Diachronic phonological asymmetries and the variable stability of synchronic contrast

Sam Kirkham and Claire Nance

- Secondary palatalisation contrasts vary in diachronic stability across sonorants
- We investigate the synchronic stability of secondary palatalisation contrasts in Scottish Gaelic
- Support Vector Machines classify compact representations of acoustic and articulatory data
- Rhotics are best classified word-initially and worst word-finally
- Dynamic information is crucial to phonological contrast, but varies in magnitude across sonorants
- Variable synchronic stability of contrasts complicates potential trajectories of diachronic change

Preprint submitted to Journal of Phonetics *July 8, 2022*

Diachronic phonological asymmetries and the variable stability of synchronic contrast

Sam Kirkham and Claire Nance

Abstract

This article aims to understand the development of diachronic asymmetries in phonological systems by evaluating the variability stability of synchronic contrasts. We focus on sonorant systems involving secondary palatalisation, grounded in the claim that palatalised laterals are more common than palatalised rhotics cross-linguistically. Our analysis reports acoustic and articulatory data on Scottish Gaelic, a Celtic language with a large sonorant inventory contrasting palatalised, plain and velarised phonemes across laterals, nasals and rhotics. We summarise high-dimensional dynamic characteristics of the acoustic spectrum and midsagittal tongue shape using a two-stage data reduction process and use these coefficients as inputs for training a Support Vector Machine. This trained model classifies unseen data in terms of its phonemic identity, which reveals that rhotics are classified best word-initially and worst word-finally, with nasals always classified better than laterals. We find that dynamic information substantially improves acoustic classification, but only improves articulatory classification for some sonorants. We propose that the variable synchronic stability of palatalisation contrasts complicates potential trajectories of diachronic change in Gaelic.

Keywords: Sound change, palatalisation, Scottish Gaelic, sonorants, synchronic variation, diachronic change, ultrasound

1. Introduction

In this article, we investigate whether diachronic and typological asymmetries in phonological systems are reflected in the variable stability of synchronic contrasts. It is widely predicted that the diachronic instability of some phonological contrasts is a consequence of a larger pool of synchronic variability. This is because such variability is hypothesised to facilitate misperception-based sound change (Ohala, 1981) and can also weaken the robustness of phonemic categories, leading to potential neutralisation over time (Bybee, 2015). But does the propensity of a phonological contrast towards diachronic neutralisation necessarily mean that it will be less robust *at a given point in time*? An assumption underpinning many theories of sound change is that we can observe the tendencies of diachronic change through examination of synchronic data, with the hypothesis that there is a tight link between the two at any point in time (Labov, 1994, 21). This suggests that a greater tendency towards diachronic neutralisation should also be evident in synchronic data. In this study, we examine claims about the diachronic trajectories of typologically unusual sound systems and whether the variable stability of synchronic contrasts is predictable from the attested sound changes. We also speculate on whether variable synchronic stability between phonological categories might be able to tell us something about future trajectories of sound change, especially in light of existing diachronic predictions.

A particularly good case study for examining variable diachronic and synchronic stability is the cross-linguistic system of contrasts that fall under the banner of secondary palatalisation. Previous research shows that some secondary palatalisation contrasts in consonants are more unstable than others (Kochetov, 2005; Iskarous and Kavitskaya, 2018). Palatalised rhotics, in particular, are cross-linguistically rare and prone to merger with non-palatalised rhotics (Hall, 2000), but laterals seem more robust to sound change (Iskarous and Kavitskaya, 2010). Word-final palatalisation contrasts are also more unstable than word-initial contrasts (Padgett and Ní Chiosáin, 2018). Importantly, previous work shows that the robustness of palatalisation contrasts may vary depending on the features analysed; for example, nasals may be more distinctive than laterals in format transitions, but laterals have a more distinctive spectral shape than nasals, with rhotics being least distinct in both analyses (Iskarous and Kavitskaya, 2018). This suggests that palatalisation contrasts are multi-dimensional and temporally distributed, potentially as a consequence of a high number of phonological categories existing together in a relatively narrow phonetic space.

The fact that some sonorant contrasts are diachronically less stable than others cross-linguistically makes them an ideal candidate for assessing claims about variable diachronic trajectories using synchronic data. In this study we wish to further understand why some sonorants show greater stability than others and, in doing so, we focus on palatalisation contrasts in Scottish Gaelic (Celtic), which contrasts palatalised, velarised and plain sonorants across laterals, rhotics and nasals. Notably, Scottish Gaelic has retained a larger system of sonorants in comparison to closely-related Irish and Manx. In this study, we take seriously the dynamic nature of sonorant contrasts, building upon our previous work that has focused on selective sampling of a limited number of timepoints. We show that this previous research may underestimate the extent of contrast that is present in the Scottish Gaelic sonorant system; contrasts which we argue are fundamentally dynamic in nature. We further demonstrate this by comparison with analyses that focus only on a sonorant 'steady-state', which illustrates how some contrasts may be more dynamic in nature than others.

1.1. Dynamics of secondary palatalisation

Secondary palatalisation involves overlap between a palatal gesture and the consonant's primary place of articulation, which contrasts with 'full palatalisation', where the consonant's primary place of articulation is changed (Bateman, 2007, 2). Some languages, such as Russian and Scottish Gaelic, have extensive secondary palatalisation contrasts across the consonant system, such that almost every consonant has a palatalised and non-palatalised counterpart (see Yanushevskaya and Bunčić 2015 for description of Russian, and Nance and Ó Maolalaigh 2021 for description of Scottish Gaelic). For this reason, all consonant palatalisation pairs in Russian, Scottish Gaelic and other languages with this system are considered to contrast in secondary palatalisation even though the secondary palatalisation contrast may at times manifest as a change in primary place/manner.[1]

In terms of articulation, the most widely reported articulatory correlate of secondarily palatalised consonants is tongue body fronting and raising towards the palate accompanying the primary consonantal gesture (Kochetov,

[1] For example in *maide* 'stick' /matʃə/, where the orthographic 'd' is palatalised and changes from alveolar to post-alveolar place of articulation.

2002; Stoll, 2017; Bennett et al., 2018; Malmi and Lippus, 2019; Spinu et al., 2019). The fronting and raising gesture also frequently extends into the surrounding vowels (Malmi and Lippus, 2019). This tongue body fronting is often accompanied by tongue root advancement and pharyngeal expansion (Kavitskaya et al., 2009; Bennett et al., 2018), while palatographic studies additionally demonstrate that the tongue blade is spread across the hard palate to a greater extent than in non-palatalised consonants (Farnetani et al., 1991; Meister and Werner, 2015).

Capturing the acoustics of palatalisation contrasts is complex given their multi-dimensional and dynamic nature. When the tongue body is raised and fronted for a palatalised consonant, this results in a larger back cavity and raised F2, which is particularly robust in laterals (Sproat and Fujimura, 1993; Nance, 2014; Kochetov et al., 2020). The vowels surrounding palatalised consonants also tend to show raised F2 due to an /i/-like glide in the transition to/from a palatalised consonant, with such articulatory dynamics being important to the contrast (Ní Chiosáin and Padgett, 2012; Kochetov, 2017; Nance and Kirkham, 2020; Howson, 2018; Malmi et al., 2022). F1 and F3 may also be lower in palatalisation contexts (Shuken, 1980; Bennett et al., 2018; Kochetov, 2017). The multi-dimensional nature of palatalisation contrasts have led others to analyse more holistic spectral features, such as Mel Frequency Cepstral Coefficients (MFCCs) (Spinu et al., 2012; Spinu and Lilley, 2016; Spinu et al., 2018) and smoothed spectra (Kochetov, 2017; Iskarous and Kavitskaya, 2018; Nance and Kirkham, 2020). For example, cepstral coefficients have been found to significantly outperform spectral measures in classifying palatalised fricative contrasts (Spinu and Lilley, 2016;

Spinu et al., 2018).

Secondary palatalisation is a good case study for testing the relationship between diachronic neutralisation and synchronic stability, because of well-documented differences between sonorant types. Palatalised rhotics involve a retracted and stabilised tongue body for trill production (McGowan, 1992; Recasens, 2013), which comes into conflict with the tongue body advancement needed for palatalisation (Iskarous and Kavitskaya, 2018; Kochetov, 2005; Stoll, 2017). Such biomechanical constraints may lead to a larger pool of synchronic variability (Ohala, 1989), with the possibility that variants become phonologised or contrasts are neutralised over time (Beckman et al., 1992; Bybee, 2015). For example, articulatory variability may lead to ambiguity in perception, which could advance the spread of a change further when misperceived by the listener (Ohala, 1981). Such explanations are explicitly pursued in previous research on sonorant palatalisation in terms of acoustics (Iskarous and Kavitskaya, 2018) and articulation (Kochetov, 2005; Stoll, 2017), with the claim in both cases being that less robust phonemic categories are more susceptible to merger.

1.2. Palatalisation in Gaelic

Our study focuses on Scottish Gaelic, a Celtic language closely related to Irish and Manx.[2] The Scottish Gaelic language is usually referred to in English by its speakers simply as ‘Gaelic’ /galɪk/ and we refer to it as

[2]Manx is believed to have become extinct as a first language in the 1970s, following a long period of decline, but has subsequently undergone revival. It is taught in immersion schooling and is transmitted in a small number of families. See Lewin (2021) for more information on revived Manx.

