Agent-based modelling, sound change, and metaphony in Southern Italian varieties of Italo-Romance.

Lilian von Bressensdorf, Pia Greca, Jonathan Harrington

Institute for Phonetics and Speech Processing, Ludwig Maximilians University of Munich.
l.bressensdorf|greca|jmh@phonetik.uni-muenchen.de

Abstract

The study uses an agent-based computational model to test the hypothesis that contact between two dialects that are conservative and innovative as far as a sound change is concerned produces an asymmetric shift of the conservative speakers towards the innovative ones. The computational model, which is initialized with speech data from real speakers, is tested for the first time on a morpho-phonological sound change by which cues to inflectional morphology are being transferred from a suffix to a stem vowel in two Italo-Romance dialects of Southern Italy. The results based on quantifying the extent of diphthongization and the emergence of categorical contrasts marking morphological inflection provide some support for the proposed asymmetric shift in dialect contact. The analysis is more generally consistent with feedback models of sound change in which production is stochastically updated as a consequence of memorizing speech signals in speech perception.

Index Terms: sound change, coarticulation, agent-based modeling, dialect contact.

1. Introduction

The present study is concerned with testing a cognitively-inspired agent-based, computational model, the interactive-phonetic agent-based model of sound change (henceforth the IP model [1-3]) based on analyses of so-called metaphony in two southern Italian varieties of Italo-Romance [4-5]. A major concern in this study is to use the IP model to understand how spoken language communication can sporadically and rarely cause non-contrastive synchronic phonetic variation to be phonologized [6] i.e., relevant for distinguishing between word meanings.

The IP model is founded upon the idea that language usage can carry the seeds of its own change [7] and that, following many exemplar and usage-based models, the association between words, phonological units and memorised speech signals is stochastic and dependent on experience i.e., on the interlocutors and speech signals to which the individual has been exposed over the lifespan [8,9].

The further background to this study is that sound change is directional [10], i.e. a sound change X > Y, in which ">" signals the direction of the change, does not imply that Y > X occurs with equal likelihood. This directional sound change has a phonetic origin that can be magnified when so-called conservative and innovative speakers in which sound change is less and more advanced respectively interact with each other [11].

The sound change to be modelled from this perspective in the present study is known as metaphony which occurs in the dialects of the (so-called) Lausberg area in Southern Italy [12,13] and whose phonetic origin is in trans-consonantal vowel-to-vowel coarticulation [14-16]. In these dialects, the cues to inflectional morphology are variously transferred from the suffix to the stem. Thus, whereas in Standard Italian, in which there is no metaphony, the cues to inflectional morphology are carried entirely by the suffix (e.g., /dɔrm**o**, dɔrm**e**/; *I sleep*, *he/she/it sleeps*), while in the dialects on the West coast of the Lausberg area (henceforth West), the suffix can be reduced or even completely deleted with the inflectional cues being manifested as diphthongal and raising contrasts in the stem (thus [d**wo**rm(ə), d**ɔ**rm(ə)] as realisations of /dormu, dorme/ for this example).

The original context for this raising and diphthongization is a high vowel suffix, i.e. either /u/ (the Lausberg counterpart of standard Italian suffix /o/, as in the example above) or /i/, e.g., standard Italian /kɔtti, kɔtta/('good' masc. pl. and fem. sg.) but West [k**wo**tt(ə), k**ɔ**tt(ə)]. Thus, in the same way that English has two contrasting allomorphs for the morpheme LEAF depending on whether it is in a singular, /liː**f**/, or in a plural, /liː**v**/, context, some speakers of the West dialects also have contrasting allomorphs (a morpheme DORM with contrasting allomorphs /d**wo**rm, d**ɔ**rm/) as well as contrasting phonological categories (e.g., /ɔ, wo/) *between* morphemes (/d**ɔ**rm, k**ɔ**tt/ vs. /d**wɔ**rm, k**wɔ**tt/) in different suffix vowel (/e, a/ vs. /u, i/) contexts.

