ASYMMETRIES IN TONGUE-PALATE CONTACT DURING SPEECH

Naomi Rachel Miller[1], Constantino Carlos Reyes-Aldasoro[2], Jo Verhoeven[1,3]

[1]Division of Language and Communication Science, Phonetics Laboratory, City, University of London, UK
[2]Department of Electrical and Electronic Engineering, City, University of London, UK
[3]Computational Linguistics and Psycholinguistics Research Centre, University of Antwerp, Belgium
Naomi-Rachel.Miller@city.ac.uk, Carlos.Reyes-Aldasoro@city.ac.uk, Johan.Verhoeven.1@city.ac.uk

ABSTRACT

Research has shown that speech articulation tends to be asymmetrical in the transverse plane of the vocal tract. A recent meta-study of previously published electropalatograms revealed that 83% of these images show asymmetrical tongue-palate contact [1].

The present study investigated articulation asymmetry on the basis of a large number of electropalatograms acquired in a sentence-reading task at the Centre for Speech Technology Research, Edinburgh University (Mocha: Multichannel Articulatory Database). The vast majority (97.5%) of these palatograms showed some degree of left-right asymmetry, with greater contact on the left-hand side being the more common finding. Asymmetry was not strongly determined by voice or place of articulation. However, it was highly dependent on manner, with fricatives and the lateral approximant showing the greatest degree of asymmetry.

Characterisation of articulation asymmetry could improve our understanding of the speech-production process and its relationship with both neural organisation and the anatomy of the organs of speech.

Keywords: Electropalatography, articulation, asymmetry, Mocha-Timit.

1. INTRODUCTION

Electropalatography (EPG) is an instrumental technique for characterising tongue-palate contact during speech. A dentist produces a detailed plaster impression of the hard palate, upper teeth and gums. The cast is then used to manufacture a custom-made, thin acrylic palate that sits in the roof of the mouth and is held in place using wire clasps that clip over the teeth. A grid of (usually) 62 electrodes is distributed over the lower surface of the palate to capture the spatial pattern of tongue-palate contact. The electropalate is coupled to a receiver and the output, known as a *palatogram*, is a binary image representing the electrode activations. This image is refreshed at a typical frame rate of 100-200 Hz.

Visual inspection of published palatograms reveals that they frequently show lateral (left-right) asymmetry in the amount of tongue-palate contact [1]. However, articulation asymmetry in EPG has only been explicitly investigated in a small number of studies; see, for example, [2-5]. The main finding of these studies was that tongue-palate contact is often asymmetrical. However, the direction and typical extent of asymmetries in palatograms have not been studied in a systematic way. Articulation asymmetry may depend on a variety of factors, including the speaker's anatomy, their handedness, asymmetries in the manufacture of the electropalate, and the type of speech sound involved. Improved understanding of these factors would have both theoretical and practical benefits. It would contribute to a more profound understanding of motor control and motor constraints in speech production, and their potential relationship with the neural organisation of speech processing. In addition, it would provide important insights into the relationship between anatomical features of speakers and the acoustic characteristics of speech. From a practical viewpoint, standard values of asymmetry in neurotypical speakers could serve as a reference when treating speech deficiencies in which asymmetry leads to poor intelligibility of speech, e.g., dysarthria due to unilateral weakness.

The objective of the present study was to examine the effect of the speech sound on the direction and amount of asymmetry seen in electropalatograms. This is a logical starting point for a systematic EPG investigation of the factors that affect articulatory asymmetry, as it enables future experiments to focus on the most influential phonetic features. Articulation asymmetry was assessed using a set of palatograms acquired in a sentence-reading task at the Centre for Speech Technology Research (CSTR) at Edinburgh University [6]. This corpus consists of temporally-registered sound and EPG files, as well as data from other instrumental methods such as laryngography. The database was chosen due to the large number of sentences per speaker (460), making it an ideal resource for studying the direction and amount of asymmetry as a function of the phoneme. A further objective was to characterise within-speaker variability in asymmetry across different realisations of the same phoneme. For a sentence-reading task, the main cause of this variability is likely to be the phonetic context.

