Analysis of Linguistic Change in Glaswegian Dialect Using Chain Graph Models

C Alexander, L Evers, T Neocleous, J Stuart-Smith*

School of Mathematics and Statistics, University of Glasgow
School of Critical Studies, University of Glasgow*

Summary

When languages change, many different factors are involved, from group-level factors such as social background and gender, to individual speaker variability. Advanced statistical methods such as mixed effects modelling are typically used to explain language change, but the complexity of linguistic data poses a key problem: the output of such models is very difficult to interpret. This project offers a novel approach using graphical models as a visualisation tool to simplify such complex model output.

Data: Vowel Change in Glasgow

- Data are acoustic phonetic measures capturing vowel change in Glasgow over the course of the 20th century, e.g. vowel of *BOOT, GOAT, COT*, etc. (*Sounds of the City* [3] corpus).
- Acoustically, vowels are characterised by main resonances, called formants, which are measured in Hz.
- The first three vowel formants are the response variables for the model input (Figure 1).

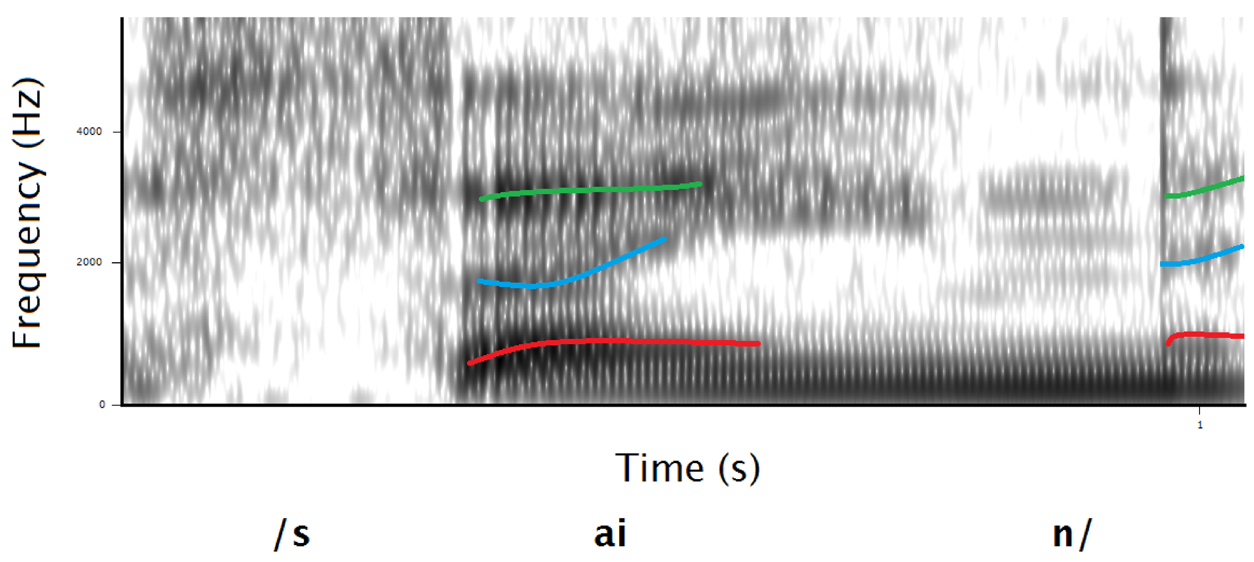

Figure 1: Spectrogram of the word *SIGN* spoken by a female Glaswegian speaker.The three coloured lines show the first three formants of the vowel /ai/.

- The explanatory variables are phonetic (**Preceding** and **Following** consonant adjacent to the vowel, e.g. sign, fine) and social (**Decade** of recording, **Age, Gender**).

Graphical Model Structure

- The graphical model is represented in the style of a chain graph model.
- An edge (line) with an arrow indicates that a factor is a significant predictor of a particular formant.
- Edges without arrows correspond to the relationships between formants, visualising their correlation structure.

To model the edges with arrows, a Bayesian mixed effects model [1] of the form:

$$p(\boldsymbol{y} \mid \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\delta}, \boldsymbol{\Omega}_\epsilon, \mathbf{X}) = \mathcal{N}\left(\boldsymbol{y} \mid \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\gamma} + \boldsymbol{\delta}, (\boldsymbol{\Omega}_\epsilon^{-1} \otimes \mathbf{I})\right),$$

where $\boldsymbol{\beta}$ are the coefficients for the explanatory variables, $\boldsymbol{\gamma}, \boldsymbol{\delta}$ are the speaker and word effect coefficients respectively and $\boldsymbol{\Omega}_\epsilon$ is the model error precision matrix.

