ABSTRACT

This study reports on an investigation into the coarticulation patterns of a set of young adult male monozygotic (identical) twins, based on a speech corpus of monosyllabic words. The monosyllabic words were of the structure CVC, and were chosen to represent a variety of consonantal and vowel contexts. Coarticulation patterns were investigated by deriving F2 locus equations [e.g., 9, 10, 11] and examining patterns of formant frequency changes (excursions) of the first three formant frequencies (F1 to F3 in Hz) by the place of articulation of the initial consonants. Rates of formant frequency change were also examined. Results indicated that there were both similarities and differences in the twins' coarticulation patterns. For example, both twins displayed similar coarticulation patterns across bilabial, alveolar, velar and glottal places of articulation, which was supported by similar F2 locus equation functions for both twins. In contrast, there were some inter-twin differences in formant frequency changes, and rates of formant frequency changes. The details of these results are presented and discussed.

1. INTRODUCTION

Speaker identification is an area of research that has received some attention within the fields of forensic phonetics [e.g., 5, 6], and the speech technology area of speaker verification [e.g., 2]. There is a group of speakers that provide an interesting source of data into how nature and nurture may determine speech and voice characteristics.

There is some acoustic evidence to suggest that the coarticulation patterns in identical twins' voices are generally very similar [7]. However, there is other evidence, which suggests that the voices of identical twins are sufficiently different to warrant them having unique voice characteristics [2, 7]. More specifically, both the extent and nature of similarities and differences in the micro-behavioural acoustic data of twins' speech and voice characteristics may provide us with valuable insights into the motor speech behaviours of monozygotic twins. However to date, only a few investigations into the similarities and differences of voice and speech characteristics of identical twins have been reported in the literature [2, 3, 7, 8].

The use of both static and dynamic acoustic phonetic parameters in both the temporal and frequency domains are particularly useful in shedding light on specific aspects of fine motor speech control. Furthermore, these parameters may be important in revealing the exact nature and extent of similarities and differences in the habitual speech characteristics of identical twins.

Coarticulation patterns in the speech of individuals may be a useful means of assessing the nature and characteristics of habitual speech patterns. The nature and extent of coarticulation can be measured, using a number of different techniques/metrics. Included in this list of these techniques/metrics are F2 Locus Equations [4, 9, 10, 11], the extent of formant frequency changes (or excursions) [4], and the temporal (or time course) patterns of formant frequency changes. Locus equations are useful phonetic descriptors [9, 10, 11] which depict the linear relationship between the F2 midvowel frequencies (plotted along the x-axis) and F2 vowel onset frequencies (plotted along the y-axis) of CVC syllables. Locus equations are expressed by simple regression functions as:

\[ F2_{vowel} = k \times F2_{midvowel} + c \]

where \( k \) represents the slope of the function and \( c \), the y-intercept. It has been established that the slopes of these regression lines vary with the place of articulation [9, 10, 11] and that the steepness of these slopes are indicative of the extent to which consonant and vowels coarticulate, with steeper slopes suggesting higher degrees of coarticulation [9, 10, 11].

This study reports on an investigation into the coarticulation patterns of a set of identical twins using F2 Locus Equations, the extent of formant frequency changes, and the rates of formant frequency changes. The coarticulation patterns of the identical twins are highlighted, presented and discussed with reference to the extent of the similarities and differences in their coarticulation patterns.

2. METHOD

2.1. Subjects

A pair of monozygotic twins participated in the study. On an impressionistic level, their voices were judged to be very similar in quality. Their accent and individual speaking styles or idiolects were also judged to be very similar. The twins were 21 year-old Southern Irish males with no speech or hearing problems, and shall be named B and D for the purposes of this study.

