

TEMPORAL-BASED SPEAKER SEX DIFFERENCES IN READ SPEECH: A PRELIMINARY STUDY

Sandra P. Whiteside

Department of Speech Science, University of Sheffield, Sheffield S10 2TA, United Kingdom.

ABSTRACT

Speaker sex differences have been shown to exist in the acoustic signals of read speech. Some of these studies have looked at fundamental frequency differences and formant frequency differences (Peterson & Barney, 1952; Wu & Childers, 1991; Childers & Wu, 1991). Other studies have investigated differences in first formant bandwidths and amplitudes (Karlsson, 1988). However, there is also some evidence to suggest that differences also exist in the temporal domain (Byrd, 1992). This paper presents the results of a preliminary acoustic-phonetic investigation into speaker sex differences in the temporal domain. It focuses on the read speech data of three men and three women speakers with a British General Northern accent. The speech material consisted of ten repetitions of five sentences (British General Northern accent APLAWD sentences) for each of the six speakers. The durations of all three hundred sentences were measured. Significant differences were found between the data of the men and women speakers. Further investigation was made into whether vowels and consonants were either elided or reduced. Each sentence was investigated. Differences were found between the consonant and vowel realisations of the men and women speakers. For example, the men speakers showed a tendency to reduce or elide vowels and consonants and the women speakers showed evidence of pausing more at syntactic boundaries. The findings are discussed.

1. INTRODUCTION

When making judgements about whether we are perceiving a woman or a man speaking, we depend on a multitude of factors which we integrate to form the final percept. These factors may relate to the style of delivery, the use of particular language, the use of particular intonation patterns and the perceived pitch of the speaker for example. Some acoustic-phonetic investigations have explored how speaker sex differences are perceived by listeners through instrumental analysis. These have shown that acoustic phonetic differences exist between the read speech of men and women speakers. Some of these studies have shown that fundamental frequency differences exist between men and women with men having, on average, lower fundamental frequencies (Aronovitch, 1976; Coleman, 1973a). This can be explained in part by their larger larynges. In addition women have been found to have, on average, higher formant frequencies (Coleman, 1976; Henton, 1986; Peterson & Barney, 1952 amongst others) because of their smaller vocal tracts. Women have also been found to have wider first formant bandwidths and lower first formant amplitudes (Fant, 1979b; Karlsson, 1988) due to differences in glottal source properties. Differences in the speaking rate of read speech have also been observed between men and women speakers of American English in the TIMIT database (Byrd, 1992). Byrd (1992) found that sex had a significant effect on speaking rate. She states that under the recording conditions used for the TIMIT database, 'women speak reliably more slowly than do men' (p.594) and that men tended to

TEMPORAL-BASED SPEAKER SEX DIFFERENCES

reduce their vowels to [ə] more often than the women (p. 595). She also found that female speakers in the TIMIT database released stops in sentence-final position more frequently and produced more glottal stops than male speakers. These findings were statistically significant.

The aim of this preliminary study is to establish whether any temporal differences exist between the read sentences of three men and three women speakers in a controlled laboratory setting. The speakers in question represent a British General Northern accent.

It has been suggested that linguistic and stylistic conventions affect both men's and women's vocal characteristics (Mattingly, 1966). This may therefore explain why women's voices sometimes sound 'smaller' and higher pitched whereas men's voices sound 'bigger' and lower pitched than the dimensions of their vocal apparatus suggest. This study also proposes to suggest tentatively whether these linguistic and stylistic conventions have an effect on the temporal aspects of read speech. On this basis in addition to investigating the durations of a specific group of sentences read by a small group of men and women, other acoustic-phonetic phenomena will be examined. These will include the incidence of: reducing consonant and vowels; eliding consonants and vowels and introducing pauses at phrase-level boundaries. Any interaction between these acoustic-phonetic observations and any effect of speaker sex will serve to illustrate that even in a controlled laboratory situation, where there is no direct interaction between subject and experimenter, men and women may adopt different speaking styles when reading sentences.

