# Modelling the spread of sound change using dynamic parameters in an agent-based model.

Jonathan Harrington

Florian Schiel

INSTITUTE OF PHONETICS
AND SPEECH PROCESSING

# Objectives, background

The overall aim is to develop a computational (agent-based) model to begin to link the phonetic conditions that give rise to sound change and its propagation around a community

## Agent-based modelling and sound change

Agent-based modelling is a computational technique that can be used to understand **how local interaction between individuals** - represented by agents who are inter-connected in a network - **can bring about global (community) change***s* (Castellano et al. 2009)

## Two sets of data

/u/-fronting in Standard Southern British
/s/-retraction in Australian English consonant clusters.

# /u/-fronting and agent-based modelling

## The hypothesis to be tested with the agent based model

When speakers with fronted /ʉ/ and retracted /u/ come into contact with each other, then the shift is predicted to be asymmetric: the retractors should shift more towards the front than the other way round.

## Relationship to synchronic variation

This is because of a synchronic asymmetry: coarticulation is more likely to cause back vowels to front than front vowels to retract.

The general idea is that this synchronic bias can be stable, but can also be magnified (leading to change) when a group with predominantly front /ʉ/ contacts another with predominantly back /u/.

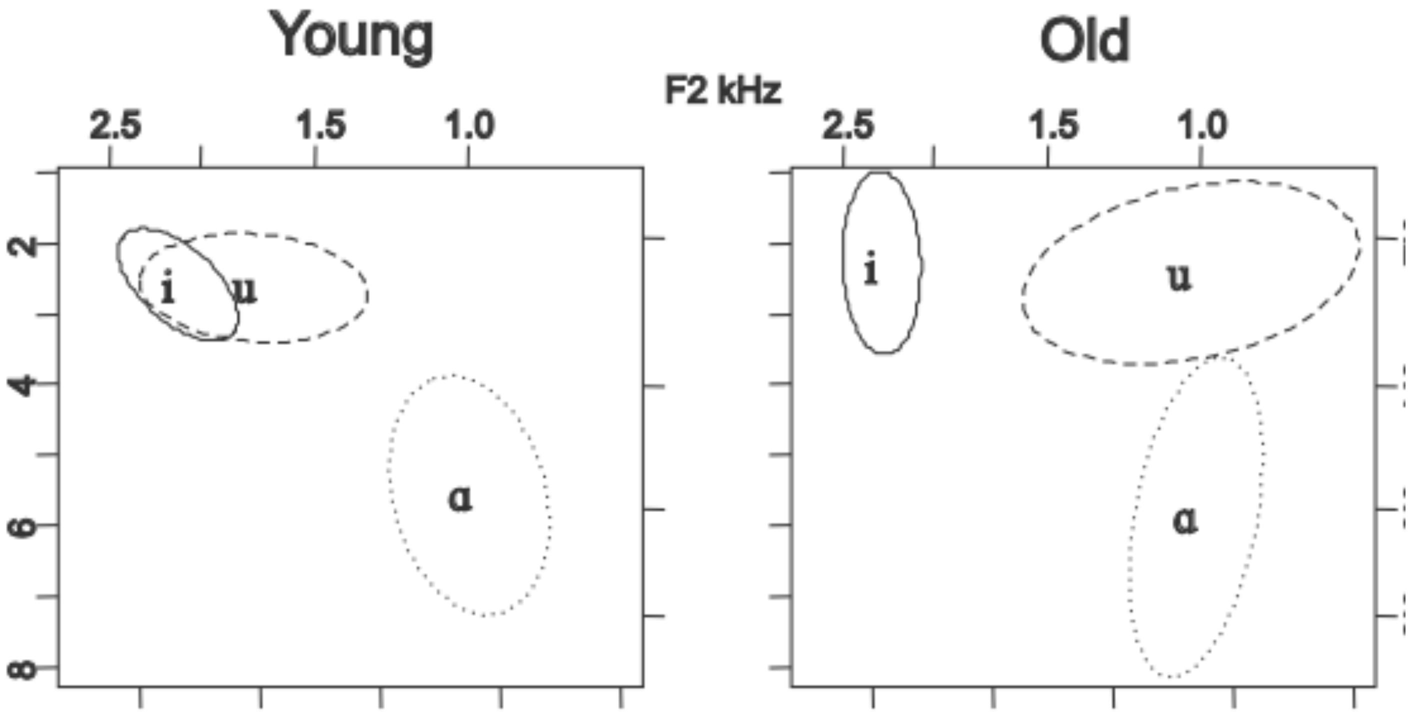
## /u/-fronting: speakers, materials

A  major different of the present ABM in contrast with other computational models of sound change is **that the input and output to the system is based on real speakers (one agent per speaker)**.