2.2. Speech stimuli

The twins were recorded using a Sony DAT recorder (model TCDD100) and high quality microphone (model ECMMS907). They were both recorded on the same day under the same conditions. They read five word lists in a sound treated room. Each word list consisted of the same 32 words but in different random orders so that five productions of each were obtained altogether. The words were Consonant-Vowel-Consonant (CVC)
monosyllables, with the consonants /b/, /d/, /g/ and /h/, and a variety of vowel contexts. The entire list of monosyllabic words was as follows: bob, bib, dad, did, dude, dead, deed, dub, gig, gag, bolog, daub, dub, dub, dad, god, dog, good, dug, bead, bed, bid, bird, head, heard, hid, had, hard, hood, hoard.¹

Each word list had 4 ‘dummy’ syllables at the beginning to give the speakers time to adjust to the task. There were also 5 ‘dummy’ items at the end of the list to allow for possible increase of speaking rate or decrease in volume or lowering of pitch that may possibly have occurred at the end of the reading. Both speakers were instructed to read out the words using their habitual reading voices, so as to avoid performance behaviour that may resulted in an unusual degree of variation in pitch, volume or speed of presentation. The coarticulation patterns of the twins’ voices reported here were based on the acoustic measurements of the five monosyllabic word lists.

2.3. Acoustic analysis

The 160 monosyllabic words (32 words x 5 repetitions) read by each twin were digitised onto a Kay Elemetrics Computerised Speech Laboratory (CSL, model 4300) with a sampling rate of 16kHz. Speech pressure waveforms and Fast Fourier Transform (FFT) spectrograms of the monosyllables were then generated and analysed using the CSL.

The following acoustic parameters were measured for the monosyllabic words. In addition, each parameter was examined by the place of articulation of the initial consonant of the words.

- Vowel duration (in milliseconds). The duration of the vowel was taken the entire duration of the vowel using the spectrograms. Vowel durations for 40 of the 320 CVC syllables (12.5%) were re-measured. The 95% confidence interval for the re-measurements was 4.7 to 16.7 msec, which compared favourably with the 95% confidence interval of 10 to 25 msec suggested previously [1].

- The first three formant frequencies (Hz). These measurements were taken at the onset and temporal midpoint for the vowel portion of each monosyllable. These measurements were obtained from the spectrograms using a hair crossed-line cursor, which provided an automatic readout at the intersection point of the cursor. Formant frequency values were measured at the mid frequency point of each formant frequency band.

- Locus equations (LEs) were generated using simple regression functions: $F_2\text{vowel} = k \cdot F_2\text{midvowel} + c$, where $k$ represents the slope of the function and $c$, the y-intercept. LEs were generated for each twin by place of articulation (bilabial, alveolar, velar¹, and glottal), and for all places of articulation combined.

- Timelag values (ms). These values were the temporal midpoint values, taken from the onset of the vowel to the mid point of the vowels, where the formant frequency values had been taken (see above description of how formant frequency parameters were taken).

- Formant frequency change (or excursion) values (in Hz). These values were calculated from the onset and mid values.

- Rate of formant frequency changes (Hz/ms). These were calculated as the rate of formant frequency change (Hz) over the timelag values (ms).

3. RESULTS

3.1. F2 Locus Equations

The scatterplots of F2 mid vowel values (Hz) plotted against F2 onset values for all places of articulation are given in Figures 1 and 2 for Twins B and D, respectively. The slope, y-intercept and R-squared values representing the locus equations for Twins B and D are given in Table 1 for all places of articulation, and by all places of articulation combined. The order of the steepness of the slope values was the same for both Twins B and D. This was as follows: Glottal > Bilabial > Velar > Alveolar. The y-intercept values for both Twins B and D showed the reverse pattern of magnitude differences by place of articulation, with the glottal and alveolar places of articulation having the lowest and highest y-intercept values, respectively. All slope and y-intercept values of Twins B and D showed overlap within the same 95% confidence intervals, therefore indicating that they were not significantly different from each other.

<table>
<thead>
<tr>
<th>Place of artic.</th>
<th>Parameter</th>
<th>Twin B</th>
<th>Twin D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bilabial</td>
<td>Slope</td>
<td>1.0</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>Y-intercept</td>
<td>-30.2</td>
<td>-53.6</td>
</tr>
<tr>
<td></td>
<td>R²</td>
<td>0.93</td>
<td>0.93</td>
</tr>
<tr>
<td>Alveolar</td>
<td>Slope</td>
<td>0.53</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td>Y-intercept</td>
<td>985.5</td>
<td>1031.2</td>
</tr>
<tr>
<td></td>
<td>R²</td>
<td>0.83</td>
<td>0.70</td>
</tr>
<tr>
<td>Velar</td>
<td>Slope</td>
<td>0.68</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>Y-intercept</td>
<td>799.1</td>
<td>398.2</td>
</tr>
<tr>
<td></td>
<td>R²</td>
<td>0.71</td>
<td>0.85</td>
</tr>
<tr>
<td>Glottal</td>
<td>Slope</td>
<td>1.21</td>
<td>1.15</td>
</tr>
<tr>
<td></td>
<td>Y-intercept</td>
<td>-380.9</td>
<td>-296.3</td>
</tr>
<tr>
<td></td>
<td>R²</td>
<td>0.92</td>
<td>0.94</td>
</tr>
<tr>
<td>All places</td>
<td>Slope</td>
<td>0.83</td>
<td>0.88</td>
</tr>
<tr>
<td></td>
<td>Y-intercept</td>
<td>370.9</td>
<td>273.1</td>
</tr>
<tr>
<td></td>
<td>R²</td>
<td>0.70</td>
<td>0.71</td>
</tr>
</tbody>
</table>

