What is an agent-based model and is it helpful in understanding sound change?

Jonathan Harrington

The IP model of sound change

The IPS has developed in the last few years a cognitively inspired, computational model sometimes known since [2] as the **I**nteractive **P**honetic (IP) model of sound change[1-10].

1. Harrington, J. & Schiel, F. (2017, *Language*)
2. Harrington, J., Kleber, F., Reubold, U., Schiel, F., & Stevens, M. (2018, *Top. Cog. Science*)
3. Stevens, M., Harrington, J., and Schiel, F. (2019, *Glossa*)
4. Harrington, J., Gubian, M., Stevens, M., and Schiel, F. (2019, *J. Acoustical Soc. America*).
5. Stevens, M. & Harrington, J. (2022, *Glossa*).
6. Cronenberg, Klingler, Kleber, Pucher (2022, *Interspeech*)
7. Gubian, M., Cronenberg, J., and Harrington, J. (2023, *Speech Communication*).
8. Cronenberg, J. (2024, doct. diss.)
9. Jochim, M. & Kleber, F. (2025, in press: in Rathcke & Kleber eds.)
10. Kapia, L. & Riverin-Coutlée, J. (in press, in Omari A., A, Sinani, L. Halala eds.)

Media attention due to [4] – see here and especially here

Recent review by Beddor (2023, section 5.2)

The IP model of sound change

has, as one of its main purposes, to understand how the **phonetic origin** and **phonetic spread** of sound change are connected.

Phonetic origin

The phonetic conditions likely to lead to sound change.

Spread of sound change

How such conditions come to 'infect' an entire speech community over time.

The IP model of sound change: foundations

Phonetics and sound change

1. sound change is directional (Garrett & Johnson, 2013)

2. there is a link between the **direction of phonetic variation** and sound change (Harrington & Schiel, 2017).

Spoken language interaction

3. Perceived speech influences speech production (e.g., Pierrehumbert, 2003) through **selective memorization**.

4. Selective memorization can under certain circumstances bring about sound change over time.

1. Sound change is directional

X > Y does not imply Y > X with equal frequency[1]

s-retraction (but rarely ʃ-fronting)

k > tʃ, but never tʃ > k

u-fronting (but rarely i-backing)

Directional sound change has a synchronic, phonetic origin:

s-retraction[2]

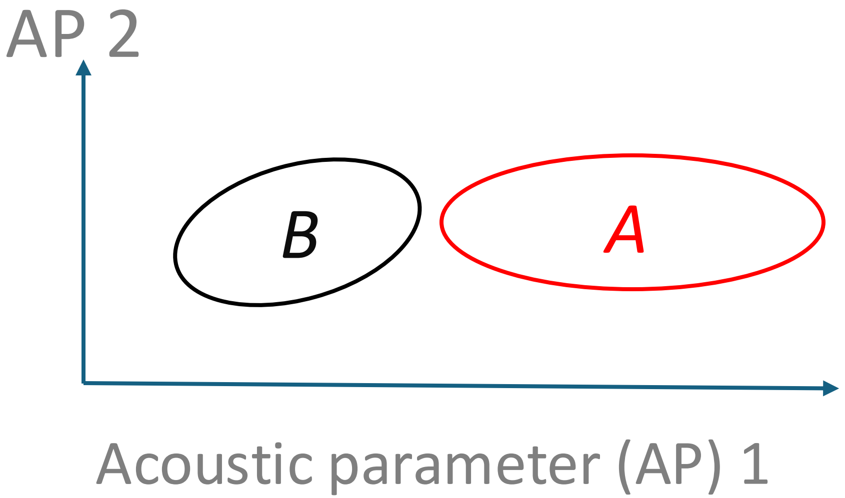
ki ⟶ ti perceptual confusions[3.]

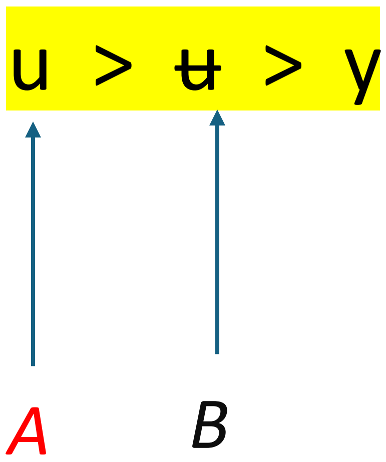
u-fronting (various)

1. garrett2011.pdf. 2. Stevens et al, 2015; Stevens & Harrington, 2016. 3. guion1998.phonetica.pdf; ohala1992.pdf

2. The direction of phonetic variation and sound change

According to the IP-model of sound change[1], phonetic variation is **directed** along a possible **path of sound change.**