# /u/-fronting: speakers, materials

27 speakers[1], 14 old (age 69.2 years), 13 young (age 18.9 yrs) of Standard Southern British.

Each produced 10 repetitions of 54 isolated words. The ones analysed in the present study included the following sets of minimal pairs: 11 words × 10 repetitions = 110 word productions per speaker:




| | /f/ | /s/ | /k/ | /h/ |
|---|---|---|---|---|
| i | feed | seep | keyed | heed |
| ju | feud | | queued | hewed |
| u | food | soup | cooed | who'd |

1. Harrington, Kleber, Reubold (2008), *J. Acoust. Soc. America*

# Acoustic parameters

A second important  difference is that in contrast all other computational models of sound change, **our agents exchange dynamic signals**.

DCT-transformed F2-trajectories (which index phonetic fronting) between voicing onset and offset in /i, ju, u/ (e.g. *feed*, *feud*, *food*)

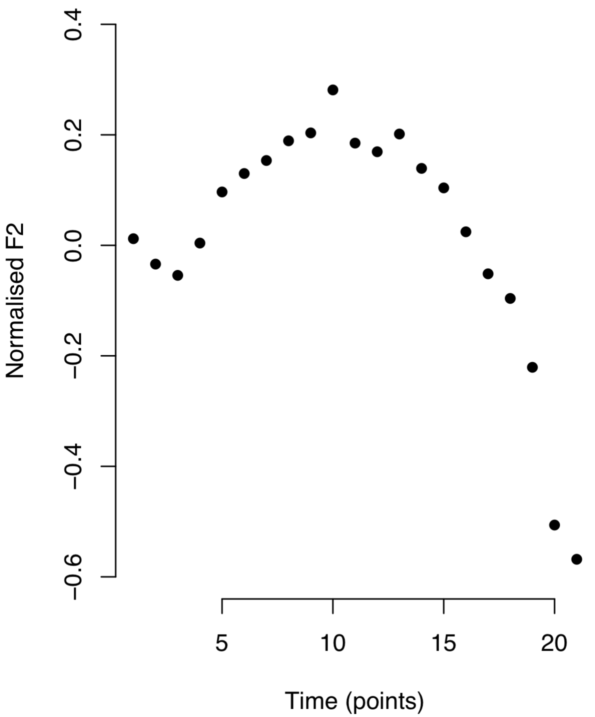
# Discrete cosine transformation

F2-trajectory

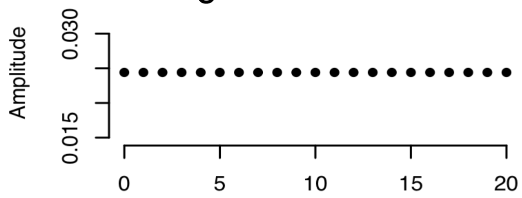
DCT-coefficients (transmitted by agents)

Amplitudes of ½ cycle cosine waves

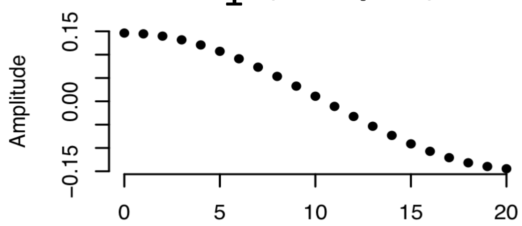
F2-reconstruction from DCT coefficeints



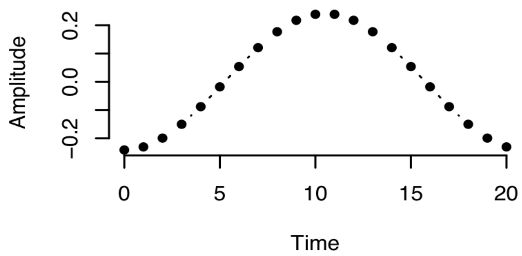

$C_0$ (mean)