Gaelic henceforth. Together, the Celtic language sub-family consisting of Gaelic, Irish and Manx is known as 'Goidelic'. The most recently available data (Scottish Government, 2015) show that there are approximately 57,600 Gaelic speakers in Scotland. Traditionally, Gaelic is associated with the north-west Highlands and Islands of Scotland, and this is where the most densely concentrated populations of Gaelic speakers live. In particular, Gaelic is associated with the chain of islands off the north-west coast of Scotland known as the Outer Hebrides or Western Isles, where around 60% of the population reported the ability to speak Gaelic (Scottish Government, 2015). A map showing the concentration of Gaelic speakers in Scotland is in Figure 1. The speakers in this study are from the Isle of Lewis, the most northerly island in the Outer Hebrides chain. The Goidelic languages are descended from Old Irish, which expanded from Ireland to Scotland and Isle of Man in early medieval times (McLeod, 2020). It is generally thought that Gaelic in Scotland had sufficiently diverged from Irish to be considered a separate language in approximately 1100 CE (Ó Maolalaigh, 2008).

The Goidelic languages all have systems of contrastive secondary palatalisation across the entire consonant system (with a few exceptions in some consonants) (Broderick, 2009; Hickey, 2014; Bennett et al., 2018; Nance and Ó Maolalaigh, 2021). In Nance and Kirkham (forthcoming 2022), we provide a historical overview of the development of palatalisation in rhotics and comparison to different Goidelic dialects. In this paper, we focus on the contrasts across the whole sonorant system. To summarise: the most extensive Goidelic palatalisation contrasts were found in Old Irish, where the system developed by approximately 900 CE (Greene, 1973; Hickey, 1995). At this time, Old Irish

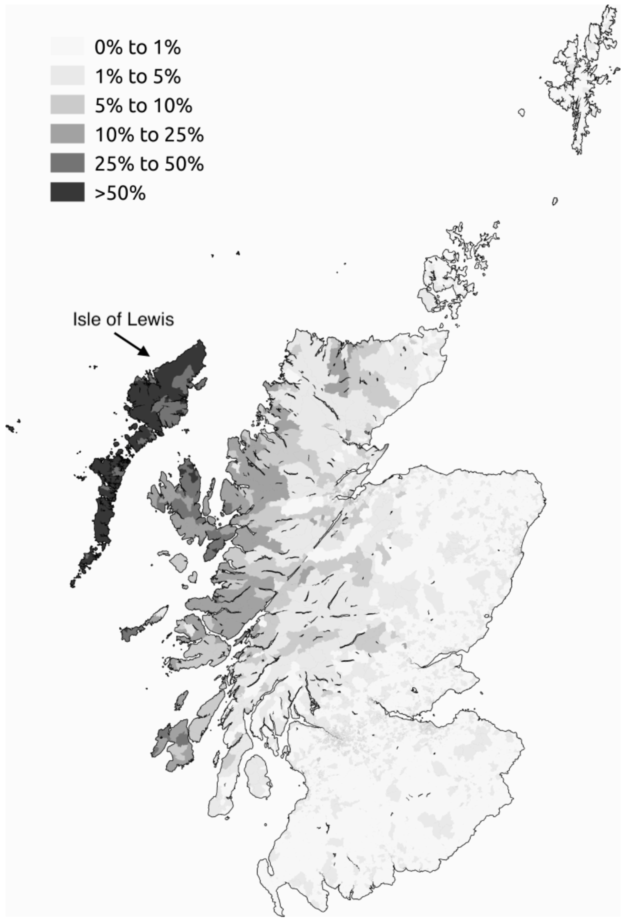

Figure 1: Map showing the concentration of Gaelic speakers in Scotland according to the most recently available figures from the 2011 National Census. Attribution: By SkateTier - Own work, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=31996352. Original figure in colour, converted to greyscale here.

sonorants contrasted in place of articulation as well as palatalisation, resulting in four different phonemes for laterals, nasals and rhotics (Thurneysen, 1946; Russell, 1995; Hickey, 1995). It is thought that a three-way contrast between palatalised, plain and velarised sonorants developed in Middle Irish (900–1200

CE) (Hickey, 1995). The Irish system has evolved since early medieval times in different ways in the modern Goidelic dialects. The most innovative dialect in this respect is Manx, where palatalisation contrasts were lost in rhotics, and reduced in laterals and nasals. At the other end of the scale are Hebridean dialects of Gaelic, including the dialect under investigation here, Lewis Gaelic. In Lewis and other Hebridean dialects, three lateral, three nasal and three rhotic phonemes are maintained.

In comparison to many of the previous studies of palatalisation, Lewis Gaelic is interesting in several respects. The majority of work carried out previously on palatalisation has examined contexts where palatalised consonants are contrasted with non-palatalised consonants, such as Russian. In Gaelic sonorants there is instead a three-way distinction between palatalised, plain and velarised. The rhotic inventory, however, has been particularly prone to reduction across Goidelic dialects, with laterals appearing most robust to sound change. This is in line with the findings discussed above for Slavic, which show that rhotics are more susceptible to change than laterals (Carlton, 1990; Iskarous and Kavitskaya, 2018).

1.3. Summary and predictions

In the current study, we investigate the extent to which palatalisation contrasts are maintained, combining dynamic phonetic evidence from acoustics and articulation in order to examine whether phonemic distinctiveness varies between laterals, nasals and rhotics. We specifically build upon previous work in the following ways. First, previous work on the asymmetry of sonorant palatalisation contrasts has focused on Russian as the language with the most extensive system of sonorant palatalisation in the Slavic family (Kochetov, 2005; Stoll, 2017; Iskarous and Kavitskaya, 2018). Here, we consider Lewis Gaelic, as the Goidelic dialect with the most extensive system of sonorant palatalisation in a completely different language family. Second, previous work in this area has considered articulation (Kochetov, 2005; Stoll, 2017) or acoustics (Iskarous and Kavitskaya, 2018) respectively, but we combine both perspectives and use a method that allows us to subject each modality to a comparable classification task. Third, much previous work has focused on static timepoints, either sonorant midpoints or specific locations of formant transitions. We take a broader approach by compressing all time-varying information that is available in the signal and using this to assess classification accuracy. This allows us to more comprehensively investigate the hypothesis that diachronically unstable contrasts are more vulnerable to synchronic neutralisation at a specific snapshot in time. Accordingly, we set out the following questions for the present study:

1. Which sonorant categories (laterals, nasals, rhotics) show the most robust phonemic contrasts?
2. Is contrast more robust in acoustic or articulatory data?
3. How do acoustic and articulatory dynamics contribute to phonological contrast?
4. What do these results tell us about the variable synchronic stability of categories and the potential diachrony of palatalisation contrasts?

We test the prediction that laterals will be best classified, followed by nasals and then rhotics, and anticipate that reduction will be more evident word-finally. In previous work on Gaelic, Nance and Kirkham (2020) show that laterals are more robust than nasals in formants at the sonorant steady-state, while Nance and Kirkham (forthcoming 2022) show that three initial rhotics are well-maintained in Gaelic, despite potential neutralisation of rhotics in word-final position. However, these studies used different methods and different features to establish contrast, as well as focusing on a small set of selective timepoints, so our present study uses a more holistic and comparable method for establishing the relative robustness of three-way contrasts across laterals, nasals and rhotics.

2. Methods

2.1. Speakers

We recorded data from twelve L1 speakers of Lewis Gaelic, all of whom were raised in Gaelic-speaking families on the Isle of Lewis (six female, six male). They acquired English either as simultaneous bilinguals or upon entering the school system. The speakers were aged 21-80 and either used Gaelic as part of their job, or had used Gaelic before retirement. All the speakers reported using more Gaelic than English in their daily lives and can be considered Gaelic-dominant bilinguals. Due to the fragility of Gaelic language transmission, even in locations such as Lewis (Munro et al., 2011), it is difficult to obtain a large sample of data from Gaelic-dominant bilingual speakers. We recognise that the data here represent a large age range, but the speakers are socially consistent in using more Gaelic than English.

2.2. Data recording and stimuli

Simultaneous acoustic and ultrasound tongue imaging data were recorded in a community centre or at the speaker's workplace. The acoustic signal was recorded using a Beyerdynamic Opus 55 headset microphone, which was preamplified and digitized using a Sound Devices USBPre2 audio interface at 44.1 kHz with 16-bit quantization. Simultaneous ultrasound data were recorded using a Telemed MicrUs system, with a 64 element probe of 20 mm radius. We used a 2 MHz probe frequency, 80 mm depth, 90% field of view and 57 scan lines, which resulted in a frame rate of ~92 Hz. The probe was stabilised using an Articulate Instruments metal headset (Articulate Instruments, 2008). The occlusal plane for each speaker was imaged by them biting on a bite plate placed behind the upper incisors and pushing their tongue up against it. Synchronization between audio and ultrasound data was achieved using the frame-level TTL pulse emitted by the ultrasound scanner. Data presentation and recording was handled using the Articulate Assistant Advanced software (Articulate Instruments, 2018).

The stimuli used for this study are shown in the Appendix (Tables 8–10). We aimed to capture laterals, nasals and rhotics in word-initial and word-final position in three vowel contexts where possible: /i a u/. This was not always possible due to the historical development of palatalisation in high front vowels. For example, there are no velarised nasals in the context of /i/ in readily-known words. The plain sonorants developed from contexts of historical lenition, and in word-initial position they still occur in contemporary lenition contexts. For an overview of changes in lenition (contemporary morphophonological changes in Celtic language word-initial consonants known as 'mutation'), see Ball and Müller (2009) or Nance and Ó Maolalaigh (2021) for Gaelic specifically. For this reason we included the word-initial plain sonorants in short phrases that would trigger mutation – e.g. *mo nathair* 'my snake' – where the possessive *mo* ‘my’ triggers mutation.