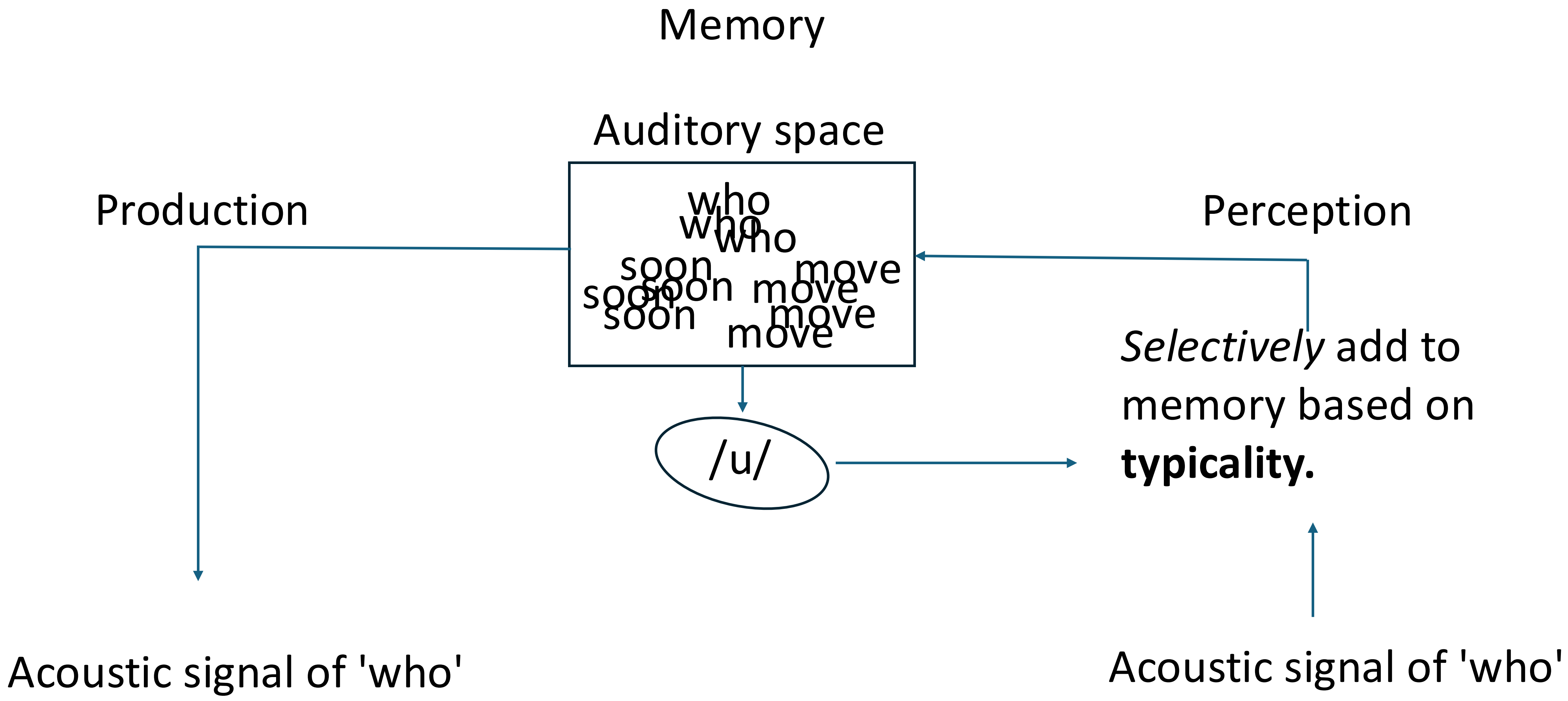




A and *B* are two speakers (or speaker groups) producing retracted (A) and fronted (B) variants of /u/ at **earlier** and **later** stages of a sound change.

1. Harrington, Kleber, Reubold, Schiel, Stevens (2018).

3. Perception influences production: selective memorization



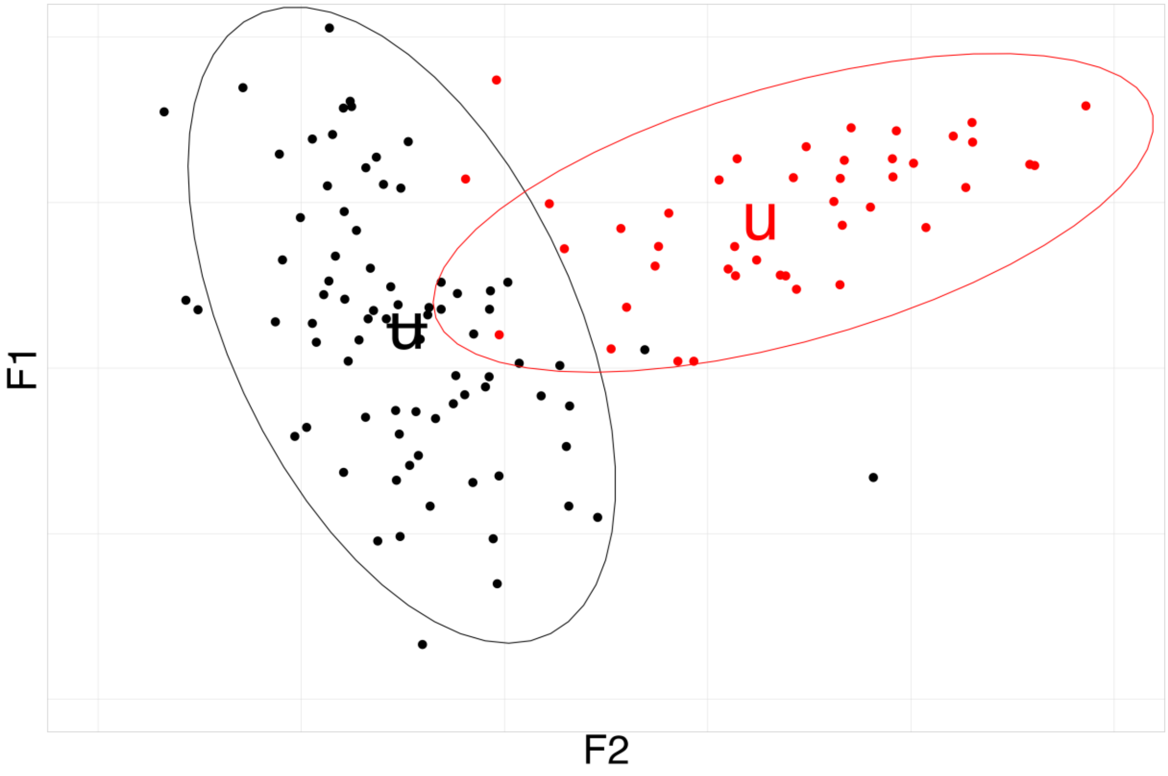
1. Todd, Pierrehumbert, and Hay (2019, *Cognition*)

3. Selective memorization through typicality

Innovative individual: fronted [ʉ]

Conservative individual: retracted [u]

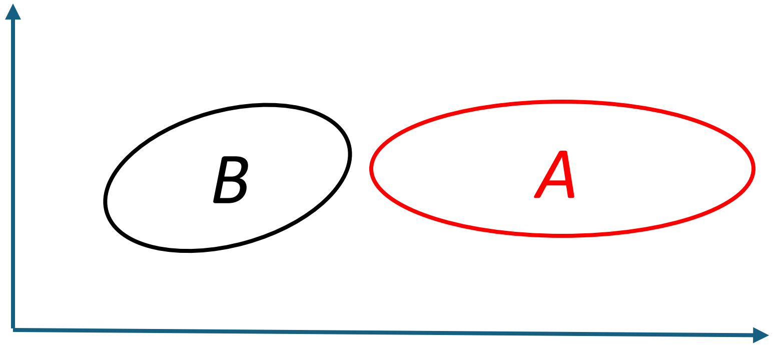
Any points within the confidence ellipses are **typical** and can be **memorized**.