Table 1: Slope, y-intercept and R-squared values representing the Locus Equations for Twins B and D by place of articulation.

¹ The underlined words indicate those CVC words that contain rhotacised (‘r-coloured’) vowels in this Irish accent.
² Values for the front and back vowel contexts were combined for the velar place of articulation.
3.2. Formant frequency changes (or formant frequency excursions) and rates of change

The mean and standard deviation values for the F1-F3 changes (in Hz) and F1-F3 rates of change (Hz/msec) are given in Table 2 for Twins B and D by place of articulation. These data were subjected to a series of paired t-tests. The results of these are given in Table 3, and indicate the following significant inter-twin differences for F1-F3 changes: F1 change for both alveolar and velar places of articulation; F2 change for bilabial and velar places of articulation; and F3 change for bilabial, alveolar and glottal places of articulation (see Figure 3). In addition significant inter-twin differences were found for: F1 rate of change for the velar place of articulation; F2 rate of change for the bilabial place of articulation; and F3 rate of change for bilabial, alveolar and glottal places of articulation.
4. DISCUSSION

The coarticulation patterns reported here for a set of identical twins revealed a number of interesting results. Twins B and D had slope and y-intercept values similar to those described previously for bilabial, velar and alveolar plosives [9, 11]. For example, the steep slopes for the bilabial plosives were indicative of high levels of anticipatory coarticulation compared to the shallower slopes for the alveolar plosives (see Table 1). In addition, both Twins B and D had the steepest slope values for the glottal fricatives, a result, which suggests the highest levels of anticipatory coarticulation occurred for this phonetic context.

The results of the F2 locus equations showed that Twins B and D showed the same overall relational patterns between mid F2 values and F2 onset values for all places of articulation (see Table 1, and Figures 1 and 2). The overlap in both slope and y-intercept values for Twins B and D for all F2 locus equations determined by 95% confidence intervals illustrated the extent of this similarity. This result may be interpreted evidence for shared patterns of learning for motor speech behaviours necessary for the production of perceptually relevant acoustic targets, such as those expressed by F2 locus equations [9]. This supports results reported previously on the coarticulation patterns of three sets of monozygotic twins based on monosyllabic words with /l/ and /r/ in initial position [7].

The formant frequency changes and rates of formant frequency changes revealed a different pattern of results. For example, although there were no significant inter-twin differences for the F2 change parameter for the alveolar and glottal places of articulation, significant inter-twin differences were found for the bilabial and velar places of articulation (see Tables 2 and 3). In addition, there were no significant inter-twin differences for the F2 rate of change parameter for all, except the bilabial place of articulation. If we also look at both the F3 change and F3 rate of change parameters, we see evidence for greater inter-twin differences. Here, significant inter-twin differences were found for all, except the velar place of articulation (see Figure 3 for Mean F3 change parameter). These patterns suggest that there were both clear similarities and differences in both the static and dynamic F2 patterns of the twins’ renditions of monosyllabic words.

This investigation has highlighted some of the ways in which a set of identical (or monozygotic) twins shared similar speech patterns, and may explain why their voices were judged to be very similar in quality. The results have also suggested some of the ways in which the more dynamic aspects of motor speech control may be investigated to assess how far microphonetic aspects of habitual speech patterns are shared by monozygotic twins, dizygotic twins, siblings and non-siblings. Such studies would further inform the nature versus nurture debate.

5. ACKNOWLEDGMENTS

Our thanks go to our informants for their time and help with this study.

6. REFERENCES