2.3. Data preparation

Acoustic landmarks were labelled manually in Praat using information from the waveform and spectrogram (Boersma and Weenink, 2020). We labelled the entire sonorant-vowel interval for all tokens, such as lateral-vowel for word-initial tokens and vowel-lateral for word final tokens. This interval was used for all analyses reported in this paper. We carried out post-hoc screening of the ultrasound data and found that only seven of the twelve speakers had consistently good images (three female, four male). As our analysis below is premised upon comparing acoustic and articulatory data, we only use these seven speakers for the analysis, resulting in 1165 tokens with parallel acoustic and ultrasound data.

2.4. Acoustic features

The acoustic features used in this analysis are Mel Frequency Cepstral Coefficients, which are highly effective at reducing the dimensionality of the spectrum while retaining linguistically-relevant features (Davis and Mermelstein, 1980). MFCCs are directly related to characteristics of the spectrum and, therefore, do have a physical interpretation, despite their complexity in the higher coefficients. For example, lower MFCCs describe global aspects of spectral shape, while increasingly higher coefficients describe increasingly finer details in the spectrum. MFCCs have previously been shown to capture phonemic palatalisation contrasts with a high degree of accuracy (Spinu et al., 2012; Spinu and Lilley, 2016; Spinu et al., 2018).

We use 6 MFCCs to summarise the acoustic spectrum, which has previously been shown to be sufficient for capturing palatalisation contrasts (Spinu et al., 2018). We sensitivity tested the effects of between 4 and 13 MFCCs and found that 6 MFCCs resulted in the strongest overall classification accuracy, although some specific models showed a small (2-4%) improvement using 8 coefficients, after which no further improvement was evident. Accordingly, for each token, 6-element MFCC vectors were calculated across each sound file using a 25 ms window and 10 ms frame shift, with a pre-emphasis coefficient $\alpha = 0.97$ and a lifter exponent of 0.6. MFCCs were subsequently extracted at 11 equally spaced points across the labelled sonorant-vowel interval and each MFCC was by-speaker normalized using z-scoring. At this stage, each token is represented by 6 MFCC trajectories, each of which is sampled over 11 points.

2.5. Articulatory features

Splines were automatically fitted to the midsagittal ultrasound data using AAA's batch fit function. A paid research assistant manually checked and corrected any obvious errors in the splines, but we did not correct minor tracking errors. All splines were then rotated and scaled to the occlusal plane. These data comprise 42 values in 2-dimensional x/y space. In order to reduce the dimensionality of the tongue splines, we fitted a Discrete Cosine Transform (DCT) to each token at 11 propotionally-spaced timepoints across the sonorant-vowel or vowel-sonorant interval. The DCT has been used for summarising whole acoustic spectra (Harrington, 2010; Nossair and Zahorian, 1991), formant trajectories (Watson and Harrington, 1999) and articulatory time series (Shaw and Kawahara, 2018) and is conceptually extendable to spatial representations, such as the ultrasound tongue spline. To this end, the ultrasound-DCT is conceptually comparable with MFCCs, as both sets of features fundamentally represent the amplitudes of cosine waves fitted to the respective signals after undergoing transformation. The DCT coefficients have a physical interpretation, with the lower coefficients being proportional to the mean (C_0), slope (C_1) and curvature (C_2) of the tongue shape, with higher coefficients representing increasingly finer detail in the shape. We fit a DCT of the form described in Harrington (2010) with m coefficients to a signal $x(n)$ with length N, where the m^{th} coefficient C_m is calculated using (1).

$$\begin{gathered} C_m = \frac{2k_m}{N} \sum_{n=0}^{N-1} x(n) \cos\left(\frac{(2n+1)m\pi}{2N}\right) \\ \text{where } k_m = \frac{1}{\sqrt{2}}, m = 0; k_m = 1, m \neq 0 \end{gathered} \tag{1}$$

We illustrate DCT compression of ultrasound tongue shapes in Figure 2, which represent smoothing using different numbers of DCT coefficients (between 2 and 10 coefficients) on a single token. We obtained the smoothed tongue shapes using an inverse DCT, which reconstructs the input signal by summing half-cycle cosine waves with the amplitudes of the corresponding DCT coefficients. The figure shows us that two coefficients $\{C_0, C_1\}$ approximates the slope of the tongue, while using between three $\{C_0, C_1, C_2\}$ and five $\{C_0, C_1, ..., C_4\}$ produces similar tongue shapes. At 6 DCT coefficients $\{C_0, C_1, ..., C_5\}$ the slight dip between the tongue tip and dorsum starts to appear, which is present in the original signal. After this, we see an increasing level of detail, but not necessarily any strikingly new information in the signal.

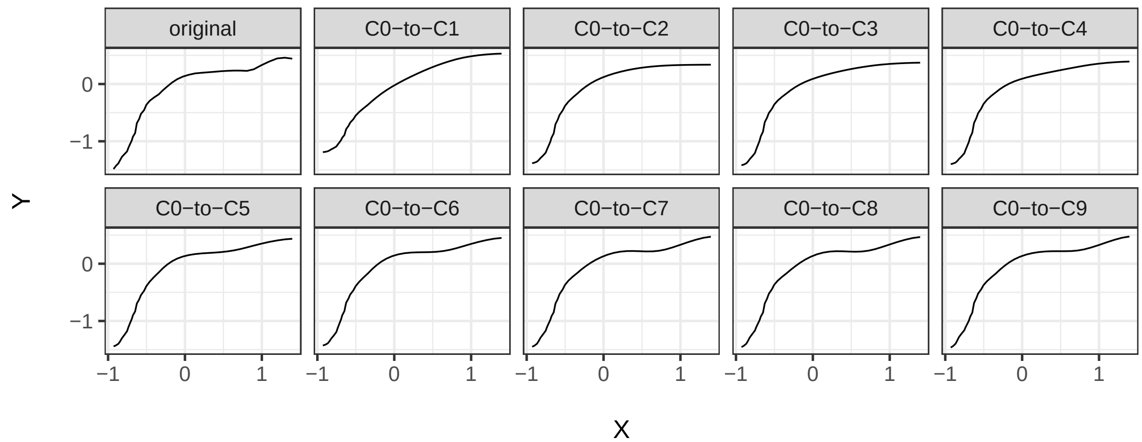

Figure 2: Original midsagittal tongue shape for one token plus DCT reconstructions of the same data using varying numbers of DCT coefficients. The tongue tip is on the right of the image and the tongue root is on the left. The token represents a single spline taken from a word-initial rhotic.

In order to empirically evaluate the number of DCT coefficients needed to summarise each tongue shape, we fitted DCTs to all tongue splines (11 per token, representing 11 time-points) with different numbers of coefficients, ranging from 2 coefficients to 10 coefficients, which gives us 9 different options to evaluate. We then conducted an inverse DCT in order to reconstruct the original signal from these coefficients, which essentially gives us a DCT-smoothed version of the original signal. Following Shaw and Kawahara (2018), we then calculate Pearson's correlation between the original signal and the DCT-reconstructed signal and plot these correlation values for different numbers of DCT coefficients. Figure 3 shows that 3 coefficients yields correlations with the original signal of $r > .95$ for all speakers. As shown above, however, there are some advantages to the higher DCT coefficients, particularly for more complex tongue tip shapes. To this end, we ran testing using the same classification analysis that we report later in this paper, examining the effects of between 4–8 DCT coefficients on classification accuracy for each sonorant*position. Laterals and nasals did not benefit from more than 5 coefficients, but the inclusion of a 6th DCT coefficient improved word-initial rhotic classification by almost 10%. We anticipate that this is because it captures the subtle tongue tip shaping depicted in Figure 2. After settling on 6 DCT coefficients, we normalized each coefficient by z-scoring each speaker's data across all productions.

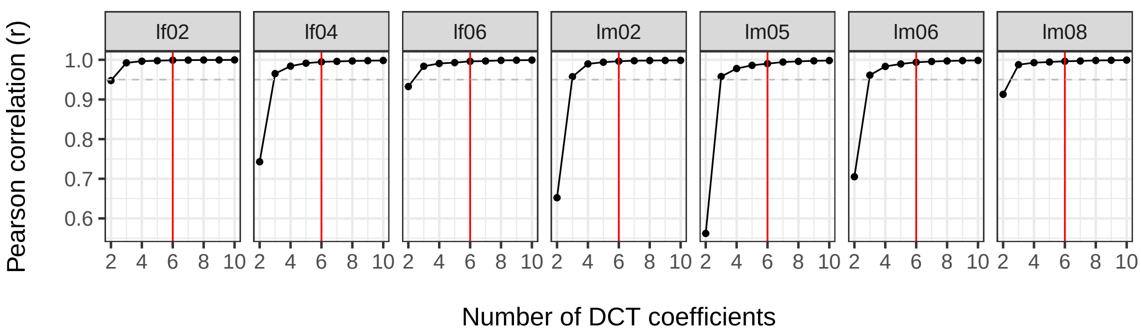

Figure 3: Pearson's correlation between the original ultrasound tongue splines and DCT-smoothed versions using varying numbers of DCT coefficients. The solid vertical line represents the final number of DCT coefficients used for the classification analysis.

2.6. Summarising high-dimensional dynamic information

At this point, the acoustic data are represented by 6 MFCC trajectories sampled at 11 points in time (= 66 points), and the ultrasound spline data are represented by 6 DCT trajectories sampled at 11 points in time (= 66 points). This already represents considerable dimensionality reduction from a time-varying power spectrum or time-varying ultrasound spline, but we conducted further dimensionality reduction of the dynamic data using an approach inspired by Nossair and Zahorian (1991). This involves fitting a Discrete Cosine Transformation (DCT) to each of the time-varying MFCC (acoustics) and DCT (ultrasound) coefficients discussed above, which allows us to summarise the shape of each of those coefficient trajectories over time (see Marin et al. 2010 for a similar approach to spectral data). This provides a higher-level set of coefficients that encodes the shape of each time-varying MFCC or DCT coefficient, each of which summarises some dynamic aspect of spectral shape or tongue shape.