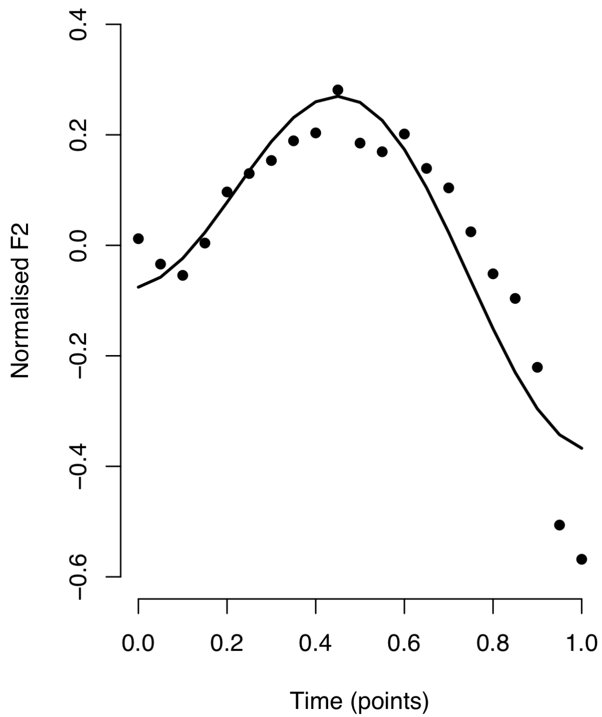



$C_1$ (slope)




$C_2$ (curvature)



# Agent-based model: general design considerations

## The lexicon, phonology, signals

Phonological classes are statistical (Gaussian) generalisations over acoustic signals

Word-specific model: each word is a statistical generalisation over its stored/remembered signals.

# Agent-talkers and listeners

## Interaction

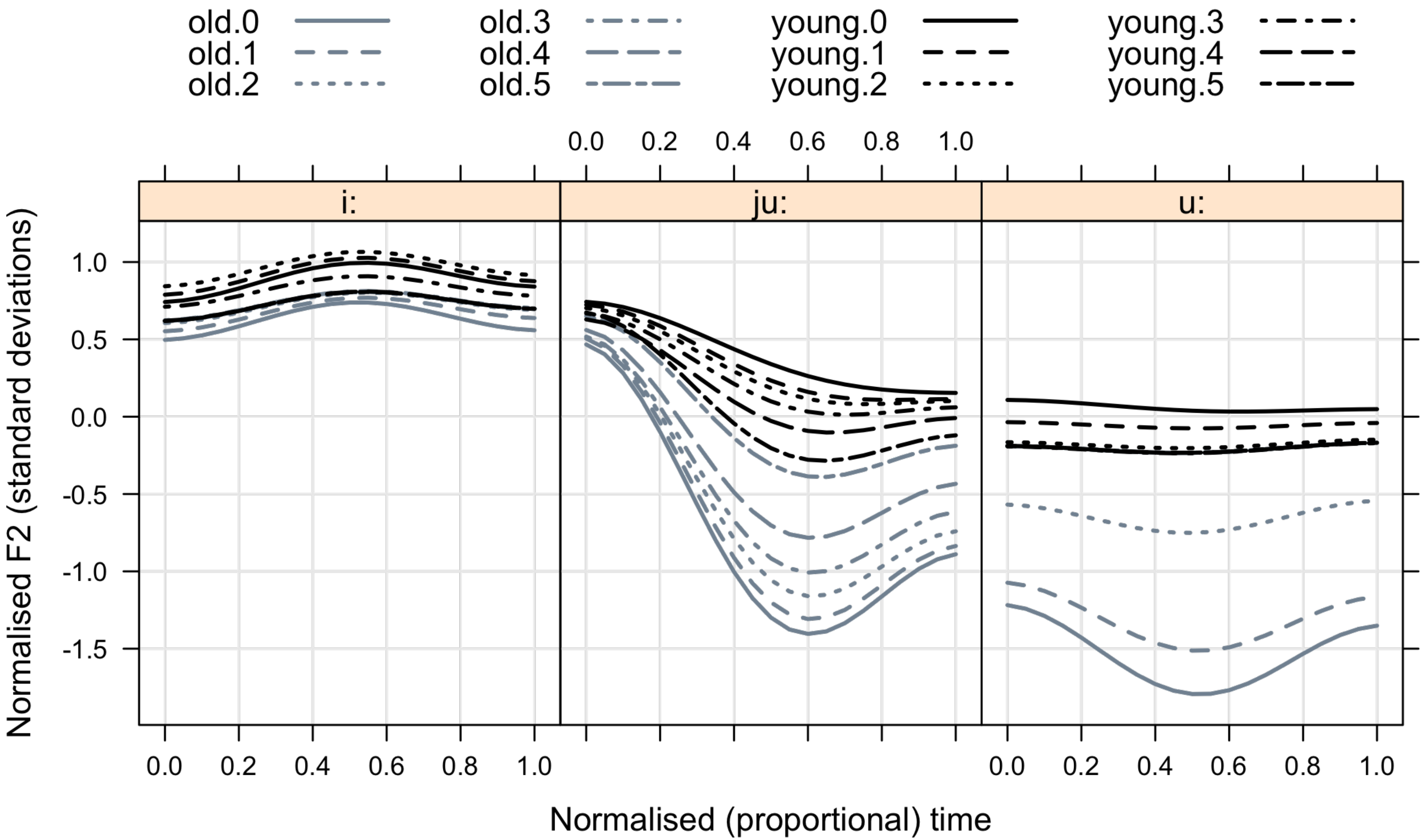
1**. randomly choose a pair of agents** that talk to each other

