INVESTIGATING SOUND CHANGE THROUGH COMPUTATIONAL AGENT-BASED MODELLING: AN R PACKAGE

Johanna Cronenberg, Michele Gubian, Jonathan Harrington, Vanessa Reichel

Institute for Phonetics and Speech Processing (IPS), LMU Munich
johanna.cronenberg | m.gubian | jmh | vanessa.reichel@phonetik.uni-muenchen.de

ABSTRACT

We introduce the R package `soundChangeR`: an implementation of a computational, agent-based model of sound change. This model rests upon ideas from exemplar theory about how the memorisation of traces of speech can result in phonetic imitation, perceptual learning, and eventually phonetic and phonological change. In theoretical views of the exemplar-based relation between speech production and perception, it has so far remained unclear whether some stored exemplars may contribute more to a listener-turned-speaker's future production than others. Using the cognitively-inspired architecture of `soundChangeR` we show that it makes a difference whether the agents preferably memorise typical and/or unambiguous exemplars. Memorising only exemplars that are typical members of their phonological class can reinforce phonetic biases, whereas memorising only unambiguous exemplars creates a repulsion between phonological classes. We argue that using such computational models can lead to new theoretical insights and we encourage readers to try `soundChangeR` on their own data.

Keywords: sound change, agent-based modelling, exemplar theory, selective memorisation

1. INTRODUCTION

Decades of research have established that sound changes are rooted in phonetic variation [1] and propagated through interactions between individuals [2, 3]. However, the exact interplay between the intra- and extralinguistic factors that may contribute to sound change at any stage of the process remains poorly understood. This is because it is impossible to know beforehand whether a sound change is going to take place, so as soon as the change is underway, it is too late to capture the circumstances which may have triggered it. Computational agent-based models (ABMs) offer an artificial world and a controlled environment which can be used to explore the cognitive, social, and linguistic factors that might play a role in the emergence and spread of a specific change that has already been observed empirically e.g. by means of apparent-time data.

The ABM presented in this paper, called `soundChangeR`, aims to link theories about the origin and spread of sound change by means of a cognitively-inspired architecture. More specifically, `soundChangeR` rests upon ideas from exemplar and usage-based models of language [4, 5, 6] about how the memorisation of traces of speech (exemplars) can result in phonetic imitation [7], perceptual learning [8], and, eventually, sound change. Therefore, and in line with other computational models of sound change [9, 10, 11, 12], agents in `soundChangeR` are the computational analogue of human speaker-listeners, i.e. they are equipped with mechanisms in order to produce and perceive exemplars of speech sounds as well as a memory that connects the two processes. In contrast to other ABMs of sound change, however, `soundChangeR` is publicly available and can be used on real speech production data to test whether the model predicts any acoustic or phonological changes. Section 2 explains the details of the agents' production-perception feedback loop and briefly mentions how the R package `soundChangeR` can be installed and used.

Exemplar-based approaches have been particularly useful in explaining phonetic shifts, such as /u/-fronting in many English varieties [13], the Northern Cities Shift [10], or the New Zealand English short front vowel shift [14]. These kinds of sound changes are hypothesised to result from updates of memorised acoustic distributions caused by consistent exposure to biased variants. However, exemplar models often remain quite vague about the relation between speech perception and speech production despite its significance in sound changes. That is, are there memorised exemplars that affect future productions more than others? For instance, do typicality and/or ambiguity in the association between exemplars and phonological classes increase the likelihood of updating learned acoustic distributions and, consequently, of shifting future productions? The exemplar-based architecture of `soundChangeR` allows us to computationally test these assumptions about the probabilistic selection of exemplars and its relation to phonetic shifts. Three simulations on artificial data will be presented to show that it makes a difference whether agents preferably memorise typical and/or unambiguous exemplars.

2. METHOD

2.1. Model Architecture

In `soundChangeR`, each agent is typically initialised with the speech production data from a real human speaker. This data should consist of relevant acoustic parameters for the sound(s) under investigation as well as the word type in which the sound was produced (e.g. formant values for the vowel in *food*). Together, the acoustic values and the lexical entry form an exemplar which is stored in the agent's memory. Before the first interaction, and then regularly throughout the simulation, each agent derives sub-phonemic classes from the stored exemplars by means of two unsupervised machine learning algorithms. The ABM follows exemplar theory's bottom-up approach to phonological knowledge: that is, sub-phonemes (SPs) in `soundChangeR` are abstractions over clouds of stored exemplars that can be re-computed when new exemplars have been memorised [15]. The resulting classes are agent-specific and can encode both classical phonological [16] as well as sub- or quasi-phonemic knowledge [17, 18].