We empirically evaluated the number of DCT coefficients needed to summarise each trajectory in the same way as for the ultrasound spline fitting, which is plotted in Figure 4. We find that 3 DCT coefficients returns correlations of $r > .9$ for all acoustic-MFCC trajectories and $r > .95$ for ultrasound-DCT trajectories, except for the 6th coefficient in both sets (MFCC6 and DCT5), which are slightly below these values. However, the MFCC/DCT trajectories are not always smooth functions of time and we avoid seeking higher correlations as we wish to avoid overfitting to the signal. Accordingly, we choose 3 DCT coefficients to represent both sets of trajectories, which captures the mean, slope and curvature of each coefficient trajectory over time. This means that each of the 6 acoustic-MFCC and 6 ultrasound-DCT dynamic trajectories is summarised by 3 DCT coefficients. As a result, each token's time-varying acoustic spectrum or ultrasound tongue spline across the sonorant-vowel interval is represented by 18 (6×3) values.

In summary, our final inputs to our model are as follows. We have

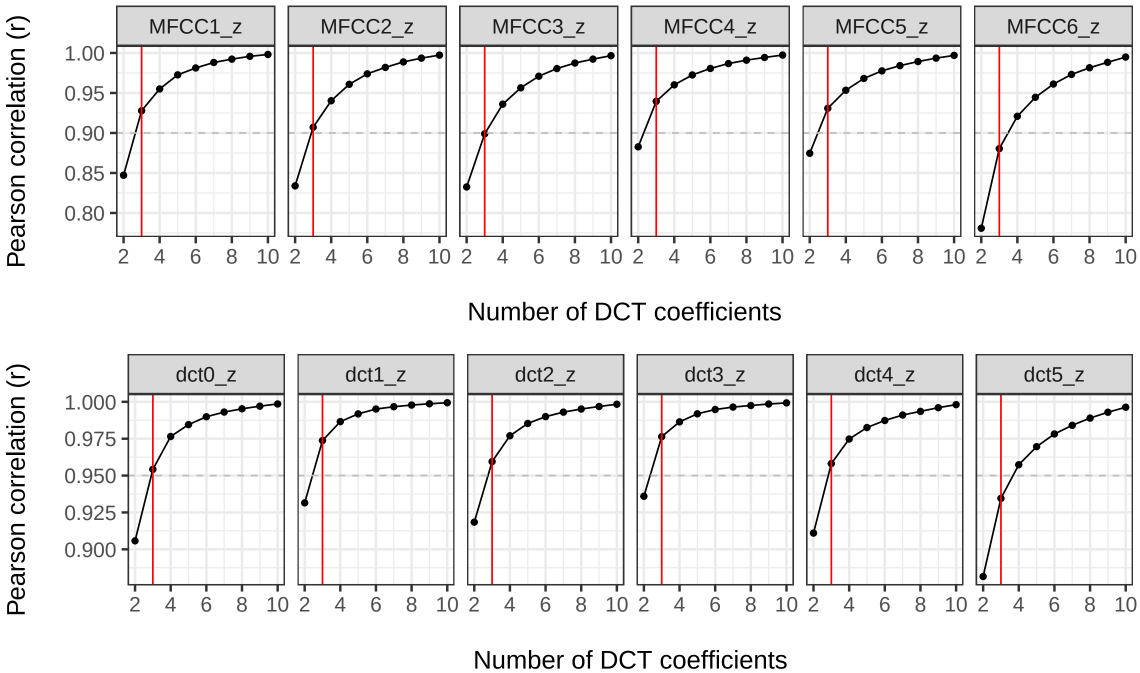

Figure 4: Pearson's correlation between dynamic acoustic-MFCC trajectories and DCT-smoothed versions using varying numbers of DCT coefficients (top) and the dynamic ultrasound-DCT trajectories and DCT-smoothed versions using varying numbers of DCT coefficients (bottom). The solid vertical line represents the final number of DCT coefficients used for the classification analysis. The dashed horizontal line represents the correlation coefficient cut-off used for selecting the number of DCT coefficients for each measure, which was based on the first 5 dynamic MFCC/DCT trajectories.

compressed a complex power spectrum sampled at 11 points in time for each token to 18 values. These values are a compressed representation of how the spectrum changes over the sonorant-vowel interval. We have also compressed time-varying ultrasound tongue splines sampled at 11 points in time for each token to 18 values, which represents how midsagittal tongue shape changes over the sonorant-vowel interval. These compressed representations correlate well with the original signals and should, therefore, capture important information in the original signals. We now turn to the details of the classification analysis.

2.7. Classification analysis

We use support vector machines (SVMs) in order to establish how robustly the three-way phonemic contrast can be classified for each sonorant, based on an initial training phase mapping phonological categories to acoustic and articulatory feature sets. SVMs are a class of supervised statistical learning models that aim to find the hyperplane that maximally separates two classes in N-dimensional space (Boser et al., 1992; James et al., 2013). The hyperplane is located at the *maximum margin*, which is the largest difference between data points of the two classes. Non-linear separation between classes is typically achieved via a kernel, whereby the data are transformed into a higher-dimensional space and linear classification is then performed in this high-dimensional space. SVMs are a binary classification method but multi-class classification can be achieved in various ways. The method we use is the one-against-one technique, in which each category is compared against one other category. This process is repeated for all combinations of categories, with each classifier voting for one category and the category with the highest number of votes being classified accordingly. SVMs have been widely applied to speech data (Clarkson and Moreno, 1999; Wang et al., 2013; Yu, 2017) and are typically reported to show good phoneme classification accuracy on acoustic and articulatory signals. One reason for this is that SVMs are concerned with the margins between classes, rather than the mean and variance of each class, meaning that a larger data set is better only insofar as the additional data better represents the boundaries between classes.

Models were fitted using the e1071 package in R (Meyer et al., 2021). We fitted separate models to each combination of sonorant type and position, such as word-initial laterals, word-final laterals, word-initial nasals, etc. Each model had phoneme as the outcome variable and the 18 dynamic acoustic features or the 18 dynamic ultrasound features as the predictor variables. Each feature set was randomly split into 80% training and 20% testing subsets. All models were fitted using a radial basis function kernel, and parameter tuning for each model was conducted on the training data only using a grid search over a range of values for $\gamma = \{10^{-6}, 10^{-5}, ..., 10^{-1}\}$ and $C = \{0.1, 1, 10\}$, with model performance evaluated using 10-fold cross-validation. The model with the optimal parameters was used to predict the phonemic identity of the 20% test data set based only on the input measurements (with separate models for acoustic and ultrasound data). In order to mitigate against splitting a small data set, we used Monte Carlo cross-validation (Picard and Cook, 1984; Kuhn and Johnson, 2013), which involved running 100 iterations of the train-test procedure for each model, using a different random train-test split each time. We then averaged over the 100 iterations to produce a final classification matrix and overall classification rates.[3] All code and data used for analyses

[3] In order to empirically determine the chance classification rate for a data set comparable in size and dimensionality to the models used here, we generated simulated data with 18 numerical variables corresponding to the 18 MFCC/DCT coefficients, each of which was populated with random values from a normal distribution $\mathcal{N}(0, 1)$ and then each observation was randomly assigned one of three phoneme labels (plain, palatalised, velarised). We then ran the same procedure described above for the real models and found an average overall classification rate of 31.5-36.94% on random data, depending on the same size, which is

in this paper is available at: `https://osf.io/dfe7g/`.[4]

3. Results

3.1. Laterals

The lateral acoustic model in Table 1 shows overall classification rates of 74.46% (initial) and 81.27% (final), which represents well above chance classification. The classification matrix for initial laterals shows that /l̪ʲ/ is the most accurately classified at 78.59%, while /l/ is the worst at 63.54%. Note that the majority of inaccurate classifications for /l̪ʲ/ in initial and final context are as /l̪ˠ/, suggesting some overlap in the correlates of velarised and palatalised lateral phonemes. Classification for word-final laterals is better than initial laterals, but word-finally /l̪ˠ/ is the most accurately classified phoneme at 91.77%. Overall, this suggests that initial and final laterals have broadly similar classification rates, with the palatalised and velarised phoneme being most similarly distinct initially and the velarised phoneme being most distinct finally.

The lateral ultrasound model in Table 2 shows overall classification rates of 73.37% (initial) and 83.04% (final), but these statistics particularly obscure considerable between-phoneme differences in classification, suggesting slightly more robust lateral contrasts in the ultrasound data. In word-initial context, /l̪ˠ/ shows rather poor classification of 59.03%, with 31.79% of productions

close to the theoretical chance level of 33.$\dot{3}$% for three-way classification.

[4]Sensitivity testing and initial modelling was carried out using Lancaster University's High End Computing facility, after which final models were fitted locally for the publicly available documentation.

	WORD-INITIAL			WORD-FINAL		
	l̪ˠ	l	l̪ʲ	l̪ˠ	l	l̪ʲ
l̪ˠ	76.55	8.92	14.54	91.77	1.29	6.94
l	27.83	63.54	8.63	22.95	70.55	6.49
l̪ʲ	20.43	0.98	78.59	25.22	1.79	72.98
	Overall: 74.46%			Overall: 81.27%		

Table 1: SVM classification matrix for lateral acoustic data. Values represent percentage correct classification (rounded to 2 decimal places).

being misclassified as /l̪ʲ/. Outside of this phoneme, the other phonemes are classified better than the acoustic MFCC data. This is also true for word-final laterals, except for /l̪ˠ/ being slightly better classified in the acoustic data (91.77% vs 89.77%).