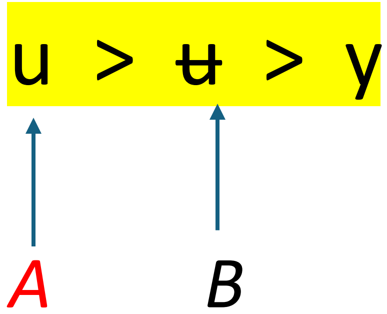


More [ʉ] tokens (black points inside the red ellipse) are memorized by the conservative individual because they are typical of [u].

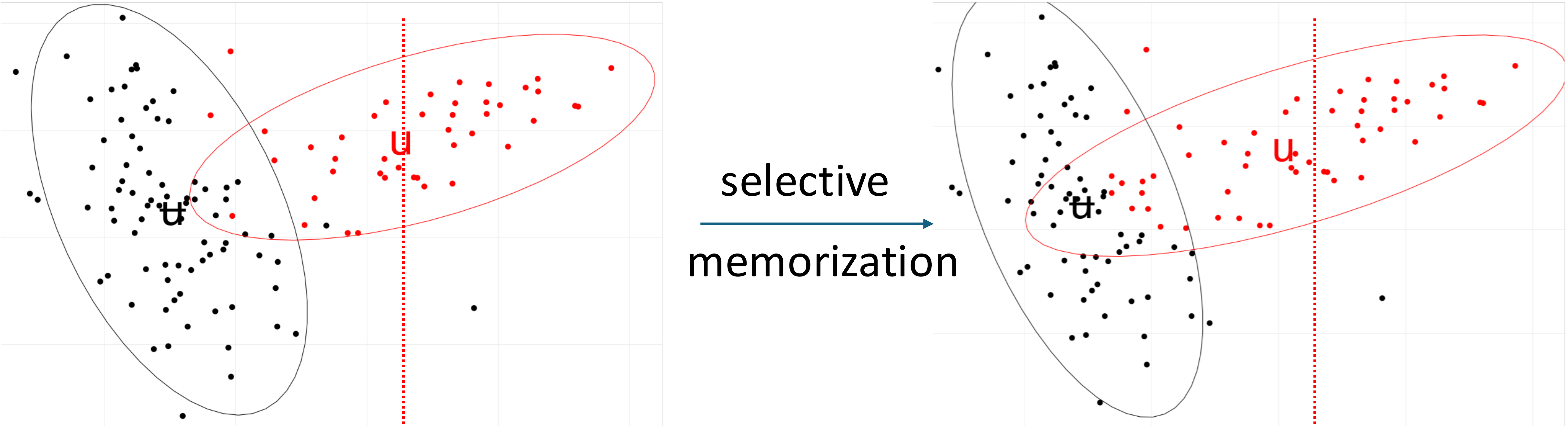
4. Selective memorization and sound change over time.

If:





then *A* shifts more towards *B* than the other way round (= sound change)



/u/-fronting and agent-based modelling

Sound change path

[u] > [ʉ] > [y]

A B

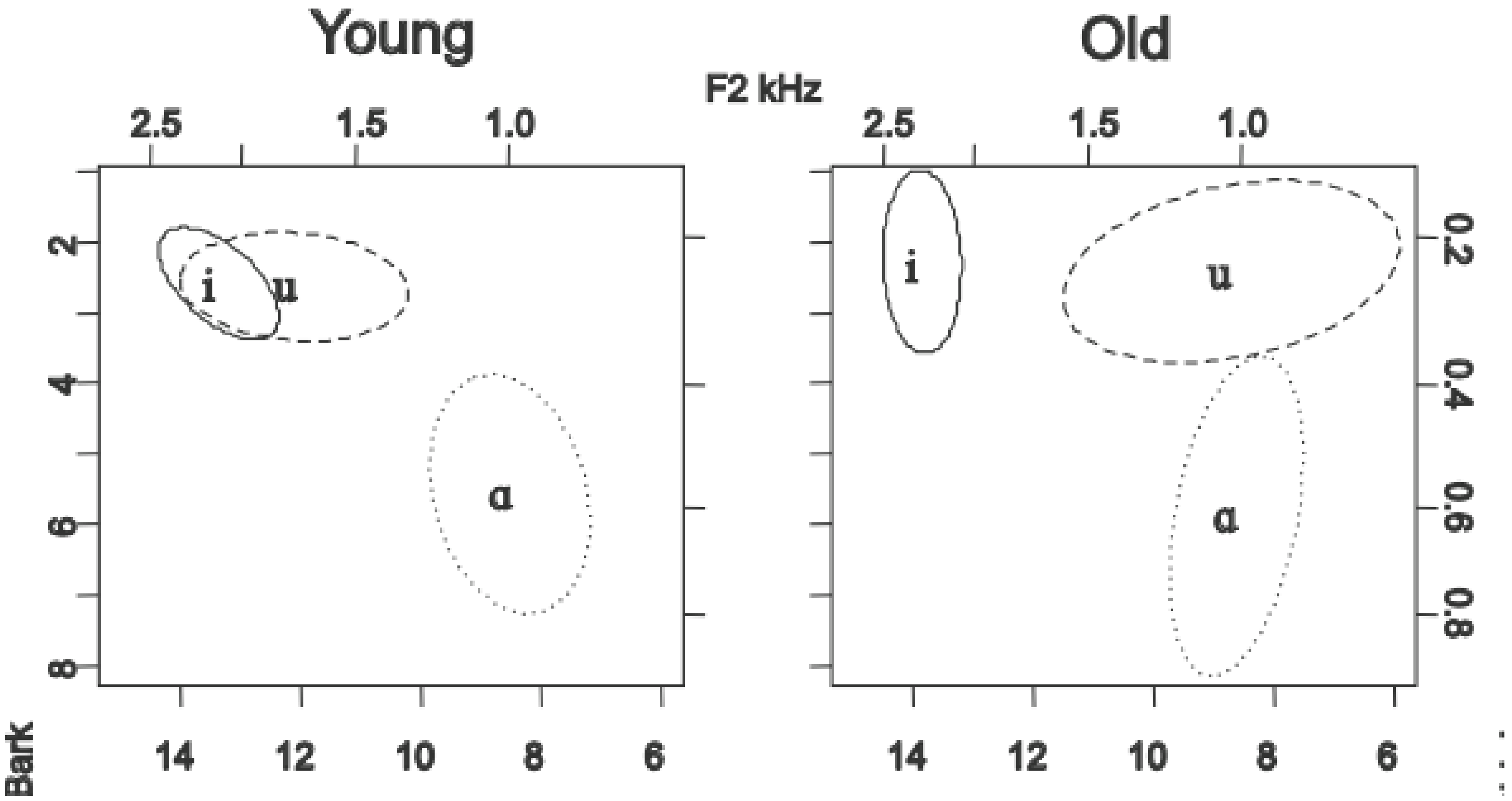
A: older speakers with retracted [u]

