smaller, but nevertheless significant, increase in F3. The decrease in fundamental frequency with increasing age is expected and consistent with earlier studies (Decoster & Debruyne, 2000; Helfrich, 1979; Linville, 2001; Linville & Fisher, 1985). Since the pattern of results for schwa in the word 'Christmas' is very similar to that of schwa in all contexts, it cannot be the case that the changes from Q50 to Q90 are an artefact either of segmental context or of word-level prosody (given that the changes have been observed in the same word). It also seems implausible that the changes could be an artefact of utterance-level prosody, if only because it is not clear what kind of utterance-level prosodic effect could bring about the nature of the observed changes (in which f_0, F1, F2, F4 decease, but in which F3 marginally increases).

As far as the formants are concerned, it seems unlikely that these changes can be explained in terms of an overall lengthening or shortening of the vocal tract because then the formants should increase or decrease proportionally by roughly the same amount, and this is very clearly not the case: there is a greater decrease in F1 than in F2 or than in F4, whereas F3 increases rather than decreases.

Another possibility is that the decrease in F1 and F2 comes about because of a smaller mouth opening (Lindblom & Sundberg, 1971) as a result of a greater schwa undershoot in the later broadcasts. The evidence that is compatible with this interpretation is that there is a significant, but very small, decrease in the duration of schwas over the 40-year period which, following Lindblom (1963), would mean that there is less time for the same extent of mouth opening to be attained in schwa productions in later broadcasts. On the other hand, this idea cannot account for the increase in F3 from earlier to later years and the size of the duration decrease (around 5 ms) does not seem large enough to account for much of the 123 Hz lowering in F1 from earlier to later broadcasts.

An F2 and F4 decrease coupled with an F3 increase is, according to perturbation theory, compatible with increasing the constriction of schwa modeled as a straight-sided tube about 1/3 way along the vocal tract from the glottis (Chiba & Kajiyama, 1941). At this distance, the standing waves in a straight-sided tube produce an F2 air-pressure zero (at 1/3 of the vocal tract length from the glottis), an F4 air-pressure zero (at 3/7 of the vocal tract length), but an F3 air-pressure maximum (at 2/5 of the vocal tract length). Since a constriction in a

straight-sided tube causes resonance frequencies to decrease and increase at air-pressure zeros and maxima, respectively, then a constriction at this distance from the glottis, which might be roughly in the upper pharynx, would produce the F2 and F4 decrease, but an F3 increase. On the other hand, there is no reason to expect that the Queen's more modern schwas in the Q90 corpus should be produced with a greater degree of pharyngeal constriction that those in the Q50 corpus and an increased upper pharyngeal constriction cannot explain the considerably greater decrease in F1 from early to later years.

It is possible that perceptual factors play a role. Perception experiments by Traunmüller (1981, 1984, 1991) have shown that the degree of perceived openness changes very little if (Bark-scaled) f_0 and F1 increase or decrease together; his conclusion is that the perception of openness depends on the difference in Bark between F1 and f_0. Perhaps, then, the Queen has compensated for the perceptual influence of a falling f_0 on vowel quality by lowering F1 and F2: that is, in order for the same phonetic quality to be maintained with a lower f_0, F1 and F2 might be shifted downwards in frequency. However, this hypothesis cannot explain why F3 should rise, nor why the drop in F1 from the 50 s to the 90 s is greater than that of F2. Also, the F1–f_0 difference in Q50 (mean = 2.68 Bark) is significantly greater ($t = 6.71$, df = 1282.8, $p < 0.0001$) than in Q90 (mean = 2.22 Bark), so it is not the case that a constant F1–f_0 difference has been maintained from earlier to later broadcasts.

An acoustic interaction between f_0 and F1 could be responsible for the lowering of F1 in the Q90 relative to the Q50 data. Since the f_0 mean in the Q50 corpus is 253 Hz, the second harmonic should be at roughly $2 \times 253 = 506$ Hz, which is close to the measured mean F1 of the Q50 schwa at 551 Hz: that is, the calculated F1 of schwa in the Q50 corpus could be tracking more or less the second harmonic. Since the mean f_0 decrease between the corpora is 64 Hz, then, if the calculated F1 is influenced by the frequency of the second harmonic, the expected F1 drop should be around $64 \times 2 = 128$ Hz which is almost exactly the same as the observed mean F1-decrease of 123 Hz from the Q50 to the Q90 schwas. The idea that the calculated F1-decrease is an acoustic artefact of the f_0-decrease cannot be ruled out, therefore.

In summary, it is not possible to explain with any degree of certainty the pattern of formant changes in schwas between the 1950s and 1990s in the Christmas broadcasts. However, it is clear from these results that any observed F1 or F2 lowering in the [iː ɪː ɪ] vowels needs to be interpreted with caution, since these changes may not be phonetic but, assuming that we do not wish to argue for a diachronic *phonetic* change to [ə], they may instead be due to some as yet elusive, long-term physiological or perceptual change in the vowels over a 40-year period.

3.2. [iː ɪː ɪ] in the Q50 and SSB corpora

3.2.1. Results

As discussed earlier, the aim here was to compare the Queen's 1950s corpus with the 1980s SSB corpus and then to use these observed differences to make a prediction about the positions of the Queen's 1990s vowels, on the assumption that the Queen's vowels have shifted towards the more modern form of RP, SSB.

Fig. 2 shows two-standard deviation ellipses for four vowel categories separately for the Q50 (left) and the SSB corpora (right). The vowel categories are [iː] (FLEECE), [ɪ] (KIT), and [ɪː] (happY), with the latter divided up depending on whether it is in a prosodically phrase-final syllable or not. The average positions of schwa are also shown. All the formant data were extracted at the acoustic vowel midpoint.

The first most obvious difference between these two sets of data is in F1 which is higher for all of the Q50 vowel categories. Taking into account the results of the preceding section, which showed an F1-lowering of schwa between Q50 and Q90, it would be premature and quite possibly incorrect, to attribute these overall differences in F1 between Q50 and SSB corpora to phonetic factors.

On the other hand, as Fig. 2 and Table 2 show, the relationship of [ɪː] to the other vowels is different in the SSB corpus than in the Q50 corpus. In the SSB corpus, both phrase-medial and phrase-final [ɪː] are acoustically between [iː] and [ɪ], whereas in the Q50s corpus, they are much closer to [ɪ].