	WORD-INITIAL			WORD-FINAL		
	l̪ˠ	l	l̪ʲ	l̪ˠ	l	l̪ʲ
l̪ˠ	59.03	9.18	31.79	89.77	1.11	9.11
l	18.88	80.81	0.31	15.05	81.28	3.67
l̪ʲ	14.29	0	85.71	21.37	1.84	76.79
	Overall: 73.37%			Overall: 83.04%		

Table 2: SVM classification matrix for lateral ultrasound data. Values represent percentage correct classification (rounded to 2 decimal places).

In summary, the laterals data show variability in classification, but with slightly better classification in word-final context and substantially above-chance classification in all cases. The models show that /l̪ˠ/ and /l̪ʲ/

are most often misclassified as each other and only very rarely as /l/. This suggests that while velarised and palatalised laterals do have some distinctive acoustic and articulatory correlates, there is a reasonable amount of overlap in these categories, which leads to occasional misclassification. The acoustic and articulatory data show relatively similar findings, except for substantially poorer classification for initial /l̪ˠ/ in the ultrasound data.

3.2. Nasals

The nasal acoustic model in Table 3 shows overall classification of 86.67% (initial) and 85.53% (final), which is higher than for laterals. Our previous work has reported less robust distinctions between nasal phonemes in Gaelic (Nance and Kirkham, 2020), but that analysis did not take formant transitions or acoustic dynamics into account. Indeed, our present analysis suggests that such dynamics are crucial to this contrast, and fitting comparable SVMs to a single time-point at the nasal steady-state reduces classification accuracy substantially (see Section 3.4).

We find that classification is relatively similar between positions. For example, /n/ is the worst classified phoneme in initial (81.22%) and final (82.32%) position, although both remain well classified. The velarised and palatalised phonemes are classified very similarly across both positions, suggesting a relatively high degree of distinctiveness between the acoustic correlates of all three phonemes.

The nasal ultrasound model in Table 4 is very similar to the acoustics model, with overall classification of 84.70% (initial) and 89.81% (final). /n/ is classified better in final position (94.68%) than in initial position (84.10%), but classification remains high in all cases.

	WORD-INITIAL			WORD-FINAL		
	n̪ˠ	n	n̪ʲ	n̪ˠ	n	n̪ʲ
n̪ˠ	87.05	6.24	6.71	86.17	0.57	13.26
n	15.02	81.22	3.76	10.52	82.32	7.16
n̪ʲ	9.08	0.58	90.34	9.96	1.64	88.40
	Overall: 86.67%			Overall: 85.53%		

Table 3: SVM classification matrix for nasal acoustics data. Values represent percentage correct classification (rounded to 2 decimal places).

	WORD-INITIAL			WORD-FINAL		
	n̪ˠ	n	n̪ʲ	n̪ˠ	n	n̪ʲ
n̪ˠ	80.13	7.92	11.95	91.79	0	8.21
n	9.29	84.10	6.61	4.58	94.68	0.74
n̪ʲ	9.01	0.90	90.10	12.27	2.09	85.64
	Overall: 84.70%			Overall: 89.81%		

Table 4: SVM classification matrix for nasal ultrasound data. Values represent percentage correct classification (rounded to 2 decimal places).

Overall, nasals show better classification than laterals in acoustics and articulation. Word-final phonemes are slightly better classified than word-initial phonemes in articulation, but this is only a small difference. This stands in contrast to our previous research, where we found weak distinctions between nasal phonemes. We propose that our current model classifies nasals very effectively due to the incorporation of dynamic information across the nasal and adjacent vowel, suggesting that cues to the three-way contrast in nasals is highly dynamic. We pursue this idea further in Section 3.4.

3.3. Rhotics

The rhotic acoustics model in Table 5 shows overall classification of 91.14% (initial) and 73.19% (final). This means that rhotics show the best average classification accuracy in initial position but the worst in final position across all sonorant types in acoustics. We find very robust maintenance of initial rhotic contrasts, with /r̪ˠ/ at 92.99%, /r/ at 90.16% and /r̪ʲ/ at 89.20. In particular, /r/ is hardly ever misclassified as /r̪ʲ/ (0.08%), which is impressive given that these results represent the average of 100 model runs, meaning that there was near-zero confusion between /r/ and /r̪ʲ/. In contrast, word-final rhotics show the poorest classification of any models, with classifications of /r̪ˠ/ = 75.14%, /r/ = 63.28% and /r̪ʲ/ = 78.41%. These misclassifications are still substantially above chance classification, but it suggests that the word-final categories have less robust phonetic correlates than word-initial categories, which leads to poorer classification accuracies.

	WORD-INITIAL			WORD-FINAL		
	r̪ˠ	r	r̪ʲ	r̪ˠ	r	r̪ʲ
r̪ˠ	92.99	5.95	1.06	75.14	4.41	20.45
r	9.76	90.16	0.08	13.55	63.28	23.17
r̪ʲ	7.87	2.92	89.20	14.11	7.48	78.41
	Overall: 91.14%			Overall: 73.19%		

Table 5: SVM classification matrix for rhotic acoustic data. Values represent percentage correct classification (rounded to 2 decimal places).

The rhotic ultrasound model in Table 6 shows overall classification of 85.07% (initial) and 65.65% (final), showing the same patterning between initial and final context but with slightly poorer performance than in acoustics. Accordingly, every phoneme is classified slightly worse than the acoustics model in both positions, except for word-final /r/, which is near identical between the two modalities. Interestingly, the robustness of word-initial classification is evidenced in the fact that /r̪ʲ/ is never misclassified as /r/ and /r/ is never misclassified as /r̪ʲ/, suggesting a categorical distinction between these phonemes in articulatory dynamics. This suggests that the palatalisation gesture in initial rhotics is highly distinct from the articulation of the plain rhotic. In contrast, there are varying degrees of confusion between palatalised and velarised rhotics, although these categories are still fairly well classified.

	WORD-INITIAL			WORD-FINAL		
	r̪ˠ	r	r̪ʲ	r̪ˠ	r	r̪ʲ
r̪ˠ	84.12	13.75	2.14	58.42	6.20	35.38
r	12.00	88.00	0	15.28	63.80	20.92
r̪ʲ	18.40	0	81.60	21.77	4.80	73.43
	Overall: 85.07%			Overall: 65.65%		

Table 6: SVM classification matrix for rhotic ultrasound data. Values represent percentage correct classification (rounded to 2 decimal places).

Overall, the most striking result for the rhotics is that while classification is the best of all models for initial rhotics, it is the lowest for final rhotics. The acoustic data for initial rhotics also outperform the ultrasound data in classification accuracy. This suggests that there exist clear correlates of the three-way contrast for initial rhotics, especially in acoustics, but much weaker phonetic correlates for the contrast in final rhotics.

3.4. Comparison between dynamic models and sonorant steady-state

Finally, we compare the models in the above sections with models fitted to the midpoint of the sonorant steady-state, which was defined in Nance and Kirkham (2020) as a labelled interval that captures relatively static formant values during an unambiguously lateral, nasal or rhotic phase. The steady-state model structure was the same as for the dynamic models, but as there is only one time-point, there are only 6 MFCCs for the acoustics and 6 DCTs summarising the ultrasound tongue shape, with no additional dynamic information. Table 7 shows the average classification accuracy for each model, with comparison between steady-state and dynamic models. To re-cap, these values represent the average classifications over 100 Monte Carlo cross-validation train-test iterations.

Table 7 shows that the dynamic models produce higher average classification accuracies in all cases, with the exception of the initial laterals acoustics model, where the dynamic model is 2.53% worse. However, the magnitude of the difference between steady-state and dynamic models is highly variable between sonorants. In acoustics, the impact of dynamics on classification is largest for nasals (24.81% higher in initial, 34.26% higher in final) and is higher than 10% for all models except initial laterals. In the ultrasound data, the differences are generally smaller, with negligible differences for laterals, final nasals and initial rhotics, but with substantial improvement for initial nasals (12.67%) and final rhotics (24.69%) when dynamic information

modality	sonorant	position	steady-state	dynamic	difference
acoustics	lateral	initial	76.99	74.46	−2.53
		final	62.28	81.27	18.99
	nasal	initial	61.86	86.67	24.81
		final	51.27	85.53	34.26
	rhotic	initial	78.84	91.14	12.30
		final	51.33	73.19	21.86
articulation	lateral	initial	68.58	73.37	4.79
		final	76.04	83.04	7.00
	nasal	initial	72.03	84.70	12.67
		final	86.81	89.81	3.00
	rhotic	initial	76.24	85.07	8.83
		final	40.96	65.65	24.69

Table 7: SVM average classification accuracies (%) for models fitted to the sonorant steady-state (steady-state) and the whole sonorant-vowel interval (dynamic). The 'difference' column represents the dynamic model accuracy minus the steady-state model accuracy, with positive values indicating % improvement for the dynamic model over the steady-state model and negative values indicating better relative performance on the steady-state model.

is included.

Overall, this comparative analysis suggests that the contrastive correlates of phonological palatalisation take on a particularly dynamic quality for all sonorants in acoustics, except for initial laterals, and also take on a dynamic quality for initial nasals and final rhotics in the articulatory data. There are fewer dynamic cues to contrast in the ultrasound data, compared with acoustics, with many sonorants not benefitting from the addition of dynamic articulatory information beyond a single theoretically-informed time-point at the sonorant steady-state.

3.5. Summary of results

We conducted classification analyses on the three-way contrast in laterals, rhotics and nasals in Scottish Gaelic, with separate models for word position and acoustic/articulatory data. We use classification accuracy as a proxy for the relative stability of each three-way contrast. In word-initial position, we find that rhotics are best classified, followed by nasals, and then laterals. This overall pattern is observed in both the acoustic and articulatory data, with the acoustic data always showing better overall classification rates. In word-final position, nasals are classified best, followed by laterals, and then rhotics. This overall pattern is observed in both the acoustic and articulatory data, with the articulatory data showing slightly better classification for final laterals and nasals, but not for rhotics. Finally, we show that incorporating dynamic information about the entire sonorant-vowel sequence improves classification accuracy by between 12.30% and 34.26% in the acoustic data, except for initial laterals, which are slightly worse when dynamics are included. However, the articulatory data show less overall improvement, with only initial nasals and final rhotics showing improvement of over 10% when dynamics are included. In the following section, we discuss the implications of these results for the role of dynamics in contrast maintenance and the stability of palatalisation contrasts.