B: younger speakers with fronted [ʉ]

Hypothesis of IP-model: When *A* and *B* speak to each other, then *A* should shift more towards *B* than *B* towards *A*.

Data for the agent-based model

was taken from Harrington, Kleber, Reubold (2008)



Speakers and materials

27 speakers 14 old (age 69.2 Jahre), 13 young (age 18.9 yrs)

They produced 10 repetitions of 54 isolated words.

For the agent-based model, 11 minimal-pair words × 10 repetitions = 110 productions/speaker were used.

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

Acoustic parameters: DCT

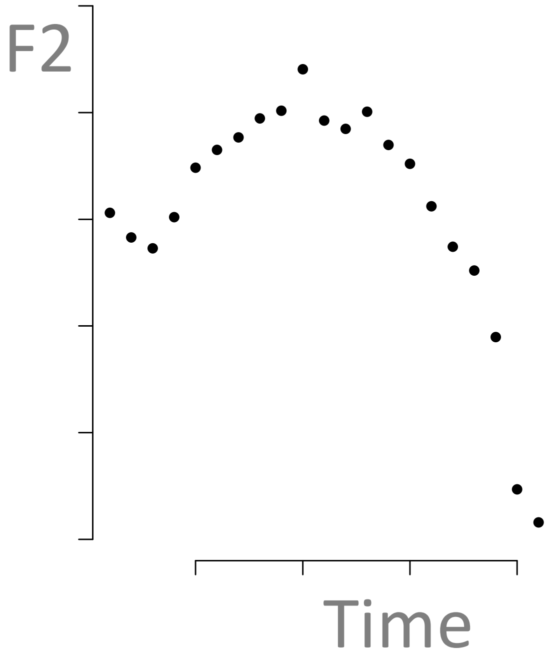
F2 trajectories were calculated between acoustic onset and offset

The discrete cosine transformation[1,2] (DCT) was applied to each trajectory to reduce it to a point in a 3D-space.

1. Watson & Harrington (1999, *J.Acoustic.Soc.America*)

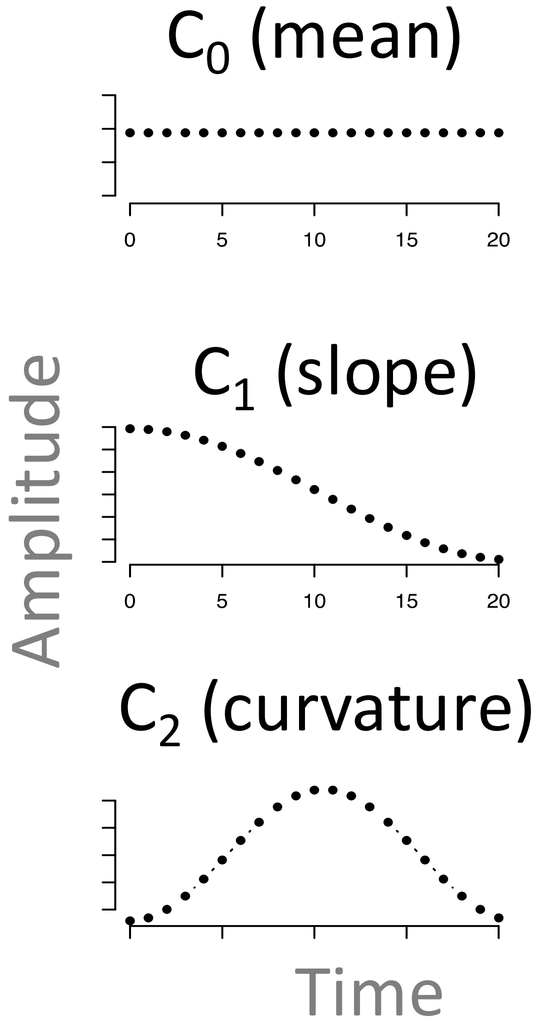
2. See examples in Emu-R

F2-trajectory

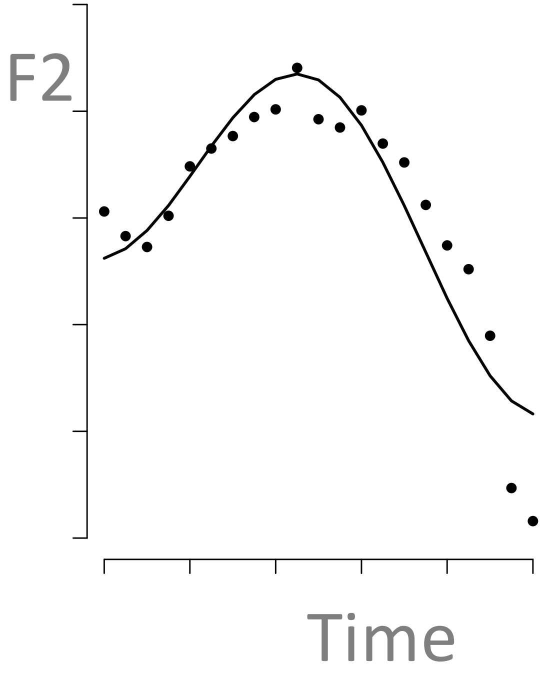


DCT-coefficients[1]