The differences between these corpora in the relative position of [ɪː] were quantified by calculating the log ratio of the Euclidean distances from the tokens of [ɪː] to the centroids of [iː] and of [ɪ] separately in the two corpora following the methodology in Section 2.4. Negative values on this log ratio measure mean that [ɪː] is closer to [ɪ], positive values that it is closer to [iː]. The difference measures were calculated for the [ɪː] tokens
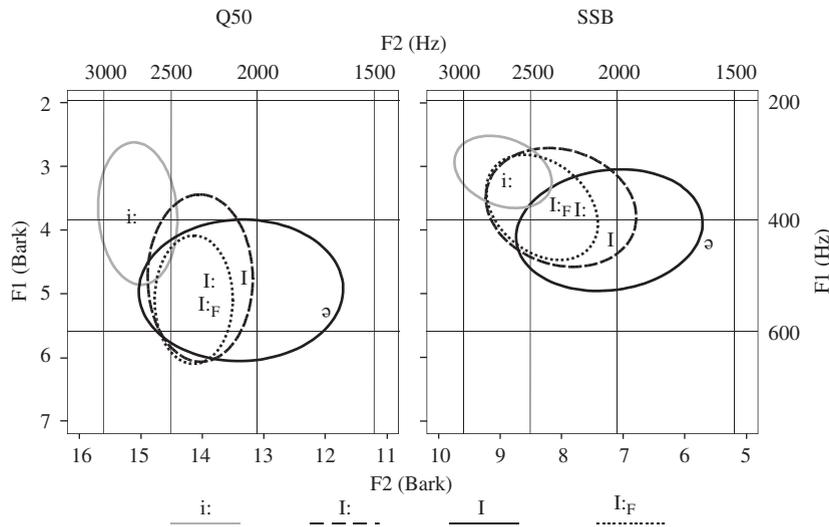
Fig. 2. Two standard-deviation ellipses in the F1 × F2 plane around four vowels in the Q50 corpus (left) and in the SSB corpus (right) for [iː] (solid gray), [ɪ] (solid), phrase-medial [ɪː] (dashed), and phrase-final [ɪː] (dotted). The average position of each of these four vowel categories is marked with the corresponding symbol at the center of the ellipse. The average position of schwa (without its ellipse) is also shown.

Table 2
Mean F1 and F2 in Hz for 5 five categories in the Q50 corpus (left two columns) and in the SSB corpus (columns 3 and 4)

Vowel	Q50		SSB	
	F1	F2	F1	F2
iː	389	2739	317	2697
ɪːF	543	2358	378	2431
ɪː	503	2322	378	2319
ɪ	525	2105	417	2054
ə	551	1696	426	1666

Phrase-medial and phrase-final [ɪː] are denoted by ɪː and ɪːF, respectively.

Table 3
Mean duration of the vowel categories (ms) in the Q50 and SSB corpora

Vowel	Q50	SSB
iː	148	113
ɪːF	170	120
ɪː	104	58
ɪ	85	52
ə	64	52

Phrase-medial and phrase-final [ɪː] are denoted by ɪː and ɪːF, respectively.

collapsed across phrasal position on F1 × F2 Bark values extracted at the midpoint. The results showed that the Euclidean distance ratio was significantly greater (Wilcoxon test, $W = 9950$; $p < 0.001$) in the Q50 corpus ($m = -0.59$) than in the SSB corpus ($m = 0.04$), i.e., that [ɪː] is relatively closer to [ɪ] in the Q50 corpus.

Table 3 shows the average duration of each vowel category. In the SSB corpus, [ə], [ɪ], and phrase-medial [ɪː] all have quite similar absolute average durations (around 50–60 ms) and they are all substantially shorter than
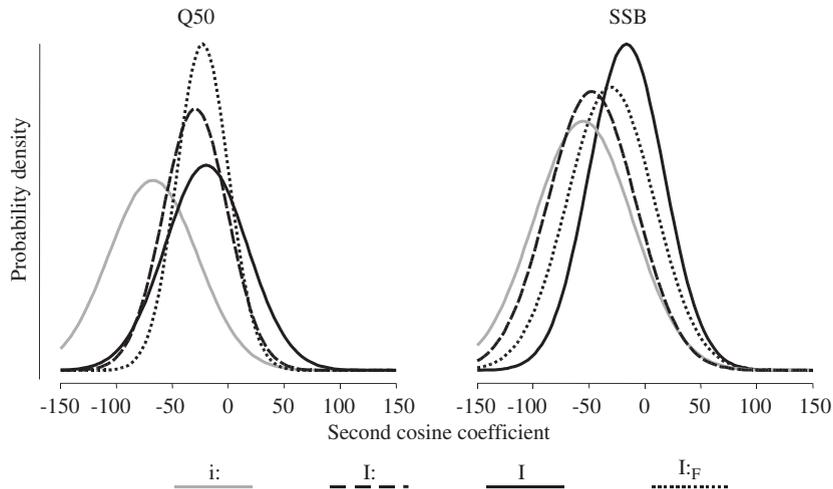
Fig. 3. Normal curves fitted to the distributions [iː] (solid gray), [ɪ] (solid), phrase-medial [ɪː] (dashed), and phrase-final [ɪː] (dotted) on the 2nd cosine coefficient calculated from the trajectory of the second formant in the Q50 corpus (left) and in the SSB corpus (right). The normal distribution of [iː] is to the left of that of [ɪ] in both corpora.

SSB [iː]. There is a very marked effect of phrase position on [ɪː] in the SSB corpus: phrase-final [ɪː] has double the absolute mean duration of phrase-medial [ɪː] and it is also longer than SSB [iː]. In the Q50 corpus, the vowel categories are all longer than in the SSB corpus. Also, whereas in the SSB corpus [iː] has over twice the duration of [ɪ], in the Q50 corpus, the average duration of [iː] is about 1.75 as great as that of [ɪ]. As far as [ɪː] is concerned, it is some 20 ms longer than [ɪ] in the Q50 corpus (whereas they have about the same duration in the SSB corpus); and, as in the SSB corpus, phrase-final Q50 [ɪː] has the greatest duration of all four vowel categories.

Finally, Fig. 3 shows fitted normal distributions of the coefficient of the 2nd cosine fitted to F2 for the four vowel categories. As discussed in 2.3, this parameter measures the extent of formant curvature: values further from zero correspond to a greater curvature in the trajectory. In both the Q50 and SSB corpora, [iː] and [ɪ] are very clearly distinguished from each other: for [iː], the 2nd coefficient has larger negative values (indicating a more curved, inverted-U shape in F2) whereas it is close to zero for [ɪ] in both corpora.

The distribution of [ɪː] is different in the two corpora. Whereas in the SSB corpus, [ɪː] falls between [iː] and [ɪ] on the parameter of F2-curvature, in the Q50 corpus, [ɪː] is hardly distinguished from [ɪ]. The relative position of [ɪː] (collapsed across the phrase position) on this parameter in these two corpora was quantified by calculating the log Euclidean distance ratio from the tokens of [ɪː] to the means of [iː] and [ɪ] following the methodology described in Section 2.4. There were significant differences between the two corpora ($t = 5.06$; df $= 350.2$; $p < 0.001$): the mean values on this log Euclidean distance ratio were -1.17 for the Q50 corpus (i.e., Q50 [ɪː] is relatively closer to Q50 [ɪ] on this parameter) and -0.12 for the SSB corpus (i.e., SSB [ɪː] is also relatively closer to SSB [ɪ] on this parameter, but not as close as in the Q50 corpus).