2. METHOD

2.1 Subjects

Three male and three female adult native speakers of English served as speakers. All speakers came from North of England. Four (2 men and 2 women) were speakers from Leeds and the remaining two were from Sheffield (1 man) and Lancashire (1 woman). All speakers represented a British General Northern accent which can be defined as a "a non-rhotic accent of Standard English characterised in the vowel system by COULD/CUD and GAS/GLASS rhyming and a tendency to retain strong vowels where RP shows weakening e.g. computer /kɒmpjʊtə/ (Wells, 1982).

2.2 Speech material

Ten repetitions of five sentences were read by three men and three women speakers. This made a total of three hundred sentences (150 for the men and 150 for the women). Both the sentences and the speakers formed part of the APLAWD (Lindsey et al, 1987) speech corpus of British General Northern (GN) accent speakers.

The list of the sentences is as follows:-

Sentence One.: *George made the girl measure a good blue vase.*

Sentence Two.: *Why are you early you owl?*

Sentence Three.: *Cathy hears a voice amongst Spar's data.*

Sentence Four.: *Six plus three equals nine.*

Sentence Five.: *Be sure to fetch a file and send theirs off to Hove.*

2.3 Recording procedures

High quality recordings were made in a sound proof studio at the University of Leeds, using recording procedures described by Lindsey et al. (1987). All recordings were made with a

Sony SL-F1E recorder and a Sony PCM (Pulse Code Modulation) F1 digital processor using high quality Sony Betamax video cassettes. Audio cassette copies of the high quality Sony Betamax video cassettes were made using a Tandberg TCD440A cassette deck. The speech pressure waveforms of all three hundred sentences were played back on a Sony Professional Walkman and digitised at a sampling rate of 11 kHz onto a Macintosh LCII computer using a Farallon™ Macrecorder and Signalyze™.

2.4 Analysis

Using Signalyze™ (Keller, 1992) the durations of each of all the three hundred sentences were measured. It must be noted at this stage that these durations did not include any of the pauses observed in the sentences as the observed pauses were subtracted from overall sentence duration. However, the incidence of pauses was noted and some of these observations are discussed below. In addition each group of sentences (60 data items for each sentence) was examined for specific linguistic and acoustic phonetic phenomena as outlined below. Some additional analyses were carried out using Voice Identification Inc. 700 series spectrograms. Observations were tested for statistical significance using a statistical package (Statview™).

Sentence 1 (/ 'mɛɪd ɪə 'gɜ:l ,meɜ:(ɪ)ə 'gʊd 'blu: 'vɔ:z/)

Firstly, the occurrence of schwa elisions in 'measure a' was examined. Whether the men and women speakers realized the utterance as ['mɛɜ:(ɪ)ə], ['mɛɜ:(ɪ)] or ['mɛɜ:(ɪ)ə] was noted using auditory and acoustic analysis. Speech pressure waveform and spectrogram patterns were time aligned and analysed to determine whether a distinct vowel was present following the fricative [ʒ]. It was hypothesised that elisions would contribute to shorter sentence durations. In addition the occurrence of pauses after 'girl' was investigated. It was predicted that the occurrence of pauses would coincide with a lengthening in the duration of 'girl'.

Sentence 2 (/ 'wɜ: 'ɑ:/ə ju: 'ɜ:l ju: 'aʊl/)

Firstly, it was noted whether 'are' was fully represented as the vowel [ɑ], reduced as the schwa [ə] or elided altogether. Auditory and acoustic analysis was used to make these decisions. The presence of a schwa was noted if there was a separate intensity peak in the speech pressure waveform. Again it was hypothesised that any reductions and/or elisions would contribute to shorter sentence durations. In addition the incidence of pauses after 'early' was investigated.