In an interaction, a pair of agents is chosen either randomly or based on their agent group. These groups can be defined by the user and contain any sociolinguistic information about the speakers that may play a role in the sound change (e.g. age, gender, or regional origin). The agent-speaker randomly chooses a word type from its lexicon, estimates a (Gaussian) distribution based on all memorised exemplars associated with that word type, and samples a new exemplar from that distribution. This word-based sampling procedure ensures that possible coarticulatory effects (e.g. that the vowel in *soup* is more fronted than that in *food*) are also present in the newly produced exemplar which, together with the corresponding word type, is transmitted to the agent-listener.

The task in speech perception is to decide whether or not to memorise the exemplar (but not recognising or categorising that exemplar). That is, we make explicit assumptions about which exemplars are more likely to impact the agent-listener's internal acoustic distributions and thus, their speech production in future interactions. The decision to memorise a perceived exemplar can be based on two criteria [9] pertaining to the sub-phoneme, SP, that is associated with the word type transmitted by the agent-speaker. The absolute criterion determines whether the Mahalanobis distance between the exemplar and the centroid of SP is below a threshold defined by the user. The reasoning behind this criterion is that atypical exemplars should not influence the listener-turned-speaker's future productions [19, 20].

The relative criterion determines whether the exemplar's posterior probability conditioned on SP is higher than the posterior probabilities conditioned on all other sub-phonemes in the agent-listener's memory. That is, only exemplars that are discriminable [9] (i.e. cannot be confused for a member of another sub-phoneme) are memorised. Since this criterion penalises acoustic ambiguity, it creates a repulsion between sub-phonemic classes, following the idea that phonological contrasts are usually maintained [21, 22]. However, given that sub-phonemes are regularly updated during a simulation, this repulsion does not preclude sub-phonemic classes from merging [23].

The two memorisation criteria can also be combined, i.e. the perceived exemplar is only memorised if it passes both criteria. Finally, if the agent-listener has memorised the perceived exemplar, it forgets (i.e. deletes) a random exemplar of the same word type. The process of forgetting exemplars ensures that the memory size does not grow to such an extent that newly memorised exemplars have no impact.

Simulations are run until there is no more observable acoustic change across the population. An analysis can then be undertaken of whether there is any evidence of change either in the acoustic makeup of the simulated sounds (in the case of a phonetic shift like e.g. /u/-fronting) or in how words are mapped to sub-phonemic classes (in the case of e.g. a split or a merger).

2.2. Usage

The R package `soundChangeR` can be downloaded from GitHub (see installation instructions: https://github.com/IPS-LMU/soundChangeR) and comes with full documentation in the form of help pages for all functions and a detailed vignette. The core function is `run_simulation()` which takes a number of arguments by means of which the user can adapt the simulation parameters. Importantly, users can submit their own speech production data in any simple table format (e.g. `.txt`, `.csv`). The dataset minimally needs to contain speaker IDs (each speaker is going to be represented by one agent), numeric-continuous acoustic features of the sounds under investigation, and the word types in which the sounds were uttered. An example of a dataset that meets these requirements is `u_fronting` which is available to the user once `soundChangeR` is loaded in R. It contains DCT-parametrised F2 values for the vowels /i, u, ju/ produced by 22 Standard Southern British English speakers [13].

2.3. Testing the Effects of Exemplar Selection

Three simulations were run with `soundChangeR` in order to demonstrate how the application of the memorisation criteria (absolute only, relative only, or both) can result both in phonetic shifts and stability [12]. Figure 1 shows the artificial dataset that was generated for this purpose. It consists of two agents, A and B, both of which are initialised with 50 exemplars each of 10 word types in a two-dimensional acoustic space. For agent A, sub-phoneme 1 (SP1, solid ellipse) is elongated and overlaps with sub-phoneme 2 (SP2, dashed ellipse) at one end along its main axis. Agent A's SP2, in turn, is similar to agent B's SP2. For agent B, both sub-phonemic classes are are compact and nearly spherical.