= Amplitudes of ½ cycle cosine waves



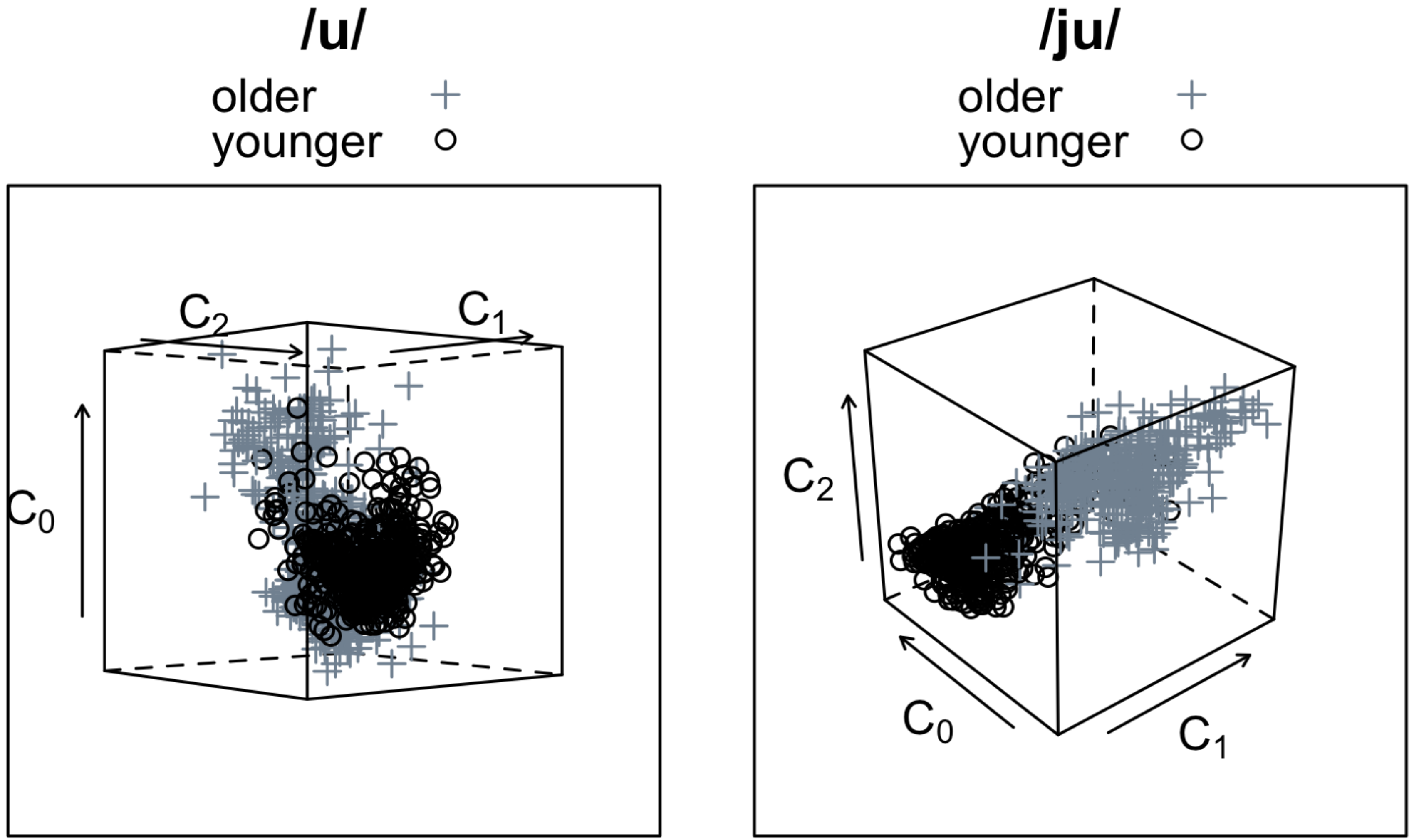
F2-reconstruction from DCT coefficients



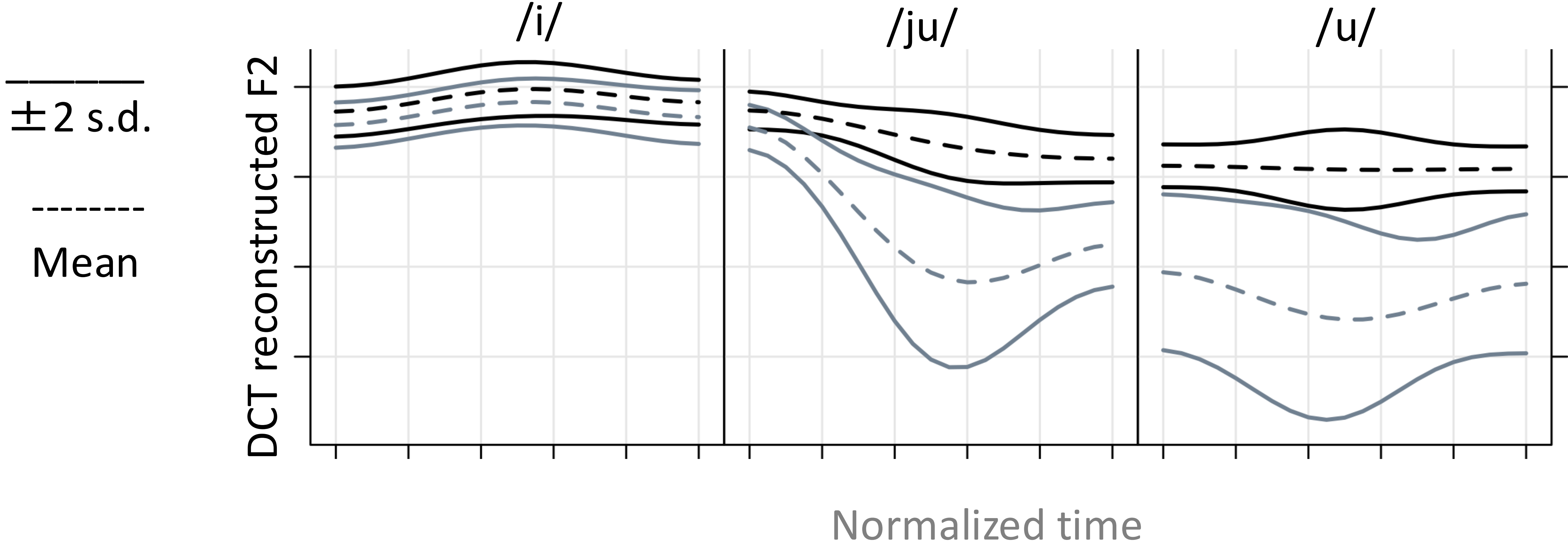
1. See examples in Emu-R

Phonetic variation of old and young

In this 3dim-DCT space, old is directed more towards young than young to old



Phonetic variation of old and young

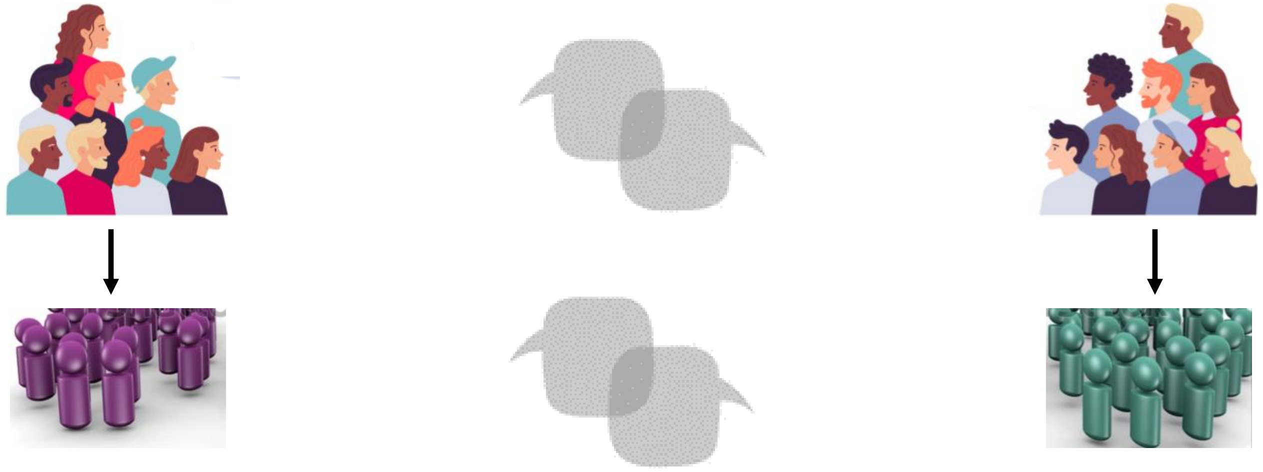


The IP agent-based computational model of sound change

Available as R package soundChangeR (Johanna Cronenberg)

https://github.com/IPS-LMU/soundChangeR

One agent per real talker



Agents can exchange dynamic signals (e.g., a DCT parameterization of an entire formant trajectory).

Agent-based model: Initialisation

22 agents, 11 old and 11 young (one agent per speaker)

Prior to interaction, each agent had stored in memory:

Word classes (11 *soup*, *seep*, *food*…)

Phonological classes: /i, ju, u/

C_0, C_1, C_2 DCT coefficients

Typically 110 *objects* per agent

An object had five components: e.g. {seep, /i/, 3 DCT coeffs.}

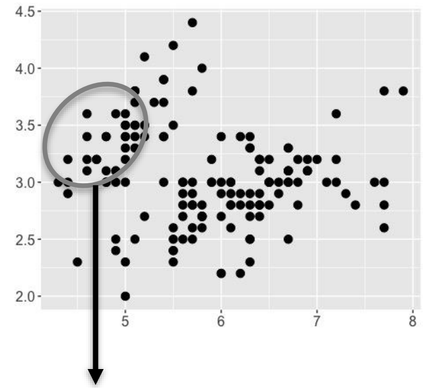