Sentence 3 (/ 'kæθɪ 'hɒz ə 'vɔ:ɪs ə'mʌŋst 'spɑ:z 'deɪə/)

Whether speakers paused after 'voice' was noted. In addition the duration of 'voice' was measured for each speaker. As for Sentence 1 it was hypothesised that the occurrence of pauses would coincide with longer durations in the word, therefore contributing to longer sentence durations. Thirdly, the duration of the /sts/ cluster in 'amongst Spar's' was measured. Whether speakers realized this cluster as a reduced form ([s s], [s]) or as a full representation([sts]) was also noted. The criteria used for these judgments included auditory and acoustic analysis. Speech pressure waveforms were used for the acoustic criteria where: [sts] was realized as fricative followed by a closure phase and a subsequent pulse/transient which was followed by a fricative; [s s] was realized as two fricatives separated by a reduction in amplitude in the speech pressure waveform and [s] was realized as a single

TEMPORAL-BASED SPEAKER SEX DIFFERENCES

fricative. Again, it was predicted that any reductions would result in reduced cluster durations and therefore contribute to reduced sentence durations.

Sentence 4 (/siks plus 'θri: i:kwɒlz 'nɑ:m/)

The vowel pair in /θri: i:kwɒlz/ was examined to see whether speakers had fully realized the vowels acoustically or reduced them to a single vowel. Auditory analyses were also used in this procedure. Additionally, the duration of /θri: i:/ was measured. It was hypothesised that the occurrence of a pause after 'three' would coincide with the full representation of the vowel cluster. On this premise the occurrence of a pause after 'three' was examined for each of the 60 sentences. Again any pauses and/or full representations would most likely be related to longer sentence durations.

Sentence 5 (/bɪ 'ʃɔ: tə 'feɪ ə 'faɪl ən(d) 'send 'θeəz of tə 'həʊv/)

For the final group of sentences, whether speakers paused after 'file' was noted. The duration of the word 'file' was also measured for each of the 60 sentences using the speech pressure waveform and supporting auditory analysis. It was hypothesized that the occurrence of a pause after 'file' would coincide with longer word durations for 'file'.

3. RESULTS AND DISCUSSION

3.1 Sentence 1

Table 1 shows the mean sentence durations of Sentence 1 and the standard deviation values for both the men and women speakers. These indicate that overall, the women's productions are longer than those of the men, 3045.2 ms versus 2479.1 ms respectively. They also manifest a wider range in their duration values which is indicated by the higher standard deviation value (483.2 ms versus 113.9 ms for the men and women speakers respectively). A one factor ANOVA test revealed that these differences were statistically significant ($DF = 59$, $F = 38.997$, $p = 0.0001$). Subsequently a one factor ANOVA was carried out to see whether there was any correlation between the occurrence of schwa elisions and sentence durations for all speakers. A very highly significant difference ($DF = 59$, $F = 28.17$, $p = 0.0001$) was found between the number of schwa elisions in the men's and women's data. The mean duration of the 23 sentences which contained the elisions was 2438.4 ms compared to 2963.3 ms for the 37 sentences which did not contain any elisions. A one factor ANOVA test was carried out to see whether these schwa elisions were linked to speaker sex. This revealed that there was no significant correlation between the occurrence of elisions and speaker sex ($DF = 59$, $F = 0.62$, $p = 0.4343$). From this it was deduced that other speaker phonetic realizations were contributing to the significant differences in the sentence durations.

It was conjectured that the occurrence of pauses after 'girl' would coincide with longer durations of 'girl'. On this premise a one factor ANOVA test was carried out to see whether there was any link between pausing after 'girl' and longer durations. A one factor ANOVA test revealed that there was a lengthening effect on 'girl' when pauses were present ($DF = 59$, $F = 136.017$, $p = 0.0001$). In the 11 cases where pauses were present, 'girl' had a mean duration of 441.8 ms (s.d. 45ms) whereas in the 49 cases where pauses were not present the mean duration was 251.1 ms (s.d. 49.8ms). All the pauses had been produced by the women speakers. The mean value of the duration of 'girl' for the women speakers was 334.2 ms (s.d. 99ms) whereas the mean duration value of 'girl' for the men was 237.9 ms (s.d. 37.8ms).