3.2.2. Discussion

In the F1 × F2 plane, SSB-[ɪː] occupies a position between SSB-[ɪ] and SSB-[iː] whereas Q50-[ɪː] is considerably closer to Q50-[ɪ] than to Q50-[iː]. The extent of F2-curvature for [ɪː] is about intermediate between that of [iː] and [ɪ] in the SSB-corpus but in the Q50 corpus, the F2-trajectory of [ɪː] is flatter and more like that of [ɪ]. Taken together, these formant differences are consistent with the view that SSB-[ɪː] is more peripheral, more like [iː], and tenser than Q50-[ɪː]. Therefore, if Q90-[ɪː] has shifted in the direction of SSB-[ɪː], then it should be acoustically closer to Q90-[iː] in the F1 × F2 plane and it should have a more curved F2-trajectory compared with that of Q50-[ɪː]. These predictions are tested in the next section.

It is more difficult to interpret phonetically the [ɪː] duration differences between the Q50 and SSB corpora. Phrase-final [ɪː] is very long in both corpora, and (predictably) much longer than phrase-medial [ɪː]. The main difference is in the relationship of phrase-medial [ɪː] to [ɪ]. In the SSB corpus, these vowels have roughly the

same duration, whereas in the Q50 corpus, phrase-medial [ɪː] is a good deal longer than [ɪ]. If Q90-[ɪː] has become more like SSB-[ɪː], then the durational differences between Q90-[ɪː] and Q90-[ɪ] should be less than in those in the Q50 corpus, although such a change, if found, might well be due to an effect of speech tempo, rather than to a shift in the phonetic quality of [ɪː].

3.3. [iː ɪː ɪ] in the Q50 and Q90 corpora

3.3.1. Results

Fig. 4 and Table 4 show, compatibly with the results obtained for the Queen's schwas in Section 3.1, that F1 is lower for all vowel categories in the Q90 corpus. Secondly, the extent of lowering in [ɪː ɪ ə] is about twice that of [iː] and, as a result, [ɪː ɪ ə] are all acoustically closer to [iː] on F1 in the Q90 than in the Q50 corpus.

There is also an overall F2-lowering between the Q50 and Q90 corpora, which is also compatible with the results for schwa, but here the differences between the vowel categories are more marked. The average F2-lowering from Q50 to Q90 is negligible in all three of [iː] (4 Hz) phrase-final [ɪː] (14 Hz), and phrase-medial [ɪː] (23 Hz); on the other hand, F2-lowering between Q50 and Q90 is a good deal more marked in [ɪ] (95 Hz) and in schwa (72 Hz). Since F2 has decreased in [ɪ] but not in [ɪː], it follows that [ɪː] is further from [ɪ] on F2 and therefore relatively closer to [iː] in the Q90 corpus.
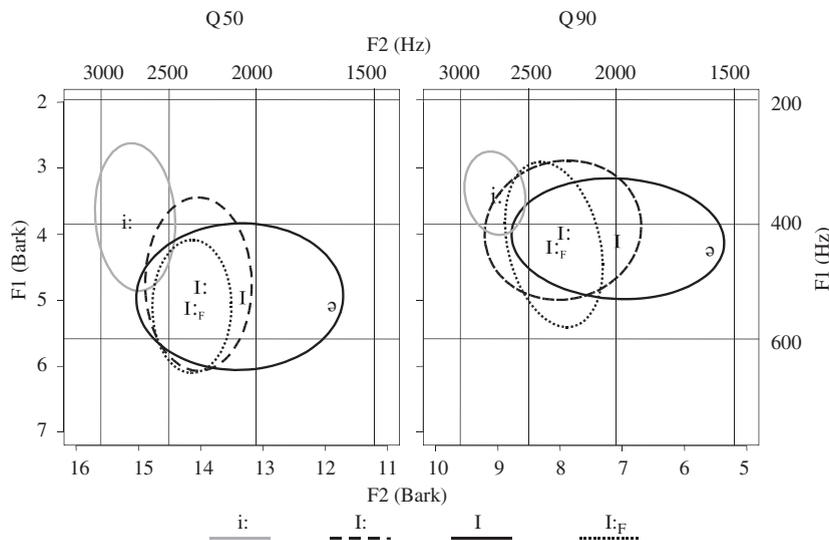

Fig. 4. Two standard-deviation ellipses around [iː] (solid gray), [ɪ] (solid), phrase-medial [ɪː] (dashed), and phrase-final [ɪː] (dotted) in the Q50 corpus (left) and in the Q90 corpus (right). The average position of schwa in both these corpora is also shown.

Table 4
Average F1 and F2 values in Hz for 5 five categories in the Q50 corpus (left two columns) and in the Q90 corpus (columns 3 and 4)

Vowel	Q50		Q90	
	F1	F2	F1	F2
iː	389	2739	349	2735
ɪːF	543	2358	435	2344
ɪː	503	2322	411	2299
ɪ	525	2105	425	2010
ə	551	1696	428	1624

Phrase-medial and phrase-final [ɪː] are denoted by ɪː and ɪːF, respectively.

The observations about the relative position of [ɪː] were quantified by calculating log Euclidean distance ratios as discussed in Section 2.4. Based on the two-parameter F1 × F2 space, these calculations showed that [ɪː] was significantly closer to [iː] than to [ɪ] (Wilcoxon test, $W = 10815$, $p < 0.0001$) in the Q90 corpus ($m = -0.18$) than in the Q50 corpus ($m = -0.59$). However, it is also clear from Fig. 4 and Table 4 that this difference comes about, not necessarily because [ɪː] has moved closer to [iː] over the 40-year period, but instead because [ɪ] and [ə] have moved away from both [ɪː] and [iː].

Turning now to duration, all vowel categories are on average shorter in the Q90 than the Q50 corpus (Table 5). The average percentage decrease is: [iː] (6.1%), phrase-final [ɪː] (13.5%), phrase-medial [ɪː] (14.4%), [ɪ] (24.7%), and [ə]: (6.3%). The greatest change is therefore in [ɪ] which, in the Q90 corpus, has an average duration that's almost the same as that of schwa. But there is no evidence from the duration data that [ɪː] is closer to [iː] in the Q90 than in the Q50 corpus.