Agent talker ← randomly paired → Agent listener

Agent listener's decision

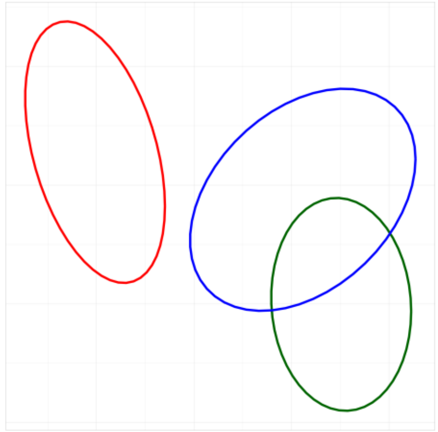
Gaussian model over memorized exemplars of /ju/, /i/, /u/.

1. Random selection of word e.g. *food*

food



3. Add x as an exemplar of *food* to memory, based on typicality and discriminability to listener's /u/



2. Generate one DCT-triplet x, from a Gaussian model of /u/ formed from all *food* tokens of that speaker

x

4. If added, randomly remove a *food* token from memory

For 60,000 interactions

Discriminability

Only memorize a token *x* of a phonological class *P* as long as it is not probabilistically closer to any other phonological class.

Discriminability

has **no influence** on the memorization of phonetic variants of the **same** phonological class (e.g. [ʉ] and [u] of food or of feud)

preserves the distinction between **different** phonological classes: that is, it prevents /i, ju, u/ (e.g., *feed*, *feud*, *food*) from merging i.e., collapsing, into a single phonological class).