Figure 1: State of agent A's and B's representations of sub-phonemic classes SP1 (solid) and SP2 (dashed) in the acoustic space before the simulation. The dots represent stored exemplars.

The data was designed to mirror a state of a sound change at which one group of speakers (represented by agent B) has already adopted a new variant of a sound (SP1) while another group (represented by agent A) is affected by a phonetic bias that skews that sound in the direction of the change. These configurations can often be observed both in apparent-time data and across regional varieties of which one has undergone a sound change faster than another.

The two agents shown in Figure 1 interacted with each other following the architecture described in 2.1, while applying either the absolute, the relative, or both memorisation criteria in perception, until there were no more relevant acoustic changes.

3. RESULTS

Figure 2 shows the outcomes of the three simulations. When only the absolute memorisation criterion was applied (Figure 2a), agent A clearly adapted SP1 to that of agent B. This is because the exemplars of SP1 produced by agent B were very likely to fall within agent A's solid ellipse, i.e. they were typical enough to become members of A's SP1. Exemplars of SP1 produced by agent A, on the other hand, were likely to have low values of both Feature 1 and 2 (see agent A's solid ellipse in Figure 1), and thus would have been too atypical according to the absolute criterion to be incorporated into agent B's narrow SP1. Neither agents' SP2 shifted acoustically compared to the baseline depicted in Figure 1, but narrowed over the course of the simulation.

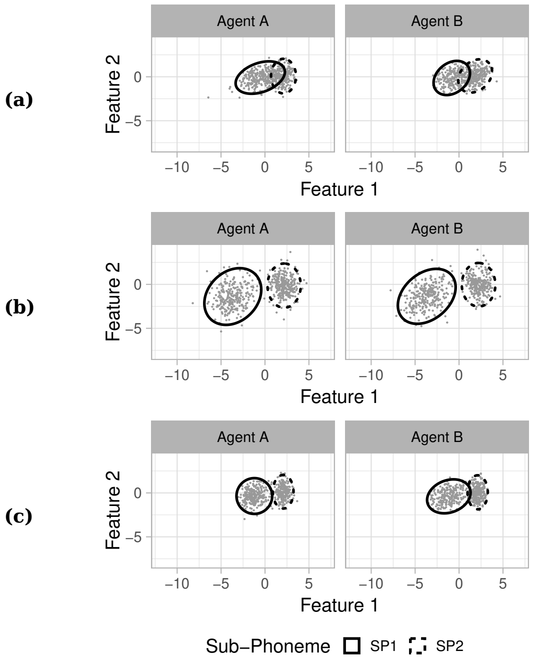

Figure 2: State of agent A's and B's representations of sub-phonemic classes in the acoustic space after the simulation. The dots represent stored exemplars. **(a)** Only the absolute criterion, **(b)** only the relative criterion, and **(c)** both memorisation criteria were applied to decide whether perceived tokens should be memorised.

When only the relative memorisation criterion was applied (Figure 2b), both agents ended up with two non-overlapping sub-phonemic classes. In this case, the centroid of agent B's SP1 moved towards that of agent A's SP1. This is because agent B accepted agent A's exemplars of SP1 which were likely to fall into an unambiguous part of the acoustic space, i.e. a part where the acoustic distribution of SP2 is not in competition with that of SP1. Agent A, on the other hand, rejected those exemplars of SP1 produced by agent B that were located in the part of the acoustic space where agent A's SP1 and SP2 overlapped. As in the first simulation, neither agents' SP2 changed.

Finally, when both memorisation criteria had to be passed in order for an exemplar to be memorised (Figure 2c), agent A's SP1 shifted more towards agent B's SP1 than vice versa, and both the overlap between the sub-phonemes as well as their variance were reduced. Thus, this simulation displays characteristics of the previous two simulations.

4. DISCUSSION

The three simulations presented in section 3 had visibly diverse outcomes despite starting from the same input data. The only difference between the simulations was the criterion applied for the selective memorisation of perceived exemplars. Both the absolute and relative criterion can cause phonetic shifts and stability (i.e. importantly, `soundChangeR` does not inevitably predict change), but both also have side effects.