4. Discussion

4.1. Variable stability of synchronic contrasts

A consistent finding in this study is that nasals have higher classification accuracy than laterals. We did not predict this based on the previous Gaelic research, but there are good reasons to believe this result, the most obvious of which is the inclusion of dynamic information in our models. Formant transitions are well known to be a strong cue to place of articulation, particularly for nasals (Malécot, 1956; Wright, 2004), which is due to the weakening of the upper formants due to nasal anti-formants in the spectrum. Indeed, Iskarous and Kavitskaya (2018) find nasals to be more distinctive than laterals in formant transitions. The inclusion of dynamic information for nasals is, therefore, a plausible reason for why we find better acoustic contrast in nasals than laterals, in contrast to Nance and Kirkham (2020), where we only analysed formants at the sonorant steady-state. This is supported by our finding that laterals are classified better than nasals in our steady-state models, but that nasal classification drastically improves when we incorporate dynamic information across the sonorant-vowel interval. From this, we can conclude that the three-way nasal contrast in Gaelic is fundamentally dynamic in nature and likely more so than for laterals or rhotics, due to the relevant cues to contrast being more temporally distributed for nasals.

We predicted that rhotics would show the weakest classifications, based on previous research (Kochetov, 2005; Stoll, 2017; Iskarous and Kavitskaya, 2018). This is true word-finally, but certainly not word-initially, which is in line with our previous work on Gaelic. In Nance and Kirkham (forthcoming 2022) we report strong evidence of contrast in initial rhotics based on low-dimensional phonetic information, such as formant frequencies, so it is unsurprising that we also find good classification for rhotics when we take even more information into account. We do find, however, that final rhotics are classified comparably worse than any other sonorant, which supports the tendency towards contrast neutralisation in final rhotics. It is well-known that codas contain weaker acoustic cues for place of articulation than onsets (Ohala, 1990; Wright, 2004). Gaelic is unusual in having an overall VC structure, similar to Irish (Hammond et al., 2014; Ní Chiosáin et al., 2012), but, despite this, the proposal that acoustic cues are weaker in syllable-final position remains and is backed up by perceptual research. For example, Kochetov (2002) and Ní Chiosáin and Padgett (2012) both find that listeners are less likely to distinguish palatalised and non-palatalised pairs in VC contexts compared with CV contexts. This factor may explain the tendency for initial rhotics to show more robust distinctions than final rhotics, but this logic does not appear to extend to laterals or nasals, which show similar classification between positions and sometimes slightly better classification in final position.

We now briefly comment on how our model compares with human listeners; in other words, can Gaelic speakers accurately perceive phonemic identity from similar acoustic information to what we analyse here? Listeners can distinguish palatalised and non-palatalised consonants with high accuracy (Kochetov, 2002; Ní Chiosáin and Padgett, 2012; Spinu et al., 2012), even when they do not speak a language with palatalisation contrasts. Babel and Johnson (2010) found that American English listeners performed no differently from Russian listeners at a fast-paced AX discrimination task comparing word-initial Russian palatalised and non-palatalised consonants, although Hacking et al. (2016) show that L2 English learners have greater difficulty producing the Russian contrast word-finally. Our rhotics results are in line with the above research showing better perceptual discrimination between palatalised and non-palatalised consonants in CV contexts compared with VC contexts. In summary, we consider our machine classification to be comparable to the discrimination capabilities of a human listener.

4.2. The dynamic nature of palatalisation contrasts

A major finding of this study is the extent to which the incorporation of dynamic information improves acoustic classification. This was particularly true of nasals, but, surprisingly, we find little difference between the steady-state and dynamic models for initial laterals. It could be the case that the sonorant steady-state is where the primary cues for such contrasts exist in laterals. However, we also find other insensitivities to model adjustments in the initial laterals data. For example, during sensitivity testing we found that increasing or decreasing the number of coefficients had the least effect on initial laterals. It may be that the acoustic and articulatory data used here provides an adequate representation for this context, with reasonable accuracies of 73–75%, but that the highly audible contrast we perceive for initial laterals has other acoustic and articulatory correlates that are not well captured in this study.

Despite the strong contribution of dynamics to acoustic classification, we find this to a much lesser degree with the articulatory data. This may be a consequence of dynamic non-linearity in acoustic-articulatory relations (Stevens, 1989; Strycharczuk and Scobbie, 2017; Gorman and Kirkham, 2020), whereby articulatory variation in some parts of the vocal tract does not produce proportionate change in the acoustic output, at least in terms of the parameters measured here. Another explanation could be the nature of the acoustic and articulatory representations used in this study. For instance, MFCCs capture rich details of the acoustic spectrum, whereas the midsagittal tongue shape obtained by ultrasound imaging is already a very sparse representation of the three-dimensional oral tract. Furthermore, it is possible that the the lesser contribution of dynamics to articulatory classification may be a consequence of our focus on global change in midsagittal tongue shape. It may be the case that other aspects of articulatory timing, such as the relative timing of coronal, palatalisation and velarisation gestures, represent stronger articulatory cues to contrast than overall change in tongue shape. We plan to explore this further in future research, with the aim of better understanding the articulatory dynamics of palatalisation contrasts.

Finally, we must highlight some caveats for interpreting the comparison between steady-state and dynamic models. First, the inputs to each model necessarily differ in dimensionality (6 for steady-state, 18 for dynamic). While this is an obvious consequence of incorporating time-varying information into the dynamic model, a larger number of parameters increases the possibility of overfitting and producing overly optimistic classification rates, so it would be valuable to further evaluate the effects of parameter space size on a much larger data set. We also cannot discount the possibility that the dynamic model is picking up on vowel cues that correspond to lexical items, rather than the phonetic correlates of deep phonological structure. In other words, by incorporating information from the sonorant and the adjacent vowel, we could be identifying mostly word-specific information. In part, this is unavoidable, as Gaelic has relatively few true minimal triplets for these contrasts, but it would be worthwhile testing on languages where such contrasts have a higher functional load, such as Russian. Finally, our analysis demonstrates the extent to which dynamic information contributes towards classification accuracy, but does not tell us the precise nature of this dynamic information. In future research, we plan to examine the temporal dynamics of the lingual gestures involved in Gaelic palatalisation contrasts.

4.3. The diachronic typology of palatalisation contrasts

We made the prediction that sonorants with a greater propensity towards diachronic phonological loss across a language family would show synchronically weaker contrasts. This was grounded in the principle that processes of diachronic change can be inferred from synchronic snapshots (Labov, 1994). In our case, the diachronic predictions suggested that laterals should have the highest classification rates and rhotics the lowest classification rates, given that lateral contrasts are best-maintained across the Goidelic language family and rhotics the least well-maintained. Our results only support the diachronic predictions when we focus solely on the sonorant steady-state, which is a partial and insufficient representation of palatalisation contrasts. When we take into account the dynamics of how the palatalisation gesture unfolds over time, we instead find a different set of results that interact strongly with word position. To re-cap, rhotics are best classified in initial position and worst in word-final position, with nasals being relatively well classified in all contexts, and laterals always being classified less accurately than nasals.

The word-final rhotic synchronic data, however, do pattern with diachronic trends towards neutralisation across Goidelic. Cross-linguistically, it has been shown that large rhotic inventories are subject to simplification, with palatalised rhotics particularly susceptible to loss (Hall, 2000). We anticipate that competing biomechanical demands on palatalised rhotics can lead to partial masking of the palatalisation gesture, especially in word-final position. For instance, Stoll (2017) reports more variable gestural timing in palatalised rhotics compared with laterals, which may also lead to greater overlap between rhotic categories. Given sufficient exposure, this increased overlap is likely to cause instances of misperception and subsequent recategorisation of a listener's phonological system, leading them to produce smaller distinctions between rhotic phonemes (Ohala, 1981, 1989). Moreover, if the reduced variants become recognised as acceptable by other community members, possibly due to the low functional load of the contrast, this is likely to accelerate the long-term progression of contrast neutralisation (Beckman et al., 1992; Bybee, 2015).

Nasals are especially interesting in this case as Goidelic diachronic data suggests they are retained more frequently than large rhotic systems, but less frequently than large lateral systems. In Slavic, on the other hand, palatalised nasals are very frequently maintained cross-linguistically, more so than laterals and rhotics (Carlton, 1990; Iskarous and Kavitskaya, 2010). Our data pattern more closely with the reported typology of Slavic sonorant development, with nasal phonemes produced more distinctively than laterals and final rhotics. This is surprising in light of previous research, some of which has suggested only a two-way contrast in Gaelic nasals (Ladefoged et al., 1998; Nance and Kirkham, 2020), but it may be the case that the Gaelic contrast has been maintained by temporally distributing the phonetic cues to contrast across the sonorant-vowel interval, which has not previously been investigated as thoroughly. We are unable to claim whether this is a novel development in Gaelic, but previous research on Slavic has also shown that nasals may sometimes show more robust contrasts than laterals in formant transitions (Iskarous and Kavitskaya, 2018), so it is likely that a similar pattern recurs in our data.