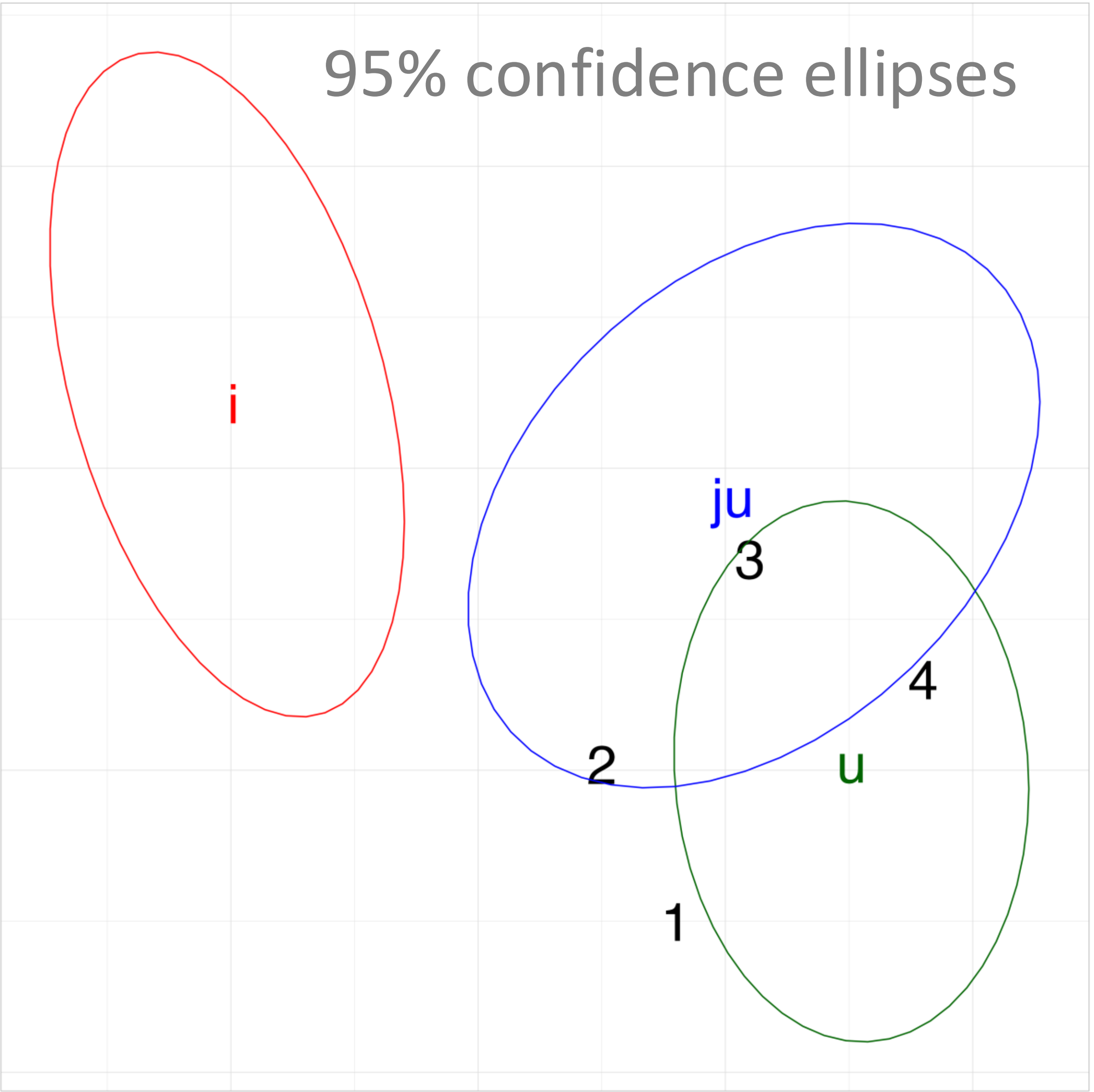
Discriminability and typicality

1-4 are four tokens of an agent speaker's productions of *food*.

The ellipses are those of an agent listener's **phonological** classes.

	1	2	3	4
typical	-	-	+	+
discriminable	+	-	-	+

Only 4 is memorized by the agent listener.



Models that were run and predictions

Three ABMs were run consisting of:

(a) 11 older speakers only

(b) 11 younger speakers only

(c) all 22 speakers

Predictions

(a, b): no change in /ju/ or /u/ because the groups are homogeneous (old: all back [u]; young: all front [ʉ]).

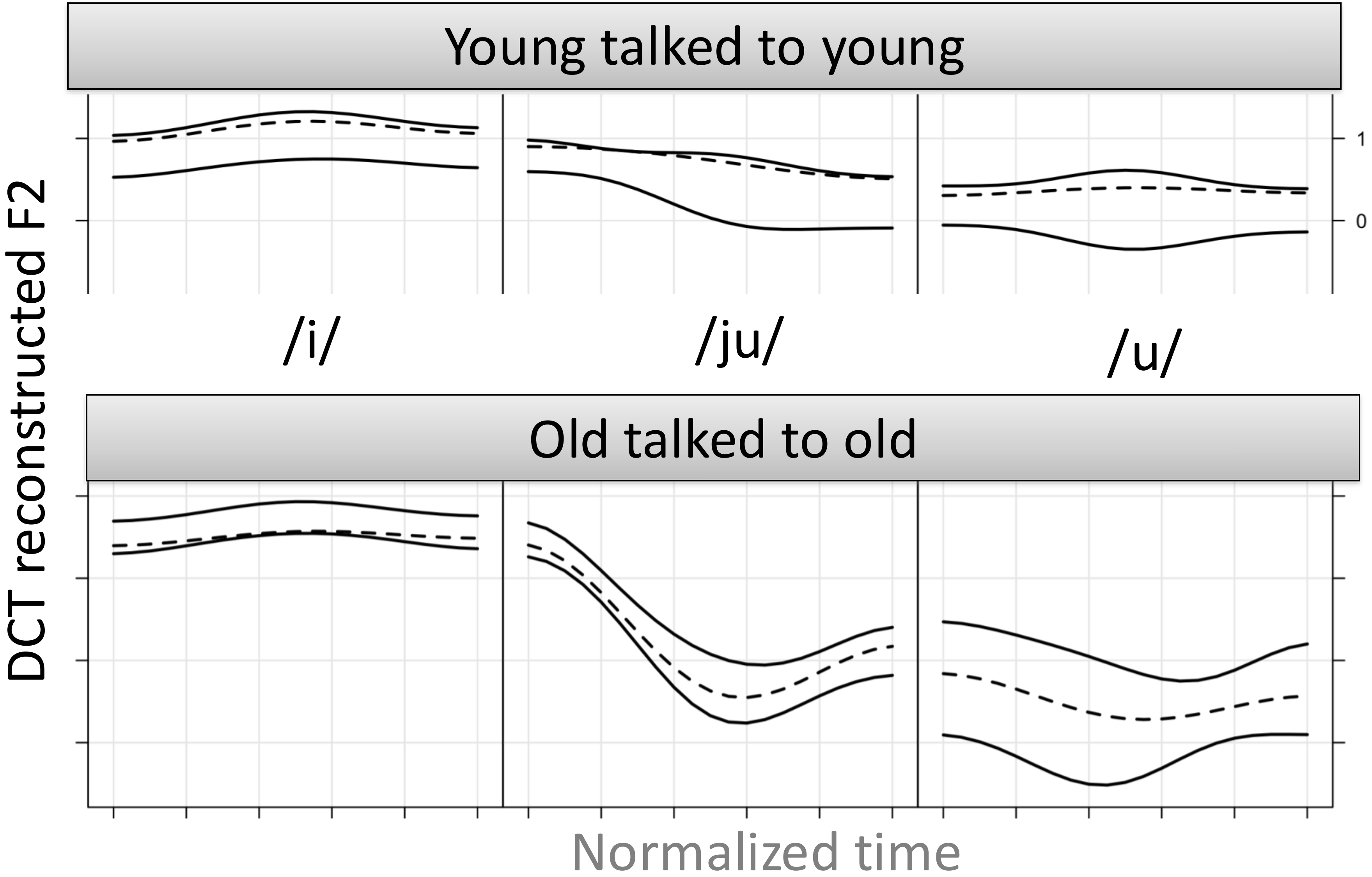
(c): /ju, u/ of old should shift towards young.