TEMPORAL-BASED SPEAKER SEX DIFFERENCES

Table 1
Mean Sentence Durations and Standard Deviations of Sentences 1 to 5 by speaker sex

Sentence	Men		Women	
	Mean (ms)	s.d.	Mean (ms)	s.d.
1	2479.1	113.9	3045.2	483.2
2	1420.6	94.1	1872.2	259.1
3	2325.9	131.2	2978.4	493.0
4	1706.6	177.2	2124.0	236.7
5	2705.1	224.0	3442.0	541.3

3.2 Sentence 2

Table 1 shows the mean sentence durations of sentence 2 and the s.d. values for both men and women speakers. These also show that overall, the women's productions are longer than those of the men (1872.2 ms versus 1420.6 ms respectively). Additionally the women's sentences show a wider range in their duration values. This is manifested in the s.d. values of 259.1ms and 94.1 ms for the women and men respectively. A one factor ANOVA revealed that these sex differences were statistically significant ($DF = 59$, $F = 80.529$, $p = 0.0001$). The investigation into schwa elisions showed that these were linked to speaker sex. A one factor ANOVA revealed that this finding was very highly significant ($F = 16.789$, $p = 0.0001$) with the women showing fewer elisions than the men (none versus 11 respectively).

On investigating the incidence of the full representation of 'are' as [a] it emerged that of the total 22 full representations, 17 were made by the women speakers, whereas only 5 had been made by the men. A one factor ANOVA test showed that this was highly significant ($DF = 59$, $F = 12.069$, $p = 0.001$). Of the total 38 reductions elisions and the realisation of 'are' as a schwa ([ə]), 25 (11 elisions and 14 schwa realisations) and 13 (all schwa realisations) represented the data of the men and women respectively. A one factor ANOVA test was used to check the effect of reductions on sentence durations. This finding was very highly significant ($DF = 59$, $F = 39.576$, $p = 0.0001$) where the mean for the 22 sentences with full representations of 'are' was 1894.3 ms (s.d. 320.7 ms) and the mean for the 38 sentences with reductions (including elisions) was 1502.9 ms (s.d. 161.7 ms). With regard to elisions, a lower but still significant effect was found on sentence duration using a one factor ANOVA ($DF = 59$, $F = 8.511$, $p = 0.005$) where the 11 sentences with elisions had a mean duration of 1422.8 ms (s.d. 101.7 ms) whereas those with no elisions (including reductions and full representations) had a mean duration of 1696.6 ms (s.d. 305.7 ms). The evidence here therefore suggests that both the reduction and the elision of 'are' contributed to the shorter sentence durations observed for the men speakers.

Given the findings for Sentence 1 it was predicted that other speaker sex differences would be found in the sentence. Consequently the incidence of pauses after 'early' was investigated. Findings revealed 17 pauses for the women speakers and one pause for the men. These results were found to be significant using a one factor ANOVA test ($DF = 59$, $F = 29.696$, $p = 0.0001$). Although no absolute duration measurements were made for 'early', visual comparisons of 700 series spectrograms suggested that the duration of 'early' was longer for

TEMPORAL-BASED SPEAKER SEX DIFFERENCES

the women speakers. In addition the women's spectrograms showed a lower tendency to assimilate the /t/ from 'early' and the /j/ from 'you'. It was noted that the women paused after 'early' and therefore realized them as two distinct segments. Conversely, the men tended not to pause and realized the two phonemes as one phonetic segment. Evidence therefore suggests that a number of factors are contributing to the temporal differences in this read sentence.

3.3 Sentence 3

Table 1 shows the mean sentence durations of Sentence 3 and the s.d. values for both the men and women speakers. These show that overall, the women's productions are longer than those of the men, 2978.4 ms versus 2325.9 ms respectively. Like Sentences 1 and 2, the women revealed a wider range in their duration values with a standard deviation of 493 ms versus a value of 131 ms for the men. A one factor ANOVA test revealed that these results were very highly significant ($DF = 59$, $F = 49.087$, $p = 0.0001$).