In summary, we find a more complex relationship between diachronic predictions and the variable stability of synchronic contrasts than we initially predicted. We believe, however, that the sociolinguistic context of Gaelic is highly informative in understanding these results. Gaelic is a minoritised language that is currently undergoing intense revitalisation. Minority languages often experience structural simplification (Dorian, 1981; Jones, 1998), but we note that speakers of Gaelic often have high levels of metalinguistic awareness about the language's phonology (Nance et al., 2016). All of the speakers in our study worked in Gaelic-essential jobs and, therefore, represent highly professional speakers of the language. The strong investment of such speakers in maintaining Scottish Gaelic also increases the likelihood of them learning to produce traditionally-reported contrasts in the language, which are often acquired through education. This sociolinguistic context, therefore, may represent one of the contributing mechanisms for the preservation of structures that would otherwise be likely to undergo loss in more typical cases of community transmission (Nance and Kirkham, forthcoming 2022). It is clear from this that identifying potential future paths of sound change in the Gaelic sonorant system will also require detailed attention to the changing sociolinguistics dynamics of Gaelic.

5. Conclusion

This study has examined the variable synchronic stability of palatalisation contrasts in light of claims that such contrasts are prone to diachronic simplification, reduction or loss. The cross-linguistic diachronic evidence suggested that laterals would show the most robust contrasts and rhotics the least robust contrasts. We do indeed find that rhotics are most poorly classified word-finally, which may reflect the diachronic trend towards contrast reduction, but we find the opposite pattern word-initially, where rhotic contrasts are highly robust. This demonstrates that some contrasts in Gaelic are robustly maintained despite intense pressures towards diachronic reduction. We do not find evidence to support the claim that laterals show more robust contrast than nasals, with both sonorants being well-classified, but with nasals showing better classification once dynamic information is taken into account. Accordingly, we find that synchronic speech production data bears a complex relationship with long-term patterns of diachronic change reported across the Goidelic languages, and it is likely that a fuller consideration of how phonological dynamics interact with changing sociolinguistic contexts will further illuminate the potential paths of sound change in Gaelic. Overall, we find evidence of weaker contrast in predictably unstable sonorants, but elsewhere we find that contrast is often more robust than previously anticipated, with the phonetic correlates of phonological structure located firmly in the temporal dynamics of the speech signal.

Acknowledgements

We are very grateful to Marianne Pouplier and two anonymous reviewers for their constructive comments on this article. Thanks also to Lois Fairclough for manual correction of the ultrasound data. This study was conducted with the assistance of two grants from the Faculty of Arts and Social Sciences at Lancaster University. We also acknowledge the use of the High End Computing facility at Lancaster University.

Appendix

Articulate Instruments, 2008. Ultrasound stabilisation headset: Users manual revision 1.5. Articulate Instruments, Edinburgh.

Articulate Instruments (Ed.), 2018. Articulate Assistant Advanced version 2.17. Articulate Instruments, Edinburgh.

Babel, M., Johnson, K., 2010. Accessing psycho-acoustic perception and language-specific perception with speech sounds. Laboratory Phonology: Journal of the Association for Laboratory Phonology 1 (1), 179–205.

Ball, M., Müller, N. (Eds.), 2009. The Celtic languages, 2nd Edition. Routledge, London.

Bateman, N., 2007. A crosslinguistic investigation of palatalization. Ph.D. thesis, University of California San Diego, San Diego.

Beckman, M., De Jong, K., Jun, S.-A., Lee, S.-H., 1992. The interaction of coarticulation and prosody in sound change. Language and Speech 35 (1), 45–58.

Bennett, R., Ní Chiosáin, M., Padgett, J., McGuire, G., 2018. An ultrasound study of Connemara Irish palatalization and velarization. Journal of the International Phonetic Association, 1–44.

Boersma, P., Weenink, D., 2020. Praat: doing phonetics by computer. Version 6.1.16, retrieved 6 June 2020 from http://www.praat.org/.

Boser, B. E., Guyon, I. M., Vapnik, V. N., 1992. A training algorithm for optimal margin classifiers. COLT '92: Proceedings of the fifth annual workshop on Computational learning theory (144–152).

Broderick, G., 2009. Manx. In: Ball, M., Müller, N. (Eds.), The Celtic languages. Routledge, London, pp. 305–356.

Bybee, J., 2015. Articulatory processing and frequency of use in sound change. In: Honeybone, P., Salmons, J. (Eds.), The Oxford Handbook of Historical Phonology. Oxford University Press, pp. 467–484.

Carlton, T., 1990. Introduction to the phonological history of the Slavic languages. Slavica, Bloomington.

Clarkson, P., Moreno, P., 1999. On the use of support vector machines for phonetic classification. IEEE International Conference on Acoustics, Speech, and Signal Processing 2, 585–588.

Davis, S. W., Mermelstein, P., 1980. Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. IEEE Transactions on Acoustics, Speech, and Signal Processing 28 (4), 357–366.

Dorian, N., 1981. Language death: The life cycle of a Scottish Gaelic dialect. University of Pennsylvania Press, Philadelphia.

Farnetani, E., Provaglio, A., Ní Chasiade, A., Faely, G., Recasens, D., Fontdevila, J., Pallarues, M. D., 1991. A study of the production and coarticulatory characteristics of palatal and palatalized consonants. In: ESPRIT II: Research Action No. 3279, Progress Report. Consiglio Nazionale delle Ricerche.

Gorman, E., Kirkham, S., 2020. Dynamic acoustic-articulatory relations in back vowel fronting: Examining the effects of coda consonants in two dialects of British English. Journal of the Acoustical Society of America 148 (2), 724–733.

Greene, D., 1973. The growth of palatalisation in Irish. Transactions of the Philological Society 72, 127–136.

Hacking, J. F., Smith, B. L., Nissen, S. L., Allen, H., 2016. Russian palatalized and unpalatalized coda consonants: An electropalatographic and acoustic analysis of native speaker and L2 learner productions. Journal of Phonetics 54, 98–108.

Hall, T., 2000. Typological generalisations concerning secondary palatalization. Lingua 110, 1–25.

Hammond, M., Warner, N., Davis, A., Carnie, A., Archangeli, D., Fisher, M., 2014. Vowel insertion in Scottish Gaelic. Phonology 31 (1), 123 – 153.

Harrington, J., 2010. Acoustic phonetics. In: Hardcastle, W. J., Laver, J., Gibbon, F. E. (Eds.), The Handbook of Phonetic Sciences, 2nd Edition. Wiley-Blackwell, Chichester, pp. 81–129.

Hickey, R., 1995. Sound change and typological shift: Initial mutation in Celtic. In: Frisiak, J. (Ed.), Linguistic typology and reconstruction. Mouton, Berlin, pp. 133–182.

Hickey, R., 2014. The sound structure of Modern Irish. Mouton de Gruyter, Berlin.

Howson, P., 2018. Rhotics and palatalization: An acoustic examination of Upper and Lower Sorbian. Phonetica 75, 132–150.

Iskarous, K., Kavitskaya, D., 2010. The interaction between contrast, prosody, and coarticulation in structuring phonetic variability. Journal of Phonetics 38, 625–639.

Iskarous, K., Kavitskaya, D., 2018. Sound change and the structure of synchronic variability: Phonetic and phonological factors in Slavic palatalization. Language 94 (1), 43–83.

James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An Introduction to Statistical Learning with Applications in R. Springer, New York.

Jones, M., 1998. Language obsolescence and revitalization: Linguistic change in two sociolinguistically contrasting Welsh communities. Clarendon Press, Oxford.

Kavitskaya, D., Iskarous, K., Noiray, A., Proctor, M., 2009. Trills and palatalization: Consequences for sound change. In: Proceedings of the formal approaches to Slavic linguistics. Yale University, pp. 97–110.

Kochetov, A., 2002. Production, perception, and emergent phonotactic patterns: A case of contrastive palatalization. Ph.D. thesis, University of Toronto, Toronto.

Kochetov, A., 2005. Phonetic sources of phonological assymetries: Russian laterals and rhotics. In: Gurski, C. (Ed.), Proceedings of the 2005 annual conference of the Canadian Linguistic Association. University of Western Ontario, London, Ontario.

Kochetov, A., 2017. Acoustics of Russian voiceless sibilant fricatives. Journal of the International Phonetic Association 47 (3), 321–348.

Kochetov, A., Petersen, J. H., Arsenault, P., 2020. Acoustics of Kalasha laterals. Journal of the Acoustical Society of America 147, 3012–3027.

Kuhn, M., Johnson, K., 2013. Applied Predictive Modeling. Springer, New York, NY.

Labov, W., 1994. Principles of linguistic change: Internal factors. Blackwell, Oxford.

Ladefoged, P., Ladefoged, J., Turk, A., Hind, K., Skilton, S. J., 1998. Phonetic structures of Scottish Gaelic. Journal of the International Phonetic Association 28 (1), 1–41.

Lewin, C., 2021. Continuity and hybridity in language revival: The case of Manx. Language in Society, 1–29.

Malécot, A., 1956. Acoustic cues for nasal consonants: An experimental study involving a tape-splicing technique. Language 32 (2), 274–284.

Malmi, A., Lippus, P., 2019. Keele asend eesti palatalisatsioonis / the position of the tongue in Estonian palatalization. Eesti ja Soome-Ugri Keeleteaduse Ajakiri 10 (1), 105–128.

Malmi, A., Lippus, P., Meister, E., 2022. Spectral and temporal properties of Estonian palatalization. Journal of the International Phonetic Association, 1–26.

Marin, S., Pouplier, M., Harrington, J., 2010. Acoustic consequences of articulatory variability during productions of /t/ and /k/ and its implications for speech error research. Journal of the Acoustical Society of America 127 (1), 445–461.

McGowan, R., 1992. Tongue-tip trills and vocal-tract wall compliance. Journal of the Acoustical Society of America 91 (5), 2903–2910.