Finally, the data on F2-curvature as measured by the 2nd cosine coefficient are shown in Fig. 5. Here there is some evidence of a rearrangement in the distribution of the vowel categories towards those seen in the SSB corpus. Recall that the F2-curvature for [ɪː] was significantly closer to that of [iː] in the SSB compared with the Q50 corpus. As Fig. 5 shows, the 2nd cosine coefficient for [ɪː] is closer to [iː] in the Q90 corpus than in the Q50 corpus, which suggests that it has greater curvature. Based on the log Euclidean distance ratio, the relative proximity of [ɪː] to the means of [iː] and [ɪ] was significantly different ($t = -4.8$; df $= 426.5$; $p < 0.001$) in the two corpora. In both corpora, [ɪː] was closer to [ɪ] as shown by the negative values on this parameter (Q50: $m = -1.17$; Q90: -0.37), but since in Q90 the value was closer to zero, [ɪː] is relatively closer to [iː] on F2-curvature than in Q50.

Table 5
Mean duration of the vowel categories (ms) in the Q50 and Q90 corpora

Vowel	Q50	Q90
iː	148	139
ɪːF	170	147
ɪː	104	89
ɪ	85	64
ə	64	59

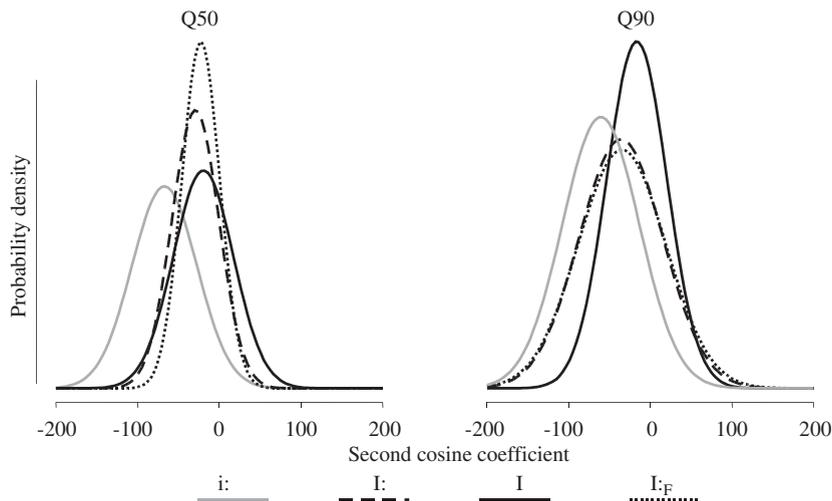Phrase-medial and phrase-final [ɪː] are denoted by ɪː and ɪːF, respectively.

Fig. 5. Normal curves fitted to the distributions [iː] (solid gray), [ɪ] (solid), phrase-medial [ɪː] (dashed), and phrase-final [ɪː] (dotted) on the 2nd cosine coefficient calculated from the trajectory of the second formant in the Q50 corpus (left) and in the Q90 corpus (right).
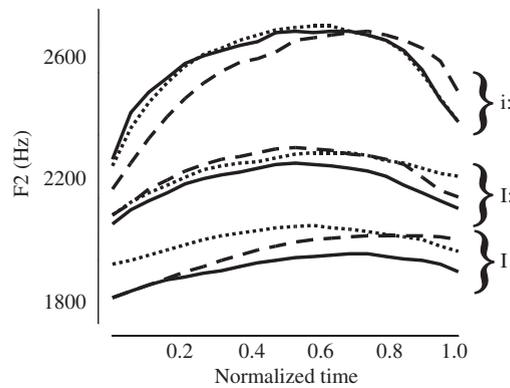
Fig. 6. Averaged, linearly time-normalized F2 trajectories in the Q50 (dotted), Q90 (dashed), and SSB (solid) corpora for [iː], [ɪː], and [ɪ].

In order to check that the differences between the year groups on this parameter were not an artefact of context, the same calculations were made but on some of the most frequently occurring word types that have a final [ɪː]. Across both Christmas broadcast corpora, the three most frequently occurring words with final [ɪː] were 'happy' ($n = 27$), 'many' ($n = 29$), and 'family' ($n = 31$). The distribution across the two corpora of these three word types was as follows: 'happy' (Q50, $n = 14$; Q90, $n = 13$); 'many' (Q50, $n = 9$; Q90, $n = 20$); 'family' (Q50, $n = 21$; Q90, $n = 10$). For all three word types, the pattern on the log Euclidean distance ratio of the 2nd cosine coefficient was the same as for all [ɪː] vowels considered together, with the following mean values: 'happy' (Q50, $m = -0.95$; Q90, $m = 0.62$); 'many' (Q50, $m = -1.14$; Q90, $m = -0.17$); 'family' (Q50, $m = -1.72$; Q90, $m = -1.12$). The differences for 'happy' on this parameter were significant (Wilcoxon test, $W = 48$, $p < 0.05$), but not for the other two words. This need not indicate a word-specific sound change, but may instead simply come about because the tokens in 'many' and 'family' are too few and too imbalanced across the year groups. But what this additional result does show is that the shift in [ɪː] towards [iː] from Q50 to Q90 is unlikely to be an artefact of segmental context, given that the shift was observed within the same word types and also taking into account that across these word types the left contexts (/p m l/ for 'happy', 'many', 'family') are different.

Fig. 6 shows F2-trajectories that were linearly time-normalized and averaged separately for the three vowel categories in the three corpora. Compatibly with the normal distributions in Fig. 5, the curvature is greatest in all three corpora for [iː], intermediate for [ɪː], and least for [ɪ]. Moreover, there is some evidence for a greater curvature in Q90 and SSB-[ɪː] than in Q50-[ɪː]: in both Q90- and SSB-[ɪː], F2 drops more extensively from the F2 peak (near normalized time point 0.5) to the end of the trajectory than in the Q50 corpus.

4. Discussion

There are three main findings in this study. The first is that there are long-term changes to the Queen's schwas over a 40-year period such that F1, F2, and F4 have lowered and F3 has raised. The average extent of lowering is greatest in F1 at just over 1 Bark and almost twice as great as the lowering of f_0. It is difficult to explain the formant changes as the result of physiological effects of vocal tract maturation—these acoustic changes cannot be accounted for by an overall increase in vocal tract length nor by an increase in the length of the pharynx relative to the rest of the vocal tract. In view of various experiments showing that the extent of vowel openness depends perceptually not just on F1, but also on the difference between F1 and f_0, then F1 may have decreased to compensate perceptually for the lower f_0 in more recent Christmas broadcasts. However, perceptual compensation still does not explain why the extent of lowering is greater in F1 than in F2 and why F3 should have risen.