On examination of the incidence of pauses after 'voice', 14 were made by the women speakers and only three by the men. A one factor ANOVA showed that this finding was significant ($DF = 59$, $F = 11.505$, $p = 0.0013$). From this, a difference in the duration of the men's and women's productions of 'voice' was predicted. Results confirmed this with an average duration of 483.8 ms for the women speakers and 391.3 ms for the men. As was observed for the sentence durations, the women showed a greater variability in these duration values with a higher standard deviation value of 126.9 ms, versus 45.3 ms for the men speakers. These duration values were found to be highly significantly different using a one factor ANOVA ($DF = 59$, $F = 14.132$, $p = 0.0004$). The evidence from this data sample therefore seems to suggest that the incidence of pauses is not only related to speaker sex but that the incidence of pauses also has an effect on the durations of words preceding a pause. These results agree with those found for Sentence 1. That a number of factors had contributed to the duration differences between the men and women speakers prompted the decision to investigate the phonetic realisation of /sts/ in 'amongst Spar's'. In addition the duration of the cluster was measured for each sentence.

Results showed that on average, the duration of the women's clusters was 308.4 ms versus 145.0 ms for the men. Standard deviation values of 172.2 ms and 39.6 ms were found for the women and men respectively. A one factor ANOVA revealed a very highly significant difference between the two speaker sex groups for the cluster durations ($DF = 59$, $F = 25.67$, $p = 0.0001$). Whether the speakers had fully realized the cluster as [sts] or reduced it to [s s] or [s] was noted. Of the 41 reductions found, 12 were made by the women whereas 29 were made by the men. A one factor ANOVA showed that this was very highly significant ($DF = 59$, $F = 34.208$, $p = 0.0001$). This suggested a greater tendency for the women in this data sample to fully realise consonant clusters in read speech. Moreover, these full realizations were linked to their longer sentence durations. This is further supported by the observation that the average duration of the 41 reduced clusters was 159.2 ms while the 19 that were fully realized had an average duration of 372.4 ms. A one factor ANOVA showed this to be very highly significant ($DF = 59$, $F = 47.823$, $p = 0.0001$).

3.4 Sentence 4

Table 1 shows the mean sentence durations of Sentence 4 and the standard deviation values for both speaker groups. Again, the women's sentences are longer than those of the men with

TEMPORAL-BASED SPEAKER SEX DIFFERENCES

values of 2124.0 ms and 1706.6 ms respectively. As noted previously, the women showed a greater variability in their durations with a standard deviation value of 236.7 ms versus that of 177.2 ms for the men. These results were found to be very highly significant using a one factor ANOVA test ($DF = 59$, $F = 59.753$, $p = 0.0001$).

Investigations into whether the vowel pair in /θri: ikwɔl/ was fully represented or reduced to a single vowel revealed that of the 20 reductions, 19 were made by the men speakers. An ANOVA revealed that this was very highly significant ($DF = 59$, $F = 39.479$, $p = 0.0001$), therefore indicating a greater tendency for the women in this study to fully realise vowels in read speech. Durations of /θri: i/ revealed mean values of 543.5 ms (s.d. 87.6 ms) and 365.5 ms (s.d. 85.8 ms) for the women and men respectively. This was found to be very highly significant using a one factor ANOVA ($DF = 59$, $F = 63.214$, $p = 0.0001$).

The occurrence of a pause after 'three' (/θri: /) was investigated. It was predicted that any occurrence of pauses would coincide with an increase in duration. A one factor ANOVA revealed that this was the case ($DF = 59$, $F = 34.952$, $p = 0.0001$). For the 27 cases preceded by a pause, a mean duration of 538.1 ms (s.d. 82.2 ms) was found versus 386.2 ms (s.d. 110.8 ms) for the 33 cases which were not preceded by a pause. Results also showed that of the 27 pauses that were observed, 20 were made by the women and only 7 were made by the men. An ANOVA showed that this was highly significant ($DF = 59$, $F = 13.576$, $p = 0.0005$). The evidence here, therefore suggests that because the women paused more at the phrase boundary, they fully realized the vowels. However in those cases where there were no pauses, full vowel realizations were manifested as two distinctive peaks in the speech pressure waveforms. Moreover, auditory analysis confirmed the presence of two vowels. On average therefore, the men in this data sample paused less frequently than the women, they also reduced the vowel pair which contributed to their shorter sentence durations.