A recent study [13] shows that the Lausberg dialects differ in the extent to which metaphony has progressed. In the dialect spoken in and around the village of Mormanno (henceforth MM), some 50 km inland from the coastal region where the West dialects are spoken, the suffix was found to be less likely to be reduced or deleted; and while there was some raising, there was no diphthongization in the stem vowel as in the West dialect. Overall, the study in [13] found the transfer of cues to morphological inflection from the suffix to the stem to be weaker in MM than in West.

1.1. Predictions

If the West is at a more advanced stage of metaphony, then interacting agents in the IP model should cause the MM group to adopt the metaphonic characteristics of West to a greater extent than West will of MM. Two predictions were tested in this regard. Firstly, MM stem vowels should become more diphthongal in a high vowel suffix context following interaction (whereas those of the West should not become more monophthongal). Secondly, the extent of phonological categorization is predicted to increase for MM in the direction of the West dialect.

2. Method

2.1. Speakers, materials, acoustic parameters

The real speech data used to initialize the agents of the computational model in both dialects involved 28 MM (age range: 22-88 years; 15 F) and 26 West (age range: 27-92 years; 10 F) speakers. The materials included 45 words that were constituted from 21 lexical morphemes (1 verb, 6 adjectives, 14 nouns) containing a primary stressed vowel that varied between monophthongal [ɔ,o] and diphthongal [wo]. The 23 morphemes occurred in the context of between 2-4 inflectional suffixes that marked number and gender in the case of nouns and adjectives, and 1st, 2nd, or 3rd person, present tense singular in the case of the verbs. We emphasize here that 'the vowel' in this study *always refers to the (primary stressed) vowel of the first lexical morpheme* (e.g., the vowel variants of morphemes COTT, DORM etc.), and not to the second inflectional morpheme, i.e. the suffix vowel, which was typically reduced or deleted and not analyzed acoustically. The participants produced each of the 51 words in a picture-naming task resulting in a total number of just over 4000 word tokens that were available in this study.

The speech data were digitized at 48 kHz and semi-automatically segmented with the MAUS forced alignment system [17]. Formant frequencies were calculated between the acoustic onset and offset of the vowel with a 25 ms window and 5 ms frame shift and also manually checked and corrected. The first three coefficients of the discrete cosine transformation (DCT) that are proportional respectively to the F1 mean, linear slope, and curvature [18] were obtained from *z*-score speaker-normalized [19] and linearly time-normalized formants between the acoustic onset and offset. The DCT coefficients were used to reconstruct a DCT-smoothed equivalent of each F1 trajectory. The focus in this study is F1 because metaphonic effects are manifested predominantly as changes in vowel height [13].

2.2. Agent-based model

The software used for the agent-based model was from the R package `SoundChangeR` [20], a cognitively-inspired computational model in which words, phonological classes, and acoustic parameters stand in a stochastic relationship to each other [9] and in which, as in other computational models with a similar architecture [21,22], sound change comes about because an agent's memorized acoustic tokens that are input into speech production can be updated as a consequence of perceiving and selectively memorizing new acoustic speech signals produced by other agents.

2.2.1. Agents

Since the 1-2 repetitions per word were insufficient for building a stochastic association between vowels and the acoustic data, the original 54 speakers were collapsed into 13 agents of which 7 agents were based on pooled MM and 6 on pooled West speakers. The sample size for each word and agent was further increased by a factor of 10 using the SMOTE algorithm [2,23], which expands a cluster of data points but without changing substantially its initial statistical properties. After the application of these two measures, there was a mean of 79 acoustic tokens per agent per word and a mean of 3730 acoustic tokens per agent, such that each acoustic token consisted of the first three DCT coefficients calculated over F1 between the acoustic vowel onset and offset (2.1).