The second finding is that 1980s SSB or mainstream RP has a tenser final [ɪː] vowel compared with the Queen's 1950s [ɪː], where 'tenser' is defined as a greater proximity in the F1 × F2 plane to [iː] and a more
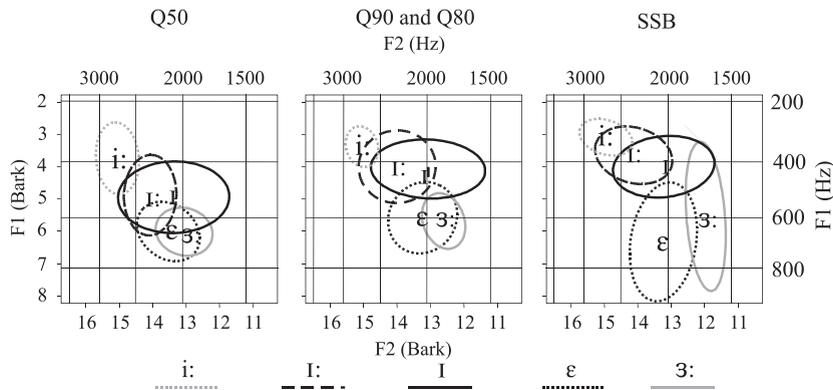
Fig. 7. Two standard-deviation ellipses in the F1 × F2 plane around [iː] (dotted gray), [ɪ] (solid), [ɪː] (dashed), [ɛ] (dotted), and [ɜː] (solid gray) in the Q50 (left), Q90 (middle), and SSB (right) corpora.

curved F2-trajectory. The generalization from this finding is that a more modern, mainstream RP has a tenser [ɪː] than its more conservative and/or aristocratic counterpart.

The third finding is that the Queen's [ɪː] has become 'tenser' in later broadcasts, where 'tenser' has the same definition as above, but the shift, while significant, is marginal. This third finding shows a far smaller shift in the direction of SSB compared with the monophthongal or diphthongal changes reported in Harrington et al. (2000a, b, 2005).

Although there is only a very marginal acoustic shift in the Christmas broadcast [ɪː] towards a tenser vowel, some of the 1950s [ɪː] vowels seem to have a more open quality than those in later broadcasts: a selection of [ɪː] vowels from the 1950s and 1990s corpora are in q50.wav and q90.wav respectively.[2] Perhaps the more open quality of the 1950s [ɪː] has something to do with a shift, not just in relation to [iː] and [ɪ], but relative to some of the mid and central vowels. Unfortunately, the mid vowels in these 1995–2002 broadcasts have not yet been acoustically analyzed, so this question cannot yet be completely resolved. However, it is possible to make use of some of the mid vowel data from Harrington et al. (2000a, b) in which they, as well the other RP monophthongs, were analyzed in Christmas broadcasts from the 1950s (1952, 1954, 1957) and from the 1980s (1983, 1985, 1988). The left and right panels of Fig. 7 shows the three vowel categories from the present study together with the 1950s and 1980s [ɛ] (DRESS) and [ɜː] (NURSE) spaces from Harrington et al. (2000b).[3] The important point to note from this figure is that in both the Queen's later broadcasts and in the SSB data, [ɪ ɪː] are acoustically further away from the mid vowels [ɛ ɜː] on the F1-axis; in the Q50 corpus by contrast, there is a much greater overlap between these two sets of vowel spaces. Taken together, these data suggest that RP [ɪ] may have risen phonetically in the last 50 years relative to the mid vowel space and that this raising is also manifested in the Christmas broadcasts, although not to the same extent as in the SSB speakers: that is, the F1 difference between [ɪ] and [ɛ ɜː] is greatest in SSB, least in the Queen's 50s broadcasts and intermediate in the Queen's later broadcasts.

The only suggestion in the literature that KIT has raised is in Bauer (1985), who comments (based on auditory impressions) that RP [ɪ] is becoming a closer vowel in stressed syllables. On the other hand, raising of the KIT vowel may not be especially perceptible, both because [ɪ] is short and because the raising may have been gradual over 50 years and not that extensive (or certainly not as extensive as the [ɪ]-raising in Australian English—see, e.g., Harrington, Cox, & Evans, 1997). However, raising *is* likely to be much more perceptible in happY, both because [ɪː] is often long (especially so in prosodically phrase-final position) and because any raising of this vowel, which already seems to have been phonetically fronter than [ɪ] in the 1950s (see the left panel of Fig. 7), is more likely to push it into part of the vowel space that has the greatest probabilistic

---

[2]The sound files can currently be downloaded from the on-line version of this article at Supplementary data doi:10.1016/j.wocn.2005.08.001.

[3]RP is non-rhotic. The vowel in 'herd' is a long, central, mid vowel.

association with FLEECE. Perhaps, then, it is the raising of the KIT vowel in all contexts (including [ɪ] and [ɪː] studied here) that has led to the more specific perceptual association of happY with FLEECE. The increased perceptual association between these two vowels might then have been the trigger for more modern RP speakers to produce [ɪː] with a fronter vowel quality as well, thus producing an even greater overlap, both in height *and* fronting, between [ɪː] and [iː] in more modern times.

The Guest Editors have pointed out to me that this argument that happY has become associated with FLEECE because of the greater phonetic proximity between KIT and FLEECE in RP does not seem to be compatible with the distribution of present-day New Zealand English (NZE) front vowels. In NZE, there is happY-tensing on the one hand (i.e., happY is close phonetically to FLEECE), but a very marked phonetic difference between KIT and FLEECE on the other (KIT is a central vowel in present-day NZE, whereas the vowel in FLEECE is high and front—see, e.g., Gordon et al., 2004; Watson, Harrington, & Evans, 1998; Watson, Maclagan, & Harrington, 2000). The centralization of KIT in NZE is related to the raising of the front vowels /æ/ (TRAP) and /ɛ/ (DRESS) that has taken place in the last 60 years (Bauer, 1979, 1986; Maclagan & Hay, 2004) and there is now evidence that this was a push-chain originating from the raising of TRAP and DRESS (Gordon et al., 2004). If NZE KIT raised slightly before it centralized (perhaps under pressure from the raising of TRAP and DRESS), then there would have been a stage in its development when it was phonetically quite close to FLEECE. Therefore, the conditions for NZE happY to become associated with FLEECE would have been met perhaps some 50–60 years ago, assuming at that time that NZE KIT and NZE FLEECE were phonetically close to each other. Once NZE happY became associated with FLEECE, then there would be no reason for happY to follow KIT along its subsequent path of centralization. Compatibly, NZE happY-tensing seems to have been slightly in advance of KIT centralization (Gordon et al., 2004). In summary, there is not necessarily any contradiction between RP and NZE as far as the proposed mechanism is concerned by which happY has become associated with FLEECE.