2. **The agent produces a token** (e.g. *few*): generate one sample from a 3D-Gaussian distribution in a DCT space formed from existing *few* items stored in memory.

3. **The listener absorbs the token into memory**, but only if it is more likely to belong to the listener's corresponding phonological category (i.e. to /ju/ in this example, rather than /i, u/).

4**. Listener memory loss** Remove the probabilistically most marginal *few* item (in one version) or the oldest *few* item, if the incoming token is memorised (to avoid indefinitely filling up memory with perceived tokens).

Results: older agents' /ju, u/ shifts more towards those of younger agents than the other way round



## Background to /s/-retraction

### English retraction before /Cr/

/s/ retraction in /ʃtr/ in many varieties of Am. English (e.g. Baker et al, 2011) incl. Philadelphia, Labov 2001; New York, Kraljic et al. 2008; southwest Louisiana, Rutter 2011. Also NZE: Warren, 1996; Standard British English: Cruttenden, 2014…[2]

### Data for the present study

22 speakers of Australian English who show varying degrees of /s/-retraction in /sCr/ contexts.

### Motivation for the present study

If these speakers – represented by agents – talk to each other will the tendency for /s/ to retract in /str/ be magnified through interaction, leading to its recategorisation as /ʃtr/?

## Agent-based modelling: speakers and agents

One agent per speaker. Each agent had initialised in memory:

- Words: **2 /s/**: *seam*, *sane*; **6 /str/** (*stream*, *strange*, *stretch*, *strand*, *stray*, *strap*; **2 /ʃ/** (*sheep*, *Shane*)

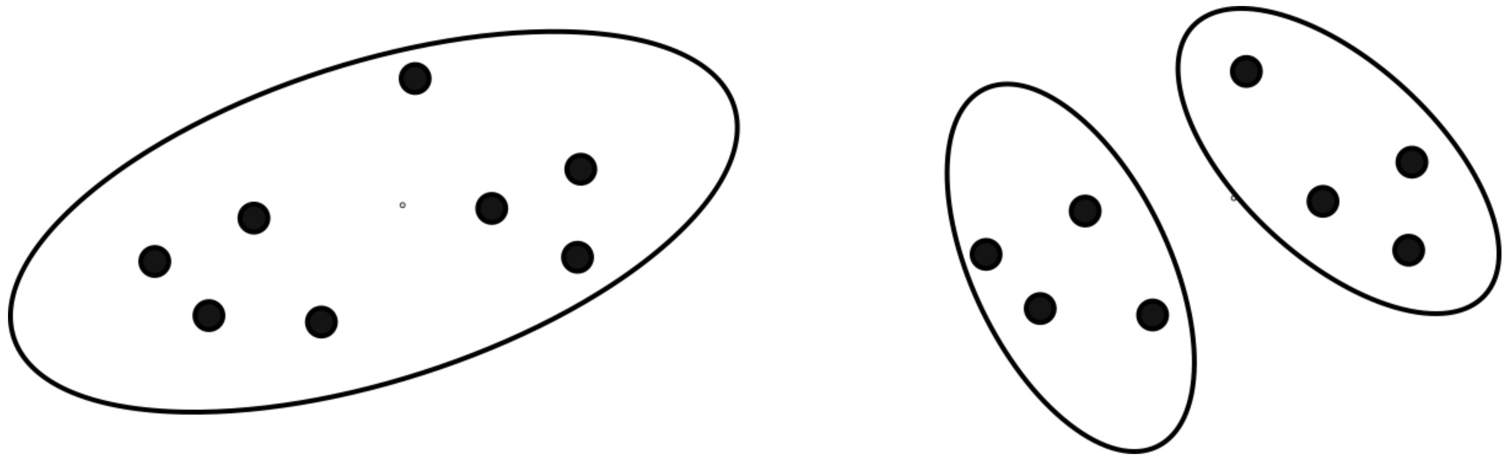
- A flexible **phonological association between words and signals** – to allow (re)-categorisation to be variable and listener-specific.

  e.g. Beddor (2009)[1]: some listeners categorise vowel-nasal sequences as VN, some as ṼN some maybe even as Ṽ.