2.2.2. Automatic derivation of morpho-phonological classes

Since the degree of allomorphic contrast within lexical morphemes is a factor to be tested, the method in [2] was used to derive them automatically from the speech signal (rather than presuming that they exist *a priori*). This was a two-step approach applied separately to each agent and separately to each morpheme's vowels - e.g., see Fig. 1, to all the vowels of BON ('good') in the four suffix contexts together). In the first step (Fig. 1), a clustering algorithm based on Gaussian mixture modeling [24] was applied to the DCT coefficients of all vowels of a given morpheme to derive any number of *acoustic clusters*. Secondly, using an algorithm based on non-negative matrix factorization [13,25], one or more acoustic clusters were designated as a morphophonological (MP) class if they contained an allomorph that, for the most part, did not occur in any other acoustic cluster – also under the condition that an allomorph could only ever occur in one MP class. Since this derivation of MP classes was done separately for each agent, then the association between morphemes, their allomorphs, and acoustic tokens was necessarily agent-specific. This aspect of the model was designed to give expression to the idea that, as much research on sound change shows, not all individuals even of the same community categorize acoustic cues to a particular contrast to the same degree [26].

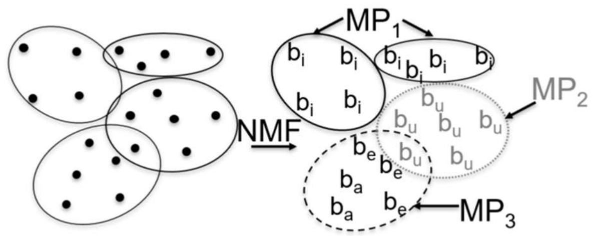

Figure 1: A schematic plot in two dimensions of acoustic tokens of the vowel (the dots) of BON *('good') organized into four acoustic clusters following the application of GMM (left) and then into three MP classes (right) for 'boni' (*b_i*),that includes two acoustic clusters, for 'bone' (*b_e*) and 'bona' (*b_a*) together (dashed ellipse) and for 'bonu' (*b_u*, grey ellipse).*

2.2.3. A single interaction in the Agent based model (ABM)

A single interaction consisted of a pair of agents randomly chosen from the pool of 13 agents (2.2.1) of which one was the *agent talker* and the other the *agent listener*. The former randomly selected one of the 51 words, w_j, and built a Gaussian model over w_j's acoustic tokens in a three-dimensional DCT space. The model was used to generate a new acoustic token, $x_{new.j}$ that, together with w_j , was transmitted to the agent listener. Subsequently, the agent listener identified in its own memory the MP class with which w_j was associated. Next, the listener memorized the acoustic token, $x_{new.j}$, only if it passed both *discriminability* and *typicality* tests. As an example of the first of these, if w_j = *boni* and if the allomorphs of BON are categorized as in Fig. 1 with *boni* in MP_1, then the discriminability test is passed if $p(x_{new.j}|MP_1) > p(x_{new.j}|MP_2)$ and if $p(x_{new.j}|MP_1) > p(x_{new.j}|MP_3)$. The typicality test prevented atypical acoustic tokens becoming members of an MP class by setting a threshold at 0.95 i.e., the typicality test was passed if $p(x_{new.j}|MP_1) > 0.95$. If the received token is memorized, another random token associated with the same word is deleted from the listener's memory to prevent the memory from enlarging.

2.2.4. An ABM run

An ABM run consisted of 42,000 interactions; this number was selected by eyeballing the interaction number beyond which change was minimal. The MP classes in 2.2.2 were recalculated at intervals of 1000 interactions.

There were five runs in total. The run that was selected for subsequent analysis was the most typical of the five. The state prior to any interaction will be referred to as the *baseline*, and the state at the end of the run as the *post-run*.

2.3. Quantification

The two quantifications described below were carried out *after* the ABM had been run (2.2.4): they therefore played no part in the derivation of the post-run from the baseline.