It is still difficult to explain, however, why the Queen seems to have participated in the height shift (although this still needs to be verified more rigorously by analyzing mid vowel data from 1995 to 2002 in the Christmas broadcasts) but only marginally in the fronting shift. An explanation could be based on the lowering of F1 from early to later years observed in schwas earlier: the F1-lowering might perhaps not be phonetic at all, but instead an effect of long-term changes to the vocal tract. On the other hand, it is clear from Fig. 7 that F1-lowering has not applied to the vowel categories in equal measure: F1 has lowered from early to later broadcasts much more in [ɪ ɪː] than in the mid vowels DRESS and NURSE and this suggests that some of the F1-lowering might be phonetic, i.e. a phonetic change in which [ɪ ɪː] have raised relative to the mid vowel space. If some of the observed F1-lowering is phonetic raising of [ɪ ɪː], and if raising took place diachronically before fronting, then raising would have had more time to have an effect on the Christmas broadcasts, especially if the bulk of the raising happened between the 1950s and 1970s. This could explain why the effect on the Christmas broadcast [ɪː] vowels of raising is greater than the effect of fronting.

One of the Reviewers has suggested that the changes observed in this paper might be due to changes of speaking style. Thus, the mode of delivery in the Christmas broadcasts might have become more casual to match the increasing informality in society in the last 50 years. In Harrington (in press), I have considered three types of evidence that are relevant to this issue. Firstly, a more casual speaking style can be expected to be hypoarticulated relative to a more formal mode of delivery, with an increase in vowel target undershoot and coarticulation (e.g., Lindblom, 1990; Moon & Lindblom, 1994). However, although Harrington (in press) found the fronting of /uː/ (GOOSE) from earlier to later Christmas broadcasts to be associated with greater coarticulatory overlap with the preceding consonant, there was also clear evidence in Harrington (in press) for increased *hyper*articulation in the phonetic lowering of /æ/ (TRAP) from earlier to later broadcasts. Secondly, a more casual speaking style and the accompanying hypoarticulation can be expected to shrink the vowel space (e.g., de Jong, 1995). In Harrington (in press), it was shown that, when /uː/-fronting is considered in relation to /iː/ (FLEECE) and /ɔː/ (THOUGHT, FORCE), there was evidence for an *expansion* of the vowel space from earlier to later years together with some evidence of chain-shifting, which again implies a phonetic change. Thirdly, the idea that the Queen introduces more casual variants in later broadcasts implies a certain degree of volition or intention i.e., that the sound changes are implemented by the Queen for a specific purpose (of sounding less formal, less aristocratic, more approachable, and so on). In order for such a deliberate, intended change of speaking style to have any effect, it must presumably be perceptible (otherwise there would

be no point in changing the speaking style) and therefore manifest itself as a change between perceptibly different phonetic variants. For example, in an analysis of Opray Winfrey's speech, Mendoza-Denton, Hay, and Jannedy (2003) show how Winfrey switches between various degrees of monophthongization of /aɪ/ depending in part on her orientation towards her discussant. But this type of switching between different phonetic variants is not what has been found in any of the longitudinal analyses of the Christmas broadcasts so far. Instead, and as shown in Harrington (in press), there is a very gradual phonetic shift over a 50-year period, such that /æ/ is incrementally lowered from one decade to the next. It may of course be that there is an incremental waning of formality with a concomitant increment in vowel openness from decade to decade, but even if such a link could be demonstrated, the size of the shift makes such a relationship unlikely: Harrington (in press) found an F2-shift in /uː/ from the 1950s to the present day of some 300 Hz, and the size of the acoustic phonetic changes in happY observed in this paper are even less. The size of these shifts, while by no means insignificant in comparing 1952 with 2002, is nevertheless very small in relation to the time period over which they have taken place and it may be so small (for example, an increase of some 60 Hz per decade in F2 of /uː/ assuming a linear change) that the shift is not perceptible from year to year. I am proposing, then, that as sound change takes place in the community it has a gradual, and perhaps imperceptible influence on adults' phonetic categories over a long period of time in a way that is reminiscent of the Neogrammarian position that sound change proceeds gradually and imperceptibly (Paul, 1880; see also Bloomfield, 1933). These incremental shifts are also predicted from an exemplar theory of speech perception (Johnson, 1997; Pierrehumbert, 2002, 2003) in which an adult's phonetic categories can be updated by auditory traces of remembered word tokens produced in the community. The further prediction from exemplar theory that the extent of phonetic change should vary depending on word frequency has not yet been tested in these longitudinal data.

At the same time, I certainly do not want to suggest that all sound change takes place in adults subliminally and involuntarily. Groups within a community clearly do introduce innovative sound changes and these may be resisted by other community members if, in the model of Lindblom, Guion, Hura, Moon, and Willerman (1995), they have a high 'social value': that is, if they represent too great a change across a social class division. Thus even if the Southern British English speaking community is increasingly replacing a syllable-final /t/ with a glottal stop, as Wells (1997) has suggested (e.g., [nɒʔ əʊnlɪ] for 'not only'), it is quite likely that such a change would be resisted by both the Queen and elder, established members of the community, because it marks such a strong shift towards a Cockney accent, i.e. towards a variant used by speakers who have traditionally been low down in the social hierarchy. Perhaps then, the incremental changes that have been found in the Christmas broadcasts are those that have a low social value. That is, these types of phonetic changes are not subject to introspection because they do not have any social significance attached to them: they are subliminal in the same way that much allophonic variation, such as the fronting or backing of /k/ is subliminal for most English listeners, until there is a reason for them to be made aware of it.

Speculating further, another explanation for the raising but not fronting of happY in these Christmas broadcasts could be that these two changes are associated with different social values. Whereas *raising* the vowel in happY may be socially unmarked, *fronting* it to the extent that it overlaps extensively with FLEECE is a marker of Estuary English (Wells, 1997), a form of RP that has been strongly influenced by Cockney and which may be resisted by the Queen.

Thus the more extensive raising, but only marginal fronting, could have the effect that the Queen does not sound old-fashioned on the one hand, but also still distinctly upper-class and very clearly differentiated from Estuary English on the other. However, further acoustic analyses especially longitudinal studies from other speakers of RP will be necessary to test these speculations further.

Acknowledgments

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.wocn.2005.08.001.

References

Bailey, G., Wikle, T., Tillery, J., & Sand, L. (1991). The apparent time construct. *Language Variation and Change*, *3*, 241–264.

Bauer, L. (1979). The second great vowel shift? *Journal of the International Phonetic Association*, *9*, 57–66.

Bauer, L. (1985). Tracing phonetic change in the received pronunciation of British English. *Journal of Phonetics*, *13*, 61–81.

Bauer, L. (1986). Notes on New Zealand English phonetics and phonology. *English World-Wide*, *7*, 225–258.

Bloomfield, L. (1933). *Language*. London: Allen & Unwin.