- $\beta_i^j \neq 0 \rightarrow$ an edge is present between variable i and formant j.
- $\beta_i^j = 0 \rightarrow$ no edge is present between variable i and formant j.
- Using precision estimates obtained from the mixed effects model, the relationship between formants can be jointly inferred using a Bayesian Gaussian Graphical Model [2].

Given a graph G, a Gaussian graphical model is defined as

$$\mathcal{M}_G = \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}^{-1}).$$

- If the precision entry for formants p and q is zero, i.e. $\omega_{p,q} = 0$, then the formants are unconnected.

Results

Figure 2 shows the best posterior model selected for the *GOAT* vowel with posterior probability 0.414.

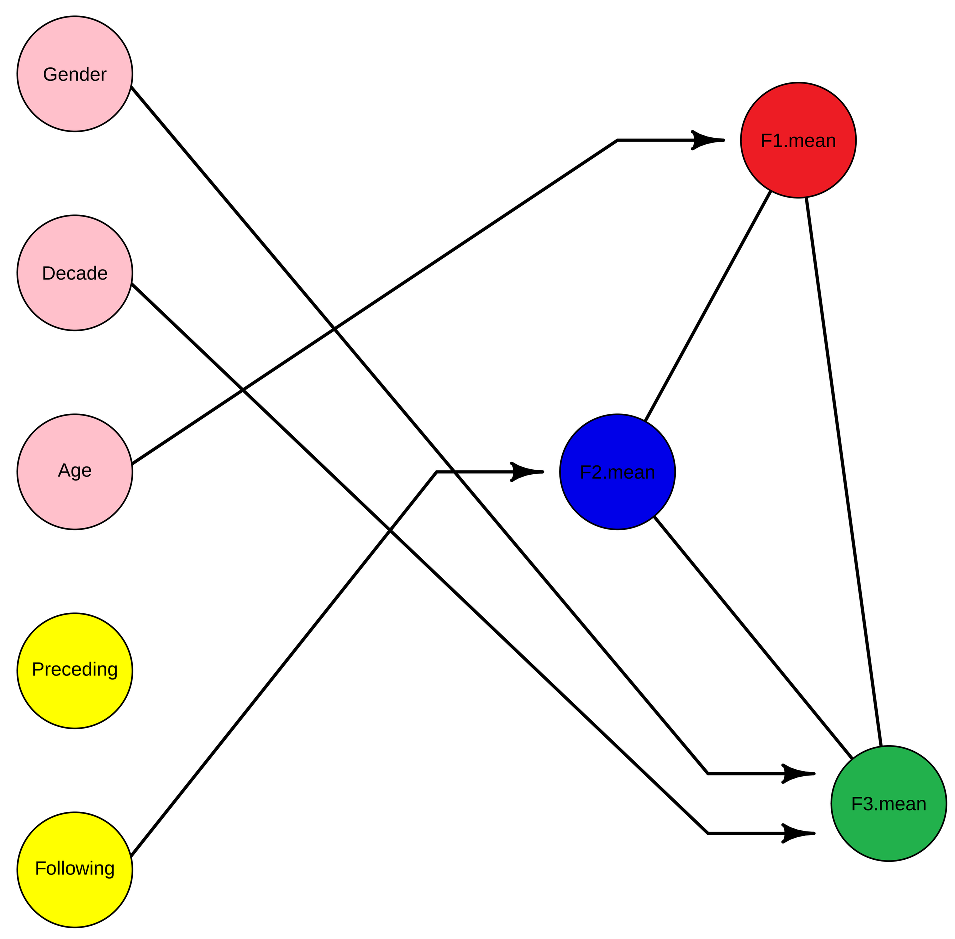

Figure 2: Best posterior model selected for the *goat* vowel.

This graphical model shows significant predictors:

- Age for F1, i.e. younger speakers are leading change in vowel quality.
- Following consonant for F2, i.e. expected effect of adjacent consonant on vowel quality.
- Decade and Gender for F3, i.e. evidence of vocal tract shape and some further indication of vowel change.

Implications

- This work extends beyond previous modelling of sociolinguistic data, which only modelled each formant individually, not accounting for the high correlation between formants.
- Modelling all formants at once is an important step towards accounting for the complexity of linguistic data and the factors constraining it.
- This also extends further to any discipline using acoustic phonetic analysis such as speech technology, forensic phonetics, clinical phonetics, which all currently model formants individually.
- The use of graphical models to represent complex model output can easily be extended beyond linguistic data. The model can be applied to any type of dataset with similar multivariate and nested features.

[1] Pinheiro, J. & Bates, D. (2000). Mixed effects models in S and S-PLUS. *Springer*. [2] Letac, G & Massam, H. Wishart distributions for decomposable graphs. *The Annals of Statistics*, 35:1278-1323. [3] Stuart-Smith, J,(2017). Changing sounds in a changing city. *E. Moore & C. Montgomery Language and a sense of place, CUP*