McLeod, W., 2020. Gaelic in Scotland: Policies, movements and ideologies. Edinburgh University Press, Edinburgh.

Meister, E., Werner, S., 2015. Comparing palatography patterns of Estonian consonants over time. In: Proceedings of the 18th International Congress of the Phonetic Sciences. Glasgow.

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., 2021. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien. R package version 1.7-8. URL https://CRAN.R-project.org/package=e1071

Munro, G., Taylor, I., Armstrong, T., 2011. The state of Gaelic in Shawbost: Language attitudes and abilities in Shawbost. Bòrd na Gàidhlig, Inverness.

Nance, C., 2014. Phonetic variation in Scottish Gaelic laterals. Journal of Phonetics 47, 1–17.

Nance, C., Kirkham, S., 2020. The acoustics of three-way lateral and nasal palatalisation contrasts in Scottish Gaelic. Journal of the Acoustical Society of America 147 (4), 2858–2872.

Nance, C., Kirkham, S., forthcoming 2022. Phonetic typology and articulatory constraints: The realisation of secondary articulations in Scottish Gaelic rhotics. Language.

Nance, C., McLeod, W., O'Rourke, B., Dunmore, S., 2016. Identity, accent aim, and motivation in second language users: New Scottish Gaelic speakers' use of phonetic variation. Journal of Sociolinguistics 20 (2), 164–191.

Nance, C., Ó Maolalaigh, R., 2021. Scottish Gaelic. Journal of the International Phonetic Association 51 (2), 261–275.

Ní Chiosáin, M., Padgett, J., 2012. An acoustic and perceptual study of Connemara Irish palatalization. Journal of the International Phonetic Association 42 (2), 171 – 191.

Ní Chiosáin, M., Welby, P., Espesser, R., 2012. Is the syllabification of Irish a typological exception? Speech Communication 54, 68–91.

Nossair, Z. B., Zahorian, S. A., 1991. Dynamic spectral shape features as acoustic correlates for initial stop consonants. Journal of the Acoustical Society of America 89 (6), 2978–2991.

Ó Maolalaigh, R., 2008. The Scottishisation of Gaelic: A reassessment of the language and orthography of the Gaelic notes in the Book of Deer. In: Forsyth, K. (Ed.), Studies on the Book of Deer. Four Courts Press, Dublin, pp. 179–275.

Ohala, J., 1981. The listener as a source of sound change. In: Masek, C., Hendrick, R., Miller, M. F. (Eds.), Papers from the parasession on language and behaviour, University of Chicago. University of Chicago, Chicago, pp. 178–203.

Ohala, J., 1989. Sound change is drawn from a pool of synchronic variation. In: Breivik, L. E., Jahr, E. H. (Eds.), Language Change: Contributions to the study of its causes. Mouton de Gruyter, Berlin, pp. 173–198.

Ohala, J. J., 1990. The phonetics and phonology of aspects of assimilation. In: Kingston, J., Beckman, M. (Eds.), Papers in Laboratory Phonology I: Between the grammar and the physics of speech. Cambridge University Press, Cambridge, pp. 258–275.

Padgett, J., Ní Chiosáin, M., 2018. The perception of a secondary palatalization contrast: A preliminary comparison of Russian and Irish. In: Bennett, R., Angeles, A., Brasoveanu, A., Buckley, D., Kalivoda, N., Kawahara, S., McGuire, G., Padgett, J. (Eds.), Hana-bana: A Festschrift for Junko Itô and Armin Mester. Department of Linguistics, University of California, Santa Cruz, Santa Cruz, CA, pp. 1–14.

Picard, R. R., Cook, R. D., 1984. Cross-validation of regression models. Journal of the American Statistical Association 79 (387), 575–583.

Recasens, D., 2013. Coarticulation in Catalan dark [l] and the alveolar trill: General implications for sound change. Language and Speech 56 (1), 45–68.

Russell, P., 1995. Introduction to the Celtic languages. Longman, London.

Scottish Government, 2015. Scotland's Census 2011: Gaelic report (part 1). National Records of Scotland, Edinburgh.

Shaw, J. A., Kawahara, S., 2018. Assessing surface phonological specification through simulation and classification of phonetic trajectories. Phonology 35 (3), 481–522.

Shuken, C., 1980. An instrumental investigation of some Scottish Gaelic consonants. Ph.D. thesis, University of Edinburgh, Edinburgh.

Spinu, L., Kochetov, A., Lilley, J., 2018. Acoustic classification of Russian plain and palatalized sibilant fricatives: Spectral vs. cepstral measures. Speech Communication 100, 41–45.

Spinu, L., Lilley, J., 2016. A comparison of cepstral coefficients and spectral moments in the classification of Romanian fricatives. Journal of Phonetics 57, 40–58.

Spinu, L., Percival, M., Kochetov, A., 2019. Articulatory characteristics of secondary palatalization in Romanian fricatives. In: Proceedings of Interspeech 2019. Graz, Austria, pp. 3307–3311.

Spinu, L., Vogel, I., Bunnell, T., 2012. Palatalization in Romanian—Acoustic properties and perception. Journal of Phonetics 40, 54–66.

Sproat, R., Fujimura, O., 1993. Allophonic variation in English /l/ and its implications for phonetic implementation. Journal of Phonetics 21, 291–311.

Stevens, K., 1989. On the quantal nature of speech. Journal of Phonetics 17 (1-2), 3–45.

Stoll, T., 2017. Articulatory analysis of palatalised rhotics in Russian: Implications for sound change. Ph.D. thesis, Ludwig–Maximilians–Universität München, Munich.

Strycharczuk, P., Scobbie, J. M., 2017. Fronting of Southern British English high-back vowels in articulation and acoustics. Journal of the Acoustical Society of America 142 (1), 322–331.

Thurneysen, R., 1946. A grammar of Old Irish. Dublin Institute for Advanced Studies, Dublin.

Wang, J., Green, J. R., Samal, A., Yunusova, Y., 2013. Articulatory distinctiveness of vowels and consonants: A data-driven approach. Journal of Speech, Language, and Hearing Research 56, 1539–1551.

Watson, C. I., Harrington, J., 1999. Acoustic evidence for dynamic formant trajectories in Australian English vowels. Journal of the Acoustical Society of America 106 (1), 458–468.

Wright, R., 2004. A review of perceptual cues and cue robustness. In:

Hayes, B., Kirchner, R., Steriade, D. (Eds.), Phonetically-Based Phonology. Cambridge University Press, Cambridge, pp. 34–57.

Yanushevskaya, I., Bunčić, 2015. Russian. Journal of the International Phonetic Association 45 (2), 221–228.

Yu, K. M., 2017. The role of time in phonetic spaces: Temporal resolution in Cantonese tone perception. Journal of Phonetics 65, 126–144.

Gaelic	**Phoneme**	**Word position**	**Vowel context**	**English**
latha	l̪ˠ	initial	a	day
lùib	l̪ˠ	initial	u	bend
càl	l̪ˠ	final	a	cabbage
cùl	l̪ˠ	final	u	back
mo litir	l	initial	i	my letter
mo leannan	l	initial	a	my darling
air an latha	l	initial	a	on the day
ann an Liurbost	l	initial	u	in Leurbost
mil	l	final	i	honey
dil	l	final	i	gravel
fuil	l	final	u	blood
càil	l	final	a	anything
dàil	l	final	a	delay
sùil	l	final	a	eye
litir	l̪ʲ	initial	i	letter
linnean	l̪ʲ	initial	i	centuries
leabaidh	l̪ʲ	initial	a	bed
Liurbost	l̪ʲ	initial	u	Leurbost
till	l̪ʲ	final	i	return (verb)
caill	l̪ʲ	final	a	lose (verb)
saill	l̪ʲ	final	a	salt (verb)
puill	l̪ʲ	final	u	ponds
ùill	l̪ʲ	final	u	oil (verb)

Table 8: Lateral word list used in this study.

Gaelic	Phoneme	Word position	Vowel context	English
nathair	$\text{n̪}^{\text{ɣ}}$	initial	a	snake
nuadh	$\text{n̪}^{\text{ɣ}}$	initial	u	new
ceann	$\text{n̪}^{\text{ɣ}}$	final	a	head
sunn	$\text{n̪}^{\text{ɣ}}$	final	u	blast
mo nighean	n	initial	i	my daughter
mo nathair	n	initial	a	my snake
mo nupair	n	initial	u	my spanner
fìon	n	final	i	wine
glan	n	final	a	clean (verb)
dùn	n	final	u	fort
nighean	n̪^{j}	inital	i	daughter
neach	n̪^{j}	initial	a	person
niucleasach	n̪^{j}	initial	u	nuclear
cinn	n̪^{j}	final	i	heads
tàin	n̪^{j}	final	i	cattle
guin	n̪^{j}	final	i	arrow

Table 9: Nasal word list used in this study.

Gaelic	**Phoneme**	**Word position**	**Vowel context**	**English**
rionnag	$r^{ɣ}$	initial	i	star
rabaid	$r^{ɣ}$	initial	a	rabbit
rudan	$r^{ɣ}$	initial	u	things
piorr	$r^{ɣ}$	final	i	pierce
as fheàrr	$r^{ɣ}$	final	a	best
cùrr	$r^{ɣ}$	final	u	corner
mo rionnag	r	initial	i	my star
mo rabaid	r	initial	a	my rabbit
riubh	r	initial	u	to you
fìor	r	final	i	really
sìor	r	final	i	eternal
far	r	final	a	where
cur	r	final	u	put
ri	r^{j}	initial	i	to
fir	r^{j}	final	i	men
sir	r^{j}	final	i	ask
gàir	r^{j}	final	a	laugh
bàir	r^{j}	final	a	goal
muir	r^{j}	final	u	sea

Table 10: Rhotic word list used in this study.