Brink, L., & Lund, J. (1975). *Dansk Rigsmål I–II. Lydudviklingen siden 1840 med særligt henblink på sociolekterne i København.* Copenhagen: Gyldendal.

Bybee, J. (2001). *Phonology and language use*. Cambridge: Cambridge University Press.

Cannadine, D. (1998). *Class in Britain*. New Haven: Yale University Press.

Cassidy, S., & Harrington, J. (2001). Multi-level annotation in the Emu speech database management system. *Speech Communication*, *33*, 61–77.

Chiba, T., & Kajiyama, M. (1941). *The vowel: Its nature and structure*. Tokyo: Tokyo Publishing.

Coggle, P. (1993). *Do you speak Estuary?*. London: Bloomsbury.

Crystal, D. (1995). Estuary English. In D. Crystal (Ed.), *Cambridge encyclopaedia of the English language* (p. 327). Cambridge: Cambridge University Press.

Decoster, W., & Debruyne, F. (2000). Longitudinal voice changes: Facts and interpretation. *Journal of Voice*, *14*, 184–193.

de Jong, K. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *Journal of the Acoustical Society of America*, *97*, 491–504.

Deterding, D. (1997). The formants of monophthong vowels in Standard Southern British English pronunciation. *Journal of the International Phonetic Association*, *27*, 47–55.

Disner, S. (1980). Evaluation of vowel normalization procedures. *Journal of the Acoustical Society of America*, *67*, 253–261.

Flemming, E. (2004). Rosa's roses: Reduced vowels in American English. *Journal of the Acoustical Society of America*, *115*, 2541.

Fletcher, J. (1991). Rhythm and final lengthening in French. *Journal of Phonetics*, *19*, 193–212.

Gimson, A. C., & Cruttenden, A. (1994). *Gimson's pronunciation of English (5th ed., revised by Alan Cruttenden)*. London: Edward Arnold.

Gordon, E., Campbell, L., Hay, J., Maclagan, M., Sudbury, A., & Trudgill, P. (2004). *New Zealand English: Its origins and evolution*. Cambridge: Cambridge University Press.

Harrington, J. (in press). Evidence for a relationship between synchronic variability and diachronic change in the Queen's annual Christmas broadcasts. In J. Cole & J. Hualde (Eds.), *Laboratory Phonology IX*. Berlin: De Gruyter.

Harrington, J., & Cassidy, S. (1999). *Techniques in speech acoustics*. Dordrecht: Kluwer Academic Publishers.

Harrington, J., Cox, F., & Evans, Z. (1997). An acoustic phonetic study of broad, general, and cultivated Australian English vowels. *Australian Journal of Linguistics*, *17*, 155–184.

Harrington, J., Palethorpe, S., & Watson, C. (2000a). Does the Queen speak the Queen's English? *Nature*, *408*, 927–928.

Harrington, J., Palethorpe, S., & Watson, C. (2000b). Monophthongal vowel changes in received pronunciation: An acoustic analysis of the Queen's Christmas broadcasts. *Journal of the International Phonetic Association*, *30*, 63–78.

Harrington, J., Palethorpe, S., & Watson, C. (2005). Deepening or lessening the divide between diphthongs? An analysis of the Queen's annual Christmas broadcasts. In W. J. Hardcastle, & J. Beck (Eds.), *A figure of speech: A Festschrift for John Laver* (pp. 227–261). Mahwah, NJ: Lawrence Erlbaum.

Helfrich, H. (1979). Age markers in speech. In K. Scherer, & H. Giles (Eds.), *Social markers in speech* (pp. 63–107). Cambridge: Cambridge University Press.

Henton, C. (1983). Changes in the vowels of received pronunciation. *Journal of Phonetics*, *11*, 353–371.

Huntley, R., Hollien, H., & Shipp, T. (1987). Influences of listener characteristics on perceived age estimations. *Journal of Voice*, *1*, 49–52.

Jacques, R., & Rastatter, M. (1990). Recognition of speaker age from selected acoustic features as perceived by normal young and older listeners. *Folia Phoniatrica*, *42*, 118–124.

Johnson, K. (1997). Speech perception without speaker normalization. In K. Johnson, & J. Mullenix (Eds.), *Talker variability in speech processing* (pp. 146–165). San Diego: Academic Press.

Jun, S.-A. (1995). Asymmetrical prosodic effects on the laryngeal gesture in Korean. In B. Connell, & A. Arvaniti (Eds.), *Papers in laboratory phonology IV* (pp. 235–253). Cambridge: Cambridge University Press.

Labov, W. (1994). *Principles of linguistic change. Vol. 1: Internal factors*. Oxford: Blackwell.

Labov, W. (2001). *Principles of linguistic change. Vol. 2: Social factors*. Oxford: Blackwell.

Laver, J. (1994). *Principles of phonetics*. Cambridge: Cambridge University Press.

Laver, J., & Trudgill, P. (1979). Phonetic and linguistic markers in speech. In K. Scherer, & H. Giles (Eds.), *Social markers in speech* (pp. 1–26). Cambridge: Cambridge University Press.

Lee, S., Potamianos, A., & Narayanan, S. (1999). Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *Journal of the Acoustical Society of America, 105*, 1455–1468.

Lenneberg, E. (1967). *Biological foundations of language*. New York: Wiley.

Lehiste, I., & Peterson, G. (1961). Transitions, glides, and diphthongs. *Journal of the Acoustical Society of America, 33*, 268–277.

Lindblom, B. (1963). Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America, 35*, 1773–1781.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In W. J. Hardcastle, & A. Marchal (Eds.), *Speech production and speech modelling* (pp. 403–439). Dordrecht: Kluwer Academic Publishers.

Lindblom, B., Guion, S., Hura, S., Moon, S.-J., & Willerman, R. (1995). Is sound change adaptive? *Rivista Di Linguistica, 7*, 5–37.

Lindblom, B., & Sundberg, J. (1971). Acoustical consequences of lip, tongue, jaw, and larynx movement. *Journal of the Acoustical Society of America, 50*, 1166–1179.

Linville, S. (2001). *Vocal aging*. San Diego: Singular Thomson Learning.

Linville, S., & Fisher, H. (1985). Acoustic characteristics of perceived versus actual vocal age in controlled phonation by adult females. *Journal of the Acoustical Society of America, 78*, 40–48.

Linville, S., & Rens, J. (2001). Vocal tract resonance analysis of aging voice using long-term average spectra. *Journal of Voice, 15*, 323–330.

Lobanov, B. (1971). Classification of Russian vowels by different speakers. *Journal of the Acoustical Society of America, 49*, 606–608.

Luchsinger, R., & Arnold, G. E. (1965). *Voice– Speech– Language*. Belmont: Wadsworth Publishing Company.