3.5 Sentence 5

Table 1 shows the mean sentence durations of sentence 5 and the standard deviation values for both the men and women speakers. These also show a longer mean duration for the women speakers (3442.0 ms) and a larger standard deviation value (541.3 ms) compared to the men speakers (2705.1 ms and 224.0 ms respectively). An ANOVA showed that this was very highly significant ($DF = 59$, $F = 47.464$, $p = 0.0001$). The occurrence of pauses after 'file' revealed that from a total of 38 pauses, 22 were made by the women and 16, by the men. An ANOVA revealed that this was not significant ($DF = 59$, $F = 2.61$, $p = 0.1116$). However an ANOVA did reveal a very highly significant difference in the mean duration of 'file' when it preceded a pause (520.3 ms, s.d. 59.4 ms). This compared to 390.2 ms (s.d. 51 ms) when it did not precede a pause. An ANOVA also revealed a very highly significant difference between the mean durations of 'file' for the women and men speakers ($DF = 59$, $F = 16.859$, $p = 0.0001$). Here values were 512.4 ms (s.d. 93.7 ms) and 432.8 ms (s.d. 49.7 ms) respectively. Evidence here therefore seems to suggest that while there is a strong relationship between pausing after 'file' and its longer duration, the occurrence of pausing here is not significantly linked to speaker sex. However, what is evident, is that the women's productions of 'file' are still significantly longer than those of the men therefore contributing to their overall longer duration of Sentence 5.

4. GENERAL CONCLUSIONS

The results of this preliminary investigation provide some acoustic-phonetic evidence that the men and the women in this data sample realise sentences differently when they are read in a controlled laboratory situation. From the evidence found for Sentence 1 for example, it appears that for this small group of speakers, there is a link between speaker sex and pausing. In addition, there is also a link between the occurrence of a pause and the duration of the word preceding a pause. The effect of phrase and sentence-final lengthening has been reported elsewhere (Klatt, 1976). The findings here differ from those of Byrd (1992) who found that there was no link between speaker sex and the occurrence of pauses. However what is interesting to note is that the findings here mirror some of the evidence of previous research which has shown that men tend to pause less frequently than women during a conversational speech setting. By not pausing men tend to dominate a conversation as this reduces both turn taking and any interruptions. Conversely women tend to pause more thus allowing themselves to be interrupted more frequently (Elyan, 1978; Tannen, 1990).

The differences between Byrd's findings and those here could be due to cultural differences between British English and American English speakers. That cultural differences exist in conversational style has been reported elsewhere (Tannen, 1990). However it is also possible that these differences are a result of individual differences in the speakers of this study. These differences are manifested in the differences in speaking rate with women tending to read at a slower rate compared to their male counterparts. It is also shown that the women display greater variability in their sentence duration. These results agree with those of Byrd (1990) who also found that in the American English TIMIT database, the men were faster speakers than the women and that the women displayed larger standard deviations than the men.

On closer examination it is apparent that the men in this study tended to either elide or reduce both vowels and consonants. On the whole, these reductions and elisions contributed to shorter sentence durations. Conversely the women showed a tendency to realise speech segments more fully. This therefore meant that the women's speech segments were on average longer than those of the men. These results agree with the finding that women enunciate more clearly than men (Kramer, 1976 as cited in Elyan, 1978). The findings here may also reflect the different strategies adopted by men and women in a conversational setting. However this is pure conjecture at this stage and reflects the need for further research. It must be stressed that while the speaker sex differences found here are compelling, these findings need to be followed up with greater numbers of speakers and a richer repertoire of speech data. Nevertheless this preliminary study indicates some of the ways in which speaker sex appears to have influenced the way in which read speech was delivered by a specific group of British English speakers (General Northern Accent) in a controlled laboratory setting.