- The possibility of splitting and merging phonological categories: each time an agent absorbed an item into memory, **each** category was tested for whether it should be split or merged.

1. Beddor (2009), *Language*

# Agent-based modelling: category splitting and merging

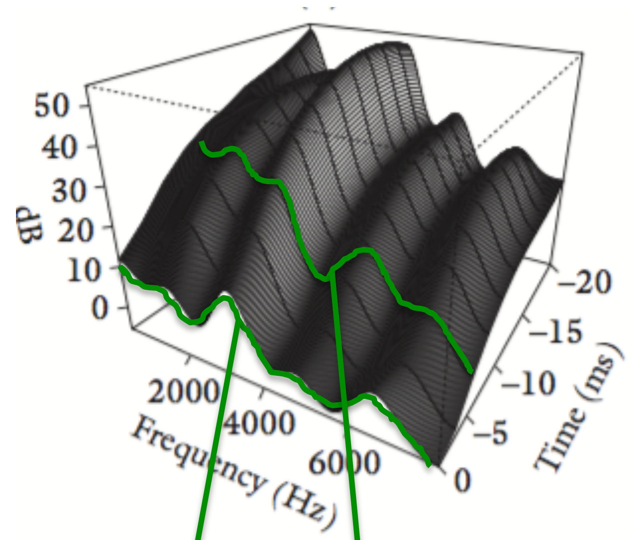


**Split**: apply *k*-means clustering to an existing category to split it into two maximally distinct categories and then determine whether the probability of category membership is greater for one category (left) or two (right).

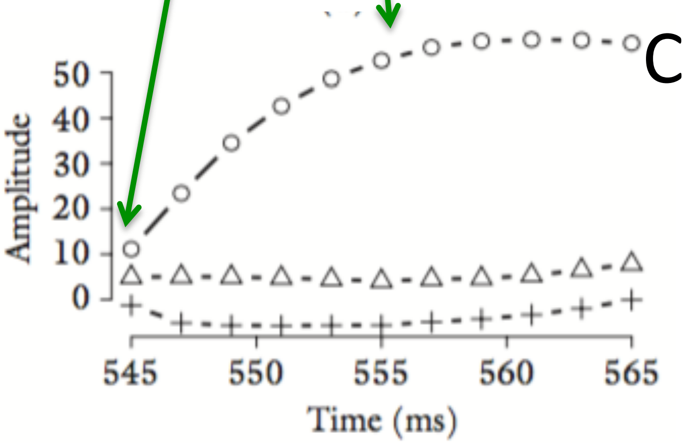
**Merge**: the reverse: test for all existing pairs of categories whether the probability of category membership is greater in single than in separate categories.

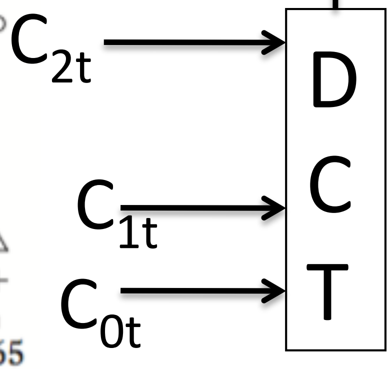
Acoustic parameters:
Transformed 2D DCT-space