Maclagan, M., & Hay, J. (2004). The rise and rise of New Zealand English DRESS. In S. Cassidy, F. Cox, R. Mannell, & S. Palethorpe (Eds.), *Proceedings of the 10th Australian speech science and technology conference* (pp. 183–188). Sydney: Australian Speech Science and Technology Association Inc.

Mendoza-Denton, M., Hay, J., & Jannedy, S. (2003). Probabilistic sociolinguistics: Beyond variable rules. In R. Bod, J. Hay, & S. Jannedy (Eds.), *Probabilistic linguistics* (pp. 97–138). Cambridge: MIT Press.

Meurer, E. M., Wender, M., Corleta, H., & Capp, E. (2004). Phono-articulatory variations of women in reproductive age and postmenopausal. *Journal of Voice, 18*, 369–374.

Milroy, J. (1994). Neogrammarian sound change from a social dialectological perspective. In G. Melchers, & N.-L. Johannesson (Eds.), *Nonstandard varieties of language* (pp. 118–130). Stockholm: Almqvist & Wiksell International.

Moon, S.-J., & Lindblom, B. (1994). Interaction between duration, context and speaking style in English stressed vowels. *Journal of the Acoustical Society of America, 96*, 40–55.

Neiman, G., & Applegate, J. (1990). Accuracy of listener judgments of perceived age relative to chronological age in adults. *Folia Phoniatrica, 42*, 327–330.

Paul, H. (1880). *Prinzipien der Sprachgeschichte (2nd ed., 1986)*. Leipzig: Niemeyer (1886).

Pierrehumbert, J. (2002). Word-specific phonetics. In C. Gussenhoven, & N. Warner (Eds.), *Papers in laboratory phonology VII* (pp. 101–139). Berlin: Mouton de Gruyter.

Pierrehumbert, J. (2003). Probabilistic phonology. In R. Bod, J. Hay, & S. Jannedy (Eds.), *Probabilistic linguistics* (pp. 177–228). Cambridge: MIT Press.

Ptacek, P., & Sander, E. (1966). Age recognition from voice. *Journal of Speech and Hearing Research, 9*, 273–277.

Rakerd, B., & Verbrugge, R. (1985). Linguistic and acoustic correlates of the perceptual structure found in an individual differences scaling study of vowels. *Journal of the Acoustical Society of America, 71*, 296–301.

Rastatter, M., McGuire, R., Kalinowski, J., & Stuart, A. (1997). Formant frequency characteristics of elderly speakers in contextual speech. *Folia Phoniatrica, 49*, 1–8.

Roach, P., Knowles, G., Varadi, T., & Arnfield, S. (1994). MARSEC: A machine-readable spoken English corpus. *Journal of the International Phonetic Association, 23*, 47–54.

Rosewarne, D. (1994). Estuary English: Tomorrow's RP? *English Today, 37*, 3–8.

Royston, P. (1982). An extension of Shapiro and Wilk's *W* test for normality to large samples. *Applied Statistics, 31*, 115–124.

Sankoff, G. (2004). Adolescents, young adults and the critical period: Two case studies from "Seven Up". In C. Fought (Ed.), *Sociolinguistic variation*. Oxford: Oxford University Press.

Sankoff, G. (in press). Cross-sectional and longitudinal studies in sociolinguistics. In U. Ammon, N. Dittmar, K. Mattheier, & P. Trudgill (Eds.), *Sociolinguistics: An international handbook of the science of language*. Berlin: De Gruyter.

Schaefer-Vincent, K. (1983). Pitch period detection and chaining: Method and evaluation. *Phonetica, 40*, 177–202.

Strange, W. (1989). Dynamic specification of coarticulated vowels spoken in sentence context. *Journal of the Acoustical Society of America, 85*, 2135–2153.

Traunmüller, H. (1981). Perceptual dimension of openness in vowels. *Journal of the Acoustical Society of America, 69*, 1465–1475.

Traunmüller, H. (1984). Articulatory and perceptual factors controlling the age- and sex-conditioned variability in formant frequencies of vowels. *Speech Communication, 3*, 49–61.

Traunmüller, H. (1991). The context sensitivity of the perceptual interaction between f_0 and F1. In *Proceedings of the 12th international congress of phonetic sciences* (Vol. 5, pp. 62–65), Aix-en-Provence.

Trudgill, P. (1974). *The social differentiation of English in Norwich*. Cambridge: Cambridge University Press.

Trudgill, P. (2001). *Sociolinguistic variation and change*. Edinburgh: Edinburgh University Press.

Ueyama, M. (1999). An experimental study of vowel duration in phrase-final contexts in Japanese. *UCLA Working Papers in Phonetics, 97*, 174–182.

Umeda, N. (1975). Vowel duration in American English. *Journal of the Acoustical Society of America, 58*, 434–445.

van Bergem, D. (1994). A model of the coarticulatory effects on the schwa. *Speech Communication, 14*, 143–162.

Watson, C. I., & Harrington, J. (1999). Acoustic evidence for dynamic formant trajectories in Australian English Vowels. *Journal of the Acoustical Society of America*, *106*, 458–468.

Watson, C. I., Harrington, J., & Evans, Z. (1998). An acoustic comparison between New Zealand and Australian English vowels. *Australian Journal of Linguistics*, *18*, 185–208.

Watson, C. I., Maclagan, M., & Harrington, J. (2000). Acoustic evidence for vowel change in New Zealand English. *Language Variation and Change*, *12*, 51–68.

Weinreich, U., Labov, W., & Herzog, M. (1968). Empirical foundations for a theory of language change. In W. Lehmann, & Y. Malkiel (Eds.), *Directions for historical linguistics* (pp. 97–195). Austin: University of Texas Press.

Wells, J. C. (1982). *Accents of English*. Cambridge: Cambridge University Press.

Wells, J. C. (1997). Whatever happened to received pronunciation? *Il Jornadas de Estudios Ingleses*, *2*, 19–28 Also available from http://www.phon.ucl.ac.uk/home/estuary/home.htm.

Whiteside, S. (2001). Sex-specific fundamental and formant frequency patterns in a cross-sectional study. *Journal of the Acoustical Society of America*, *110*, 464–478.

Xue, S., & Hao, G. (2003). Changes in the human vocal tract due to aging and the acoustic correlates of speech production: A pilot study. *Journal of Speech Language and Hearing Research*, *46*, 689–701.

Yaeger-Dror, M. (1994). Phonetic evidence for sound change in Quebec French. In P. A. Keating (Ed.), *Papers in laboratory phonology 3* (pp. 267–292). Cambridge: Cambridge University Press.

Zwicker, E. (1961). Subdivision of the audible frequency range into critical bands (Frequenzgruppen). *Journal of the Acoustical Society of America*, *33*, 248.