The absolute criterion can reinforce phonetic biases that are present in the input data. In the presented case, agent A's phonetic bias was reinforced through interactions with agent B whose variants of SP1 lay in the direction of the bias. This mechanism of `soundChangeR` is based on the idea that speakers that have already adopted a new variant of a sound may serve as attractors to speakers whose variant is skewed in that direction e.g. because of a phonetic bias [24]. The side effect of the absolute criterion is a reduction of within-category variance as acoustic outliers are rejected (although the severity of this side effect depends on the Mahalanobis distance threshold set by the user). Other ABMs of sound change instead opt for explicit production biases, i.e. a newly produced exemplar is shifted slightly in the direction of an articulatory bias before being transmitted to the listener [9, 11, 12]. While this reduces the risk of a return-to-the-mean effect, the constant application of the bias may shift the acoustic distributions further than it would in a real sound change.

The relative criterion results in phonetic shifts because of the repulsion created by the constant pressure against ambiguous exemplars. Since SP1 and SP2 partially overlap for both agents in Figure 1, new exemplars that are located in this area of the space are indiscriminable for the agents and therefore not memorised. When applied on its own, the relative criterion does not prevent even the furthest outliers from being memorised as long as they are unambiguous, which can lead to an expansion of the acoustic-phonetic space beyond plausible values. However, the third simulation showed that this side effect of the relative criterion can be counteracted by the absolute criterion while still maintaining the contrast between sub-phonemic classes as desired.

Phonetic stability was achieved for SP2 in all three simulations. This was because the agents had a similar representation of this sub-phoneme and thus, there was no force that could cause acoustic changes as a result of the interactions. This underlines the necessity for phonetically heterogeneous input data: none of the algorithms in `soundChangeR` can create the kinds of phonetic biases in speech production which are hypothesised to lead to change [25]. The forces that create changes in a simulation and result from the agents' production-perception feedback loop only take effect if the agents have sufficiently distinct representations of the sounds under investigation to begin with.

ABMs are a powerful tool for the investigation of so called complex adaptive systems, i.e. systems in which individual entities interact with one another according to some rules and can change as a response to their environment [6, 26]. Spoken language is well suited to be modelled as a complex adaptive system in which microscopic variation at the level of individual interactions leads to community level categorical change, as also shown by the recent upsurge in ABMs designed to answer specific questions in the field of sound change [9, 12]. The simulations presented in this paper showed that the way we handle the probabilistic selection of exemplars is critical in determining the direction of a phonetic shift. In general, ABMs like `soundChangeR` force us to be explicit about the assumptions that underlie the theoretical model to be tested. It is for this reason that we hope that `soundChangeR` will be applied to a variety of speech production data in the future: to gain a better understanding of the interplay of factors that impact the emergence and spread of many kinds of sound changes.

5. ACKNOWLEDGEMENTS

This research was supported by European Research Council Grant No. 742289 "Human interaction and the evolution of spoken accent" (2017-2023).

6. REFERENCES

[1] J. J. Ohala, "Sound change is drawn from a pool of synchronic variation," in *Language Change: Contributions to the Study of Its Causes*, ser. Trends in Linguistics, Studies and Monographs, L. E. Breivik and E. H. Jahr, Eds. Berlin: Mouton de Gruyter, 1989, no. 43, pp. 173–198.

[2] P. Trudgill, "Colonial dialect contact in the history of European languages: On the irrelevance of identity to new-dialect formation," *Language in Society*, vol. 37, pp. 241–280, 2008.

[3] W. Labov, *Principles of Linguistic Change. Volume 2: Social Factors*. Malden: Blackwell, 2001.

[4] J. B. Pierrehumbert, "Exemplar dynamics: Word frequency, lenition and contrast," in *Frequency Effects and the Emergence of Linguistic Structure*, J. Bybee and P. J. Hopper, Eds. Amsterdam: John Benjamins, 2001, pp. 137–155.

[5] J. Bybee, "Usage-based Phonology," in *Functionalism and Formalism in Linguistics: Volume I: General Papers*, M. Darnell, E. A. Moravcsik, M. Noonan, F. J. Newmeyer, and K. Wheatley, Eds. Amsterdam: John Benjamins, 1999, pp. 211–242.

[6] C. Beckner, R. Blythe, J. Bybee, M. H. Christiansen, W. Croft, N. C. Ellis, J. Holland, J. Ke, D. Larsen-Freeman, and T. Schoenemann, "Language Is a Complex Adaptive System: Position Paper," *Language Learning*, vol. 59, pp. 1–26, 2009.