2. METHODS

The asymmetry data were obtained by analysing palatograms from the Mocha (Multichannel Articulatory Database) – Timit (M-T) corpus provided by the CSTR [6]. This corpus consists of a phonetically balanced set of 460 sentences (example: *Those thieves stole thirty jewels*) designed to include the main connected-speech processes of English such as assimilations and weak forms. Palatograms are stored in raw binary form (8 bytes per sample) at a frame rate of 200 Hz. The EPG data are reported to be "carefully synchronised" with the audio data [6], where the latter are sampled at a rate of 16 kHz. For each sentence, the authors supply a text file (with the extension '.lab') containing phoneme segmentations performed using forced alignment [Simon King, personal communication]. Each row of the .lab file contains the identity of the phoneme, coded using the CSTR Machine Readable Phonetic Alphabet, and its start and finish times. Complete datasets (.wav, .epg and .lab files) appear to be available for five speakers of English (3 females and 2 males) with a variety of accents. All five speakers were included in the present study. In total, palatograms pertaining to 34,370 tokens (phonemes) were analysed.

The data were read into MATLAB [7] using the suite of open-source programmes presented in [8]. The asymmetry analysis focused solely on the consonant phonemes, as EPG provides limited information about vowels, except for the lateral tongue-palate contact in high front vowels. Furthermore, it was anticipated that the *manner* of articulation (which only varies in consonants) would be the most influential phonetic dimension.

For each palatogram, three indices of asymmetry were calculated. The first reflects the left-right asymmetry for the entire palate [4]:

$$I_{as} = (N_R - N_L) / (N_R + N_L) \qquad (1)$$

where N_R is the number of activated electrodes on the right-hand side and N_L is the number of activations on the left. Thus, negative values indicate more tongue-palate contact on the left, while positive values denote a right-sided bias. The degree of asymmetry is given by the magnitude of the index, $|I_{as}|$, which ranges from 0 to 1, with higher values representing greater asymmetry. The remaining two indices were based on a formula analogous to Eq. (1), but with the left-right asymmetry calculated separately for the anterior and posterior halves of the palatogram. For each instance of each phoneme (i.e., 'token'), the asymmetry metrics were calculated from a cumulated image that was obtained by summing the palatograms over the *entire duration* of the token.

3. RESULTS

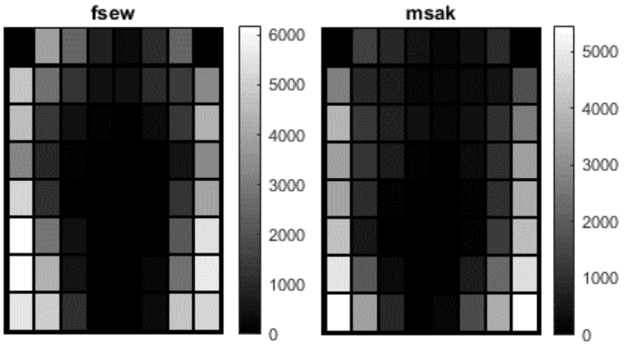
3.1. Overall direction and degree of asymmetry

On average, 97.5% of palatograms showed some degree of asymmetry (i.e., a non-zero value of I_{as}), with a range of 95.7% to 98.2% across the five speakers. The summary statistics relating to I_{as} are shown in Table 1. A one-way analysis of variance with 'speaker' as the independent variable and I_{as} as the dependent variable was highly significant [F(4, 34365) = 461.0, $p < 0.0001$]. Table 1 shows that, on average, all speakers exhibited greater tongue-palate contact on the left-hand side (i.e., *negative* mean values of I_{as}). For four out of five speakers, the mean values of the asymmetry metric for the anterior and posterior halves of the palate were both negative. For the remaining speaker ('msak'), the front and back of the palate showed opposite directions of asymmetry. This can be appreciated by inspection of Fig. 1, which presents cumulative palatograms for the speakers 'fsew' and 'msak' (chosen for display because their hand dominance was known: both right-handed). The male speaker ('msak') shows a left-sided bias at the front of the palate (the upper part of the image) and a right-sided bias at the back of the palate. This largely explains why his mean asymmetry index was the closest to zero of all the speakers.

Table 1: Summary statistics for the five speakers

Speaker	# phonemes	Mean I_{as}	St dev I_{as}
ffes	7247	-0.131	0.138
fjmw	6715	-0.078	0.111
fsew	6617	-0.104	0.096
mjjn	7201	-0.084	0.143
msak	6590	-0.037	0.167

Figure 1: Cumulative palatograms for two speakers, obtained by summing over all phonemes. The grey level indicates the cumulative number of contacts.