(a, b, c): no change in /i/ for which there is hardly any speaker variation.

No change within same speaker groups

____
±2 s.d. of **original** (prior to interaction).

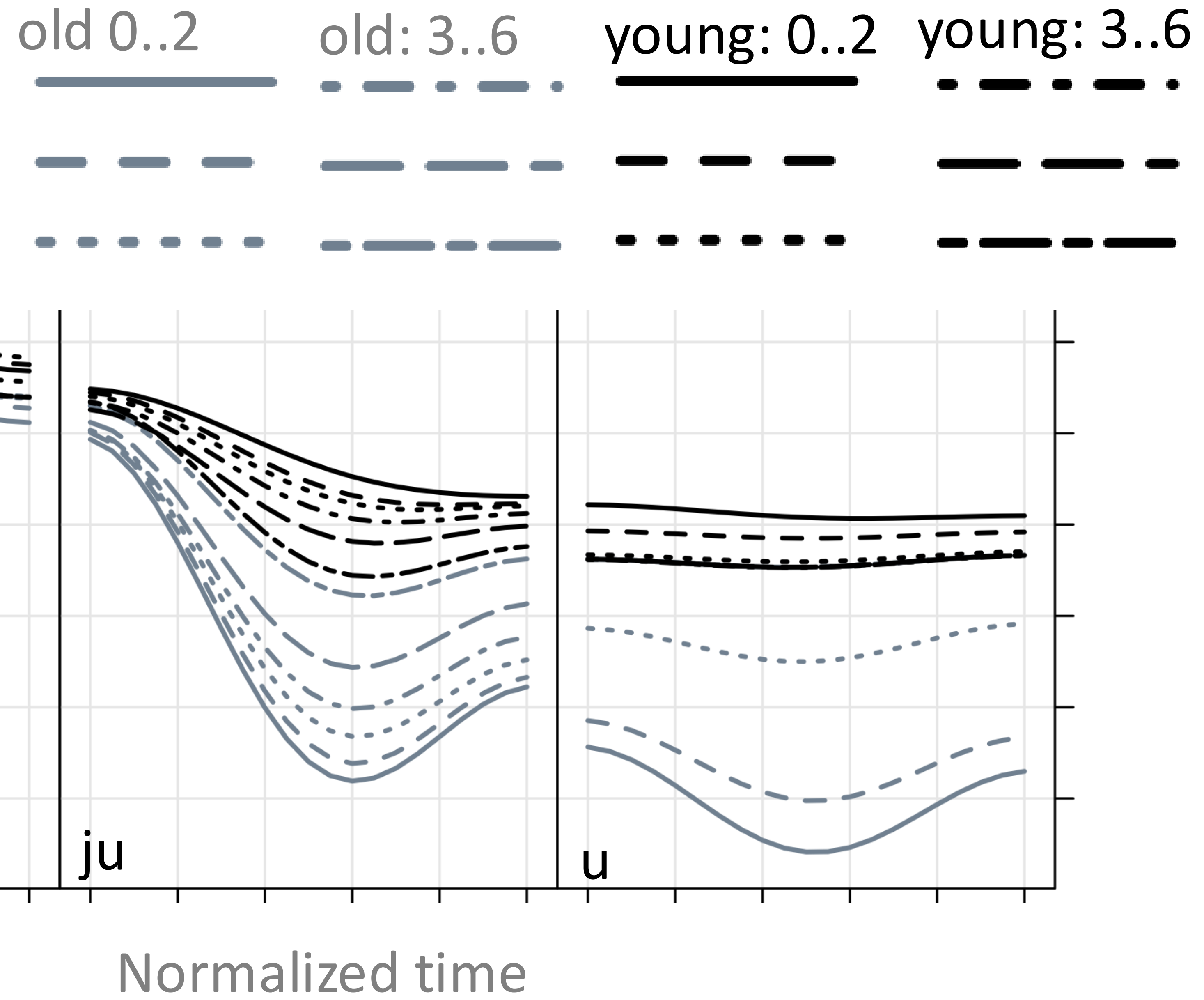
Mean after 30,000 interactions



Both groups interact: old → young in /ju, u/

0 = baseline (at start)

1..6 after 10,000 – 60,000 interactions.



Conclusions and outlook

How sound change gets going: there is interaction between (groups of) individuals who are advanced to different degrees along a path of sound change combined with selective memorization.

The model cannot explain how some individuals get to be more 'innovative' than others in the first place.

Sound change is not inevitable if there is relative phonetic homogeneity between individuals.

The model in Harrington & Schiel (2017) cannot handle phonetic mergers (neutralization) because of discriminability – but see [2,3,7] page 2.