[7] J. S. Pardo, "On phonetic convergence during conversational interaction," *The Journal of the Acoustical Society of America*, vol. 119, no. 4, pp. 2382–2393, 2006.

[8] A. G. Samuel and T. Kraljic, "Perceptual learning for speech," *Attention, Perception, & Psychophysics*, vol. 71, no. 6, pp. 1207–1218, 2009.

[9] S. Todd, J. B. Pierrehumbert, and J. B. Hay, "Word frequency effects in sound change as a consequence of perceptual asymmetries: An exemplar-based model," *Cognition*, vol. 185, pp. 1–20, 2019.

[10] J. N. Stanford and L. A. Kenny, "Revisiting transmission and diffusion: An agent-based model of vowel chain shifts across large communities," *Language Variation and Change*, vol. 25, no. 2, pp. 119–153, 2013.

[11] J. Kirby, "Incipient tonogenesis in Phnom Penh Khmer: Computational studies," *Laboratory Phonology*, vol. 5, no. 1, pp. 195–230, 2014.

[12] M. Sóskuthy, "Understanding change through stability: A computational study of sound change actuation," *Lingua*, vol. 163, pp. 40–60, 2015.

[13] J. Harrington, F. Kleber, and U. Reubold, "Compensation for coarticulation, /u/-fronting, and sound change in standard southern British: An acoustic and perceptual study," *The Journal of the Acoustical Society of America*, vol. 123, no. 5, pp. 2825–2835, 2008.

[14] J. B. Hay, J. B. Pierrehumbert, A. J. Walker, and P. LaShell, "Tracking word frequency effects through 130 years of sound change," *Cognition*, vol. 139, pp. 83–91, 2015.

[15] M. Gubian, J. Cronenberg, and J. Harrington, "Phonetic and Phonological Sound Changes in an Agent-Based Model," *Speech Communication*, vol. 147, pp. 93–115, 2023.

[16] P. Kiparsky, "Formal and empirical issues in phonological typology," in *Phonological Typology*, L. M. Hyman and F. Plank, Eds. Berlin: De Gruyter Mouton, 2018, pp. 54–106.

[17] J. M. Scobbie and J. Stuart-Smith, "Quasi-phonemic contrast and the fuzzy inventory: Examples from Scottish English," in *Contrast in Phonology: Theory, Perception, Acquisition*, P. Avery, B. E. Dresher, and K. Rice, Eds. Berlin: Mouton de Gruyter, 2008, pp. 87–113.

[18] E. Reinisch, K. I. Juhl, and M. Llompart, "The Impact of Free Allophonic Variation on the Perception of Second Language Phonological Categories," *Frontiers in Communication*, vol. 5, p. 47, 2020.

[19] M. Sumner, S. K. Kim, E. King, and K. B. McGowan, "The socially weighted encoding of spoken words: A dual-route approach to speech perception," *Frontiers in Psychology*, vol. 4, p. 1015, 2014.

[20] Y. V. Melguy and K. Johnson, "Perceptual adaptation to a novel accent: Phonetic category expansion or category shift?" *The Journal of the Acoustical Society of America*, vol. 152, no. 4, pp. 2090–2104, 2022.

[21] B. de Boer, "Self-organization in vowel systems," *Journal of Phonetics*, vol. 28, no. 4, pp. 441–465, 2000.

[22] J. Blevins and A. Wedel, "Inhibited sound change: An evolutionary approach to lexical competition," *Diachronica*, vol. 26, no. 2, pp. 143–183, 2009.

[23] M. Stevens, J. Harrington, and F. Schiel, "Associating the origin and spread of sound change using agent-based modelling applied to /s/-retraction in English," *Glossa*, vol. 4, no. 1, pp. 1–30, 2019.

[24] J. Harrington, F. Kleber, U. Reubold, F. Schiel, and M. Stevens, "Linking Cognitive and Social Aspects of Sound Change Using Agent-Based Modeling," *Topics in Cognitive Science*, pp. 1–22, 2018.

[25] A. Garrett and K. Johnson, "Phonetic bias in sound change," in *Origins of Sound Change*, A. C. L. Yu, Ed. Oxford: Oxford University Press, 2013, pp. 51–97.

[26] T. C. Schelling, *Micromotives and Macrobehavior*. New York: Norton, 1978.