3.2. Degree of asymmetry as a function of the phoneme

The relationship between the amount of asymmetry and the identity of the phoneme was explored using correspondence analysis (CA), a technique for condensing the information in a two-way contingency table. In the present study, the row variable is the phoneme, the column variable is the degree of asymmetry, and each datapoint is the number of tokens of the phoneme with the given asymmetry level (summed over all speakers). Since CA requires categorical data, four levels of asymmetry were defined: $|I_{as}| \leq 0.1$ ('small'), $0.1 < |I_{as}| \leq 0.2$ ('medium'), $0.2 < |I_{as}| \leq 0.3$ ('large'), and $|I_{as}| > 0.3$ ('very large'). CA calculates a set of coordinates representing the associations between the row and column variables. The results are plotted on a map (Fig. 2), where the two axes are conceptually similar to the first and second component in principal component analysis. In this case, 98.7% of the total Pearson Chi-square for the two-way frequency table is accounted for by the first two dimensions. The greater the horizontal (vertical) distance between the origin and a variable along Dimension 1 (Dimension 2), the greater the contribution of that variable to that dimension.

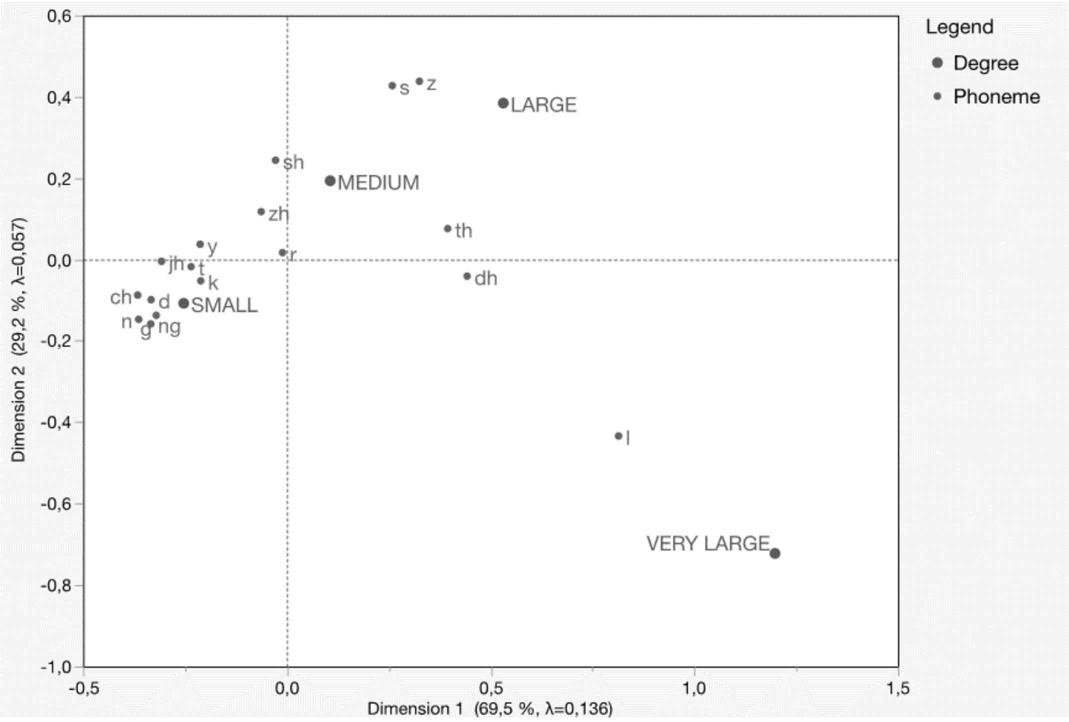
The interpretation of Fig. 2 is not straightforward and for a detailed description, the reader is referred to [9, 10]. However, the two most important principles are as follows: (1) If two phonemes are in close proximity to each other, then they have a similar relative frequency distribution (i.e., 'profile') across the four asymmetry levels. (2) A small (<< 90°) angle between the diagonal line from the origin to a given *phoneme* and the diagonal line from the origin to a given *asymmetry level* indicates that these two variables are positively associated. The strength of this association increases with the lengths of the diagonal lines. A 90° angle indicates no relationship, while ~180° implies a strong negative association.