5. REFERENCES

- Aronovitch, C. D. (1976). The voice of personality: stereotyped judgements and their relation to voice quality and sex of speaker. *Journal of Social Psychology*, 99, 207-220.
- Byrd, D. (1992). Preliminary results on speaker-dependant variation in the TIMIT database. *Journal of the Acoustical Society of America*, 92(1), 593-596.

Proceedings of the Institute of Acoustics

TEMPORAL-BASED SPEAKER SEX DIFFERENCES

- Childers, D. G. & Wu, K. (1991). Gender recognition from speech. Part II: Fine analysis, *Journal of the Acoustical Society of America*, 90(4), 1841-1856.
- Coleman, R. O. (1971). Male and female voice quality and its relationship to vowel formant frequencies, *Journal of Speech and Hearing Disorders*, 14, 565 - 577.
- Coleman, R. O. (1973a). Speaker identification in the absence of intersubject differences in glottal source characteristics, *Journal of the Acoustical Society of America*, 53, 1741-1743.
- Coleman, R. O. (1973b). A comparison of the contributions of two vocal characteristics to the perception of maleness and femaleness in the voice. *Speech Transmission Laboratory-Quarterly Progress and Status Report, Royal Technical Institute, Stockholm*, 2-3, 13-22.
- Coleman, R. O. (1976). A comparison of two voice quality characteristics to the perception of maleness and femaleness in the voice. *Journal of Speech and Hearing Research*, 19, 168-180.
- Elyan, O. (1978). Sex differences in speech style, *Women Speaking*, 4, 4-8.
- Fant, G. (1979b). Temporal fine structure of formant damping and excitation. *Speech Communication Papers, Acoustical Society of America*, 161-165.
- Henton, C. G. (1986). *A comparative study of phonetic sex-specific differences across languages*. D. Phil Thesis, University of Oxford, UK.
- Karlsson, I. (1988). Glottal Waveforms for normal female speakers, *Speech Transmission Laboratory-Quarterly Progress and Status Report, Royal Technical Institute, Stockholm*, 31-36.
- Keller, E. (1992). *Signal Analysis for Speech and Sound*. InfoSignal™ Inc.
- Klatt, D. H. (1976). Linguistic uses of segmental duration in English: acoustic and perceptual evidence, *Journal of the Acoustical Society of America*, 59, 1208-1221.
- Lass, N. J., Hughes, K. R., Bowyer, M. D., Waters, L. T. & Bourne, V. T. (1976). Speaker Identification from voiced, whispered and isolated vowels, *Journal of the Acoustical Society of America*, 59, 675-678.
- Lindsey, G., Breen, A & Nevard, S. (1987). Spar's Archivable Actual-Word Databases, *UCL Report on SPAR Project*.
- Mattingly, I. G. (1966). Speaker Variation and Vocal Tract Size, *Journal of the Acoustical Society of America*, 39, S1219A.
- Peterson, G. E. & Barney, H. L. (1952). Control methods used in the study of vowels, *Journal of the Acoustical Society of America*, 24, 175-184.
- Tannen, D. (1992). *You Just Don't Understand*, Virago Press.
- Wells, J. C. (1982). *Accents of English 2*. Cambridge University Press.
- Wu, K. and Childers, D. G. (1991). Gender recognition from speech. Part I: Coarse analysis, *Journal of the Acoustical Society of America*, 90(4), 1828-1840.

6. ACKNOWLEDGEMENTS

I would like to thank Celia Scully, Marion Shirt and Eric Brearley at the University of Leeds, UK for collecting the APLAWD database. In addition I would like to thank David Pearce at GEC Marconi, Hirst Research Centre, UK for the use of the APLAWD data in this report. I would also like to thank Peter Roach at the University of Leeds for his helpful comments during the writing of the paper.