**1.** Spectrum of a sibilant over time




**2.** DCT-triplet per spectral slice

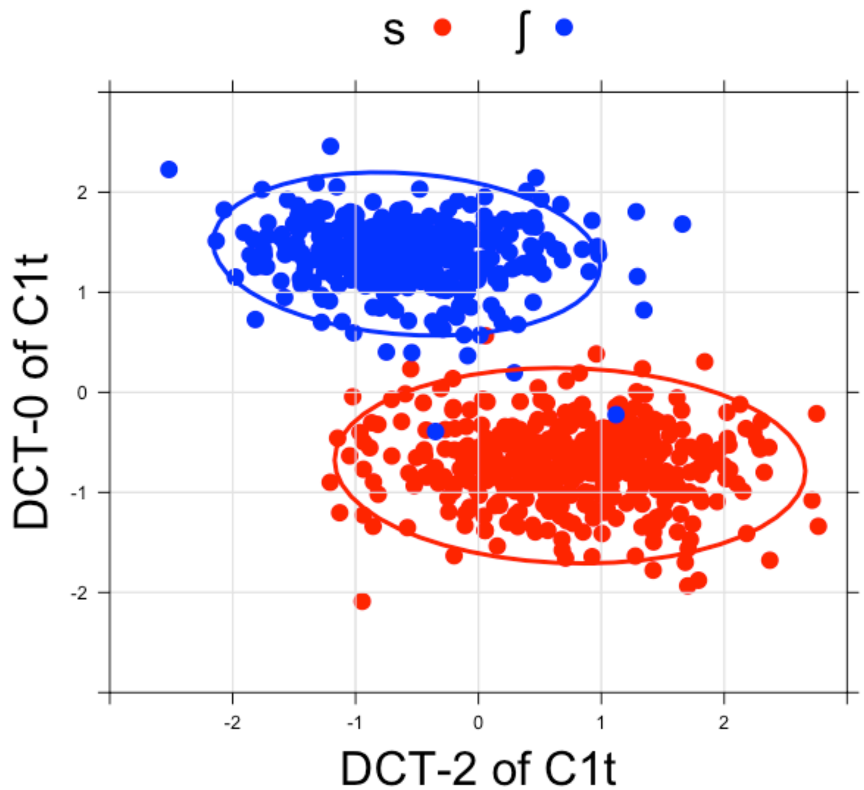



**3.** Mean, slope, curvature for each of $C_0$t, $C_1$t, $C_2$t (i.e. point in 9 dim. space)

**4.** Do 1-3. for all [s, ʃ] in –V contexts

**5.** Choose 2/9 params. that optimally separate [s, ʃ])



# Agent-based modelling: predictions

Following several thousands of interactions between pairs of agents:

## Phonological categories

Some agents will reassign /s/ in /str/ words to /ʃ/

## Acoustic signal

Some shift acoustically towards /ʃ/ for /str/ words

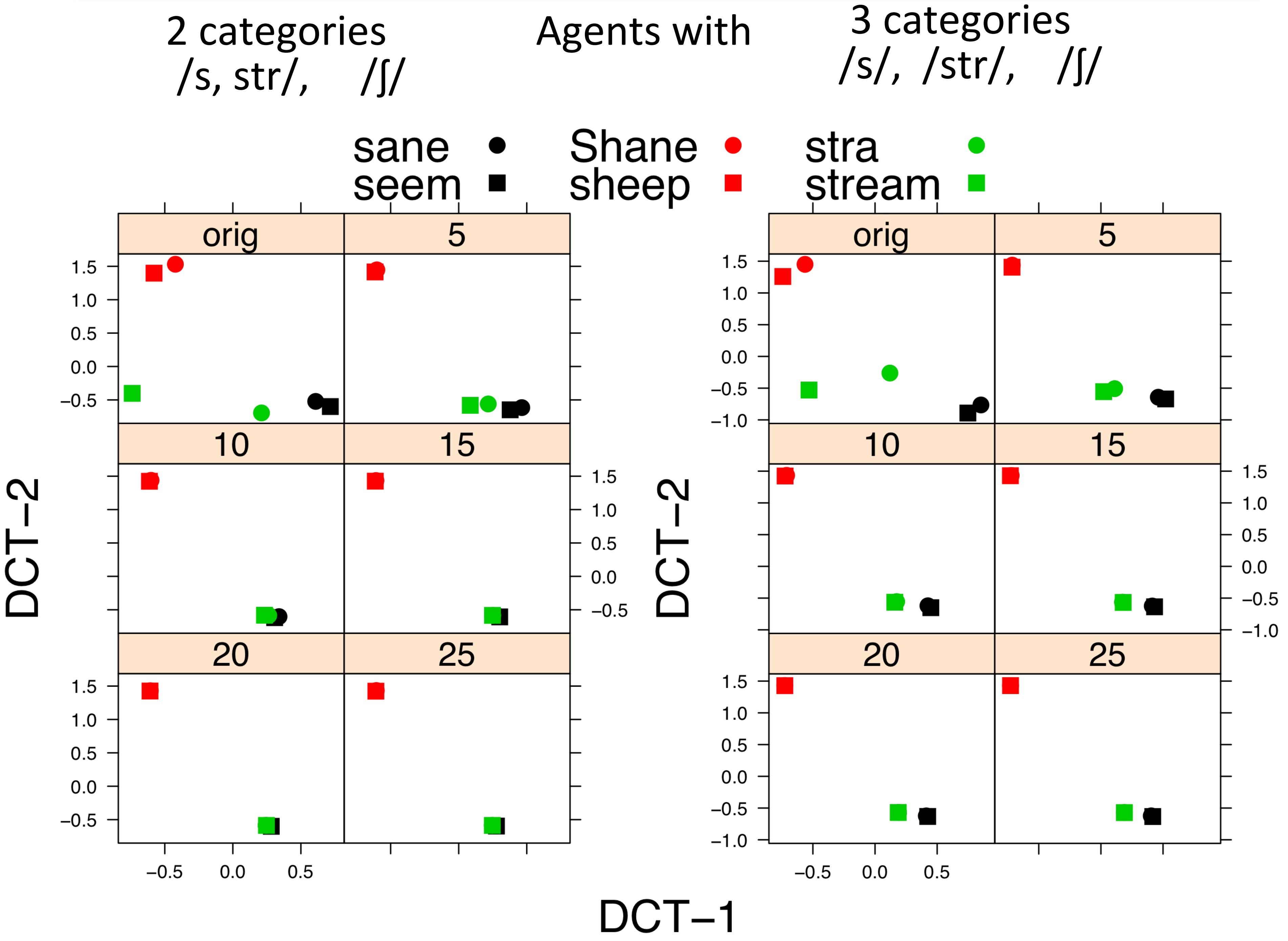
# Results: Phonological Categories

All agents begin with two categories: /s/ (containing /str/ words) and /ʃ/

After 25,000 interactions:

- 9 agents still have those two categories
- 10 agents have 3 categories: /$s_1$/, /$s_2$/, /ʃ/
  /$s_1$/ has all /str/ words, /$s_2$/ the other /s/-initial words
- 1 agent has 3 categories:
  - /s/
  - 1 category for all /str/ words except *stream*
  - 1 category containing /ʃ/ and *stream*.

# Results of simulations: No evidence that /str/-words shift to /ʃ/



# Agent-based modelling and /s/-retraction

There is some limited evidence for the expected phonological change: /str/ words break away from /s/ for some agents following interaction.

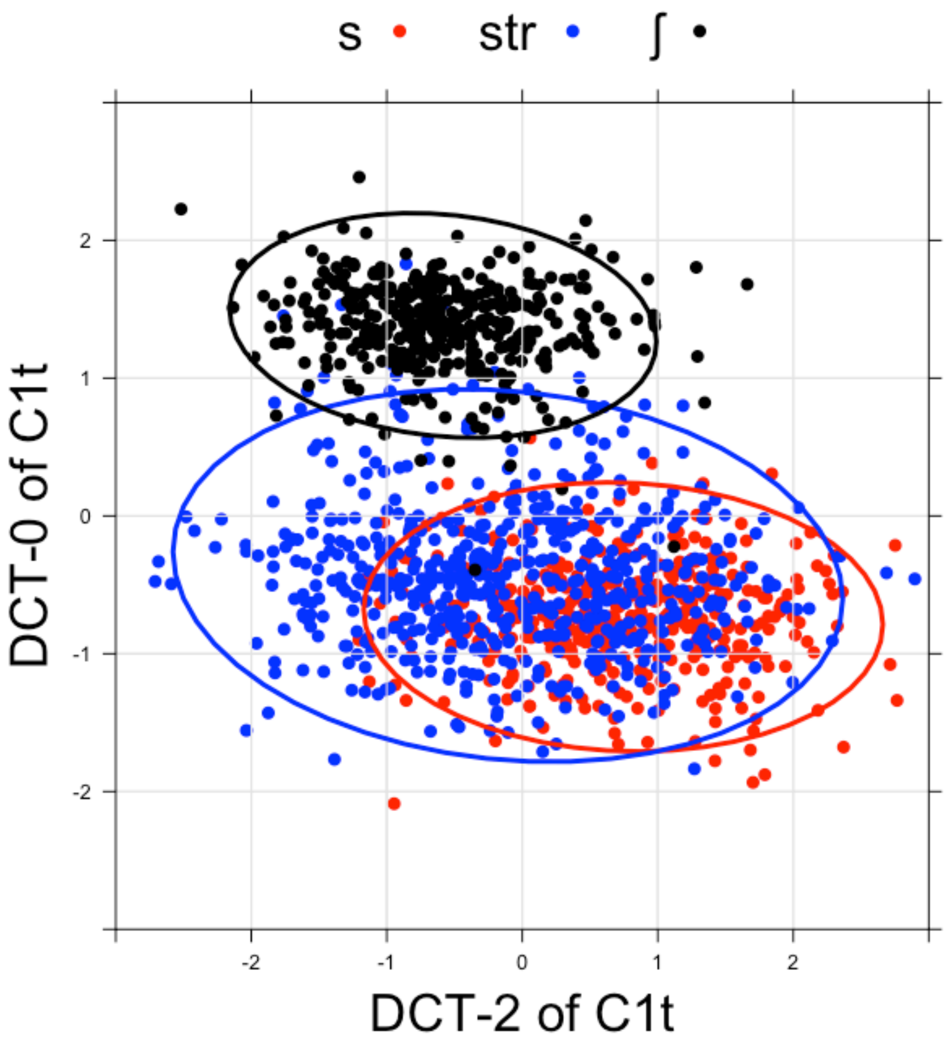
No evidence acoustically that /str/ words become more /ʃ/-like following interactions.