As far as *voice* is concerned, it can be seen that the voiceless plosives /t, k/ show more asymmetry than their voiced counterparts. The opposite trend is seen for fricatives and affricates (i.e., voiced sounds exhibit more asymmetry), with the exception of the post-alveolar fricatives. However, in general, the distances between voiced and voiceless phonemes are small. Therefore, there is not a strong association between voice and level of asymmetry.

A similar picture emerges for *place* of articulation. In plosives and nasals, there is no strong association between place and degree of asymmetry. Likewise, in fricatives, the alveolars and the dentals show similar asymmetry profiles. The only notable trends are that (1) post-alveolar fricatives show less asymmetry than fricatives produced at more anterior positions and (2) the alveolar approximant ('r') exhibits more asymmetry than its palatal counterpart ('y'). Thus, overall, there is some evidence to suggest that anterior places of articulation are more prone to asymmetry.

The strongest association is seen for the dimension *manner*. Plosives, nasals and affricates all exhibit relatively low levels of asymmetry, while fricatives and /l/ show substantially greater asymmetry. Furthermore, the lateral approximant stands in isolation, implying that its asymmetry profile does not resemble that of any other phoneme.

Figure 2: Correspondence analysis between the phoneme and the degree of asymmetry. The following symbols denote voiced-voiceless phoneme pairs: dh, th - dental fricatives; zh, sh - post-alveolar fricatives; jh, ch - post-alveolar affricates.

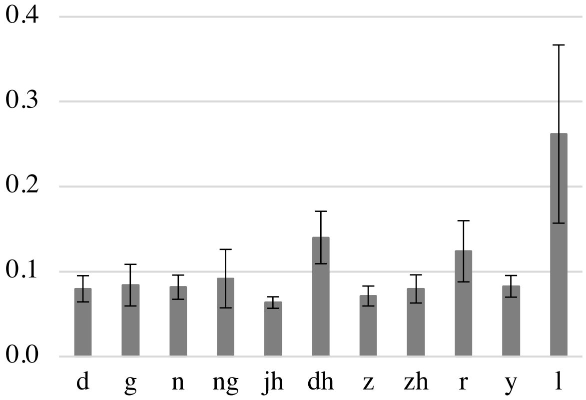


3.3. Intra-speaker variability

The final objective was to examine within-speaker variability in asymmetry values as a function of the phoneme. This represents variation in realisation of the phoneme due to the phonetic context. This information could be of theoretical interest; for example, it is likely that phonemes that are more susceptible to coarticulatory effects show greater asymmetry variability. Furthermore, variability data could be useful for planning future experiments, e.g., by allowing estimation of the sample size required to achieve a given confidence interval on I_{as}.

Intra-speaker variability was denoted by $stdev(I_{as})$, the standard deviation on the mean value of I_{as} for a given speaker and phoneme. The average value across the 5 speakers (± 1 SD) was then calculated (see Fig. 3). Voiceless phonemes are not shown, as the variability values closely matched those of their voiced counterparts. Intra-speaker variability is similar across phonemes, except for the dental fricative ('dh'), alveolar approximant and lateral approximant, all of which show higher variability.

Figure 3: Intra-speaker variability per phoneme, calculated as the mean value (± 1 SD) of $stdev(I_{as})$.

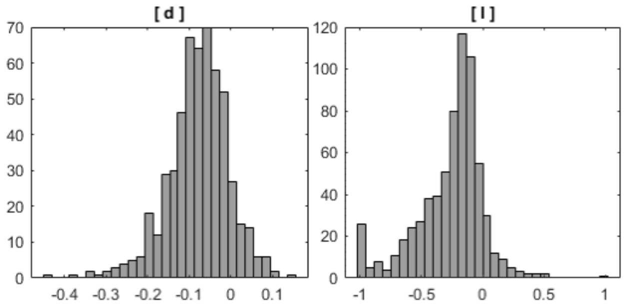


Finally, Fig. 4 shows histograms for I_{as} for the phonemes /d/ and /l/ for speaker 'ffes', chosen because her standard deviation values were similar to the averages for the cohort. It can be seen that the distribution for /l/, in particular, is non-normal, with significant skewness and outlying values.

4. CONCLUSIONS AND DISCUSSION

The vast majority (97.5%) of palatograms showed some degree of asymmetry, the predominant direction of which was towards the *left*. These findings agree with a meta-study that extracted values of I_{as} from published palatograms [1]. However, this metric is highly sensitive (a left-right difference of just one electrode is considered 'asymmetrical'). Thus, future work will determine the confidence interval associated with measures of I_{as}, due to, for example, variability in electrode sensitivity. Although the hand dominance of all five speakers in the M-T database is unknown, at least two are right-handed. Therefore, in common with [3], the present study suggests that asymmetry does not align with hand dominance.

Figure 4: Example histograms of the asymmetry index I_{as}.



The degree of asymmetry was shown to have only a weak association with voicing status and place of articulation, but a strong association with manner. In particular, fricatives and the lateral approximant exhibited greater asymmetry than other manners of articulation. In the case of fricatives, asymmetrical articulation may be a means of maximising turbulence. This would imply that it is a learnt process. The fact that lateral consonants may exhibit marked articulation asymmetry has been discussed previously [2, 3]. Figure 4 shows that tongue-palate contact during the articulation of /l/ may be unilateral.

The dental fricatives, as well as /r/ and /l/, were found to show greater variability in asymmetry than other phonemes. Approximants are known to have variable articulation, including a propensity to be influenced by adjacent sounds, as they are not characterised by precise articulatory demands [11]. The cause of high variability in asymmetry for dental fricatives warrants further investigation.

A limitation of this study was that asymmetry metrics were derived based on the entire duration of the phoneme. Thus, these metrics were strongly influenced by coarticulation. Future work should also determine asymmetry from the stable portion of the phoneme. This would minimise coarticulation and hence indicate the degree to which asymmetry is primarily a feature of the target phoneme. A further limitation lies in the fact that neither the speakers' electropalates, nor their plaster casts, were available for analysis. In future work, these devices will be scanned using computed tomography, and analysed so as to determine whether the asymmetry seen in palatograms can be attributed to anatomical factors and/or asymmetries in the palate manufacture [8].

5. ACKNOWLEDGEMENTS

This work was funded by The Leverhulme Trust.

6. REFERENCES

[1] Verhoeven, J., Mariën, P., De Clerck, I., Daems, L., Reyes-Aldasoro, C., Miller, N. 2019. Asymmetries in speech articulation as reflected on palatograms: A meta-study. *Proc. 19th ICPhS* Melbourne, Australia.

[2] Hamlet, S., Bunnell, H., Struntz, B. 1986. Articulatory asymmetries. *Journal of the Acoustical Society of America* 79, 1165–1169.

[3] Hamlet, S. 1987. Handedness and articulatory asymmetries on /s/ and /l/. *Journal of Phonetics* 15, 191–195.

[4] Marchal, A., Farnetani, E., Hardcastle, J., Butcher, A. 1988. Cross-language EPG data on lingual asymmetry. *Journal of the Acoustical Society of America* Suppl. 1, vol. 84, S127.

[5] Stone, M., Faber, A., Raphael, L., Shawker, T. 1992. Cross-sectional tongue shape and linguopalatal contact patterns in [s], [ʃ] and [l]. *Journal of Phonetics* 20, 253–270.

[6] Wrench, A. Mocha-Timit Multichannel Articulatory Database: English. http://www.cstr.ed.ac.uk/research/projects/artic/mocha.html

[7] MATLAB 8.0 and Statistics Toolbox 8.1, The MathWorks, Inc., Natick, Massachusetts, United States.

[8] Verhoeven, J., Miller, N. R., Daems, L., Reyes-Aldasoro, C. C. 2019. Visualisation and analysis of speech production with electropalatography. *Journal of Imaging* 5(3), 40.

[9] How Correspondence Analysis Works (A Simple Explanation). https://www.displayr.com/how-correspondence-analysis-works/

[10] Higgs, N. T. 1991. Practical and innovative uses of correspondence analysis. *The Statistician* 40, 183–194.

[11] Matthies, M. L., Guenther, F. H., Denny, M., Perkell, J. S., Burton, E., Vick, J., Lane, H., Tiede, M., Zandipour, M. 2008. Perception and production of /r/ allophones improve with hearing from a cochlear implant. *Journal of the Acoustical Society of America* 124, 3191–3202.