- Perhaps this largely negative result reflects the first stage of sound change following Ohala in which **there may be a listener recategorisation but no change to the acoustic signal**.

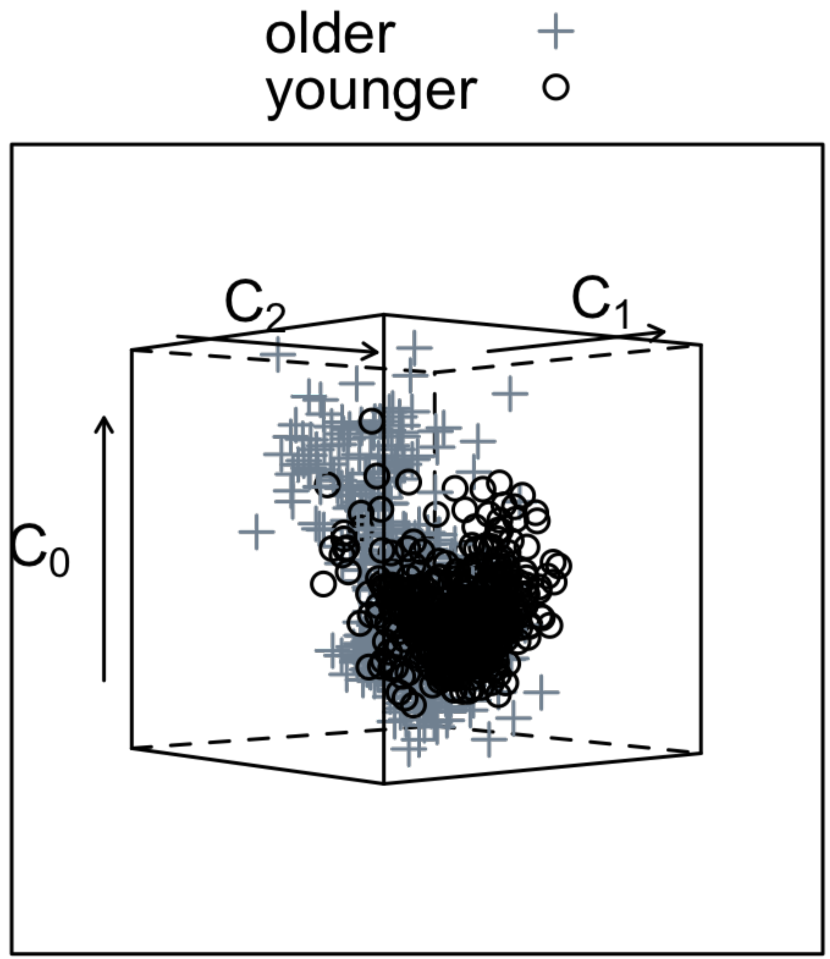
- Alternatively, it may be that the homogeneity of the speakers (all from the same town in Australia) and above all of the items (only a handful of /s, str, ʃ/ words) quash any less-well represented innovative changes.

# Agent-based modelling, sound change, direction of variation

Perhaps more importantly, the direction of variation in /str/ was **not** towards /ʃ/




Whereas the direction of variation of old group's retracted /u/ **was** towards young group's fronted /ʉ/



## Computational modelling: the origin and spread of sound change

Coarticulation, undershoot, and lenition exert biases in how sounds are distributed in a phonetic space.

Such synchronic biases can be magnified and lead to sound change through external group contact when the other group has a phonetic variant that is an extension of the naturally occurring phonetic bias.

Thus **sound change via group contact need not be socially driven** but may instead be shaped by **how phonetic variants from two groups are oriented with respect to each other in a phonetic space**.