2.3.1. Diphthongization

The extent of F1-dipthongization was quantified with the third moment of an F1 smoothed trajectory that had been reconstructed with an inverse DCT applied to the first three DCT coefficients. The third moment $m_{3.i}$ for a given vowel, *i*, is a measure of the skewness and was calculated from (1):

$$m_{3.i} = \left(\frac{\sum_{n=0}^{N} x[n](\mathrm{n} - m_{1.i})^3}{\sum_{n=0}^{N} x[n]}\right) m_{2.i}^{-1.5} \quad (1)$$

where $x[n]$ (n = 0 1, 2… N=20) is vowel *i*'s 21-point, time-normalized and DCT-smoothed F1-trajectory which was rescaled so that the trajectory's minimum value is set to zero, and $m_{1.i}$ and $m_{2.i}$ are the first and second moments of the same F1-trajectory (2,3):

$$m_{1.i} = \frac{\sum_{n=0}^{N} x[n]n}{\sum_{n=0}^{N} x[n]} \quad (2)$$

$$m_{2.i} = \frac{\sum_{n=0}^{N} x[n](\mathrm{n} - m_{1.i})^2}{\sum_{n=0}^{N} x[n]} \quad (3)$$

The extent of diphthongization in the context of high vowel /i, u/ suffixes was further analysed with a mixed model [27] and subsequent Bonferroni-corrected, post-hoc testing [28] in R [29] as in (4):

m3.i ~ dialect * state + (state|agent) + (dialect|stem) (4)

with $m_{3.i}$ as the dependent variable, with fixed factors dialect (MM, West) and state (baseline, post-run), and with random factors agent and stem (*bon-*, *cott-*, *zopp-* etc.).

2.3.2. Phonological categorization

Phonological classes were derived separately for each agent by applying the GMM and NMF methodology (2.2.2) to the vowels of all words together (rather than separately by morpheme as in 2.2.2), once in the baseline and once in the post-run. This was done in order to determine the extent to which phonological vowel contrasts emerged depending on the suffix context. For this purpose, a proportional measure of phonologization, p_{prop}, was applied in (5):

$$p_{prop} = (w_{n.s} + w_{n.d})/w_n \quad (5)$$

where w_n is the total number of unique word pairs, $w_{n.s}$ the number of same-suffixed word pairs that occurred in the same phonological class and $w_{n.d}$ the number of different suffixed word pairs that occurred in different phonological classes.

Consider the application of (5) to a vocabulary of 8 words *boni*, *cotti, dormi*, *bone*, *cotte*, *dorme*, *bona*, *cotta*. When phonologization is maximal, then the words are grouped by suffix context into three separate phonological classes: P_i = {*boni, cotti*, *dormi*}, P_e = {*bone*, *cotte*, *dorme*}, *Pa* = {*bona*, *cotta*}. Applying (5) to this example, $w_{n.s}$ = 7 (three pairs of same suffixed words in P_i: *boni* ~ *cotti*; *boni* ~ *dormi*; *cotti* ~ *dormi*; 3 pairs in P_e, and 1 pair in *Pa*) and $w_{n.d}$ = 21, so p_{prop} = (7+21)/28 = 1 when metaphony is maximally phonologized. For no metaphony in which, as in standard Italian, vowels in the stems of the types of words analyzed here have the same /ɔ/ vowel and are therefore in the same phonological class, $w_{n.d}$ = 0 (because there is only one class), so $p_{prop} = w_{n.s}/w_n$ = 7/28 = ¼. Intermediate degrees of phonologization, therefore, range between ¼ and 1 for this particular lexicon. A generalized linear mixed model was applied to quantify the influence of dialect and state on phonologization and had the form:

hit ~ dialect * state + (state|agent), family= binomial (6)

in which *hit* was T either when a same suffixed word pair was in the same phonological class or when a different suffixed word pair was in a different phonological class, otherwise F (thus, p_{prop} in (5) is the sum of T observations in *hit* divided by the number of observations). Bonferroni corrected, post-hoc tests were calculated in the case of a dialect * state interaction.

3. Results

3.1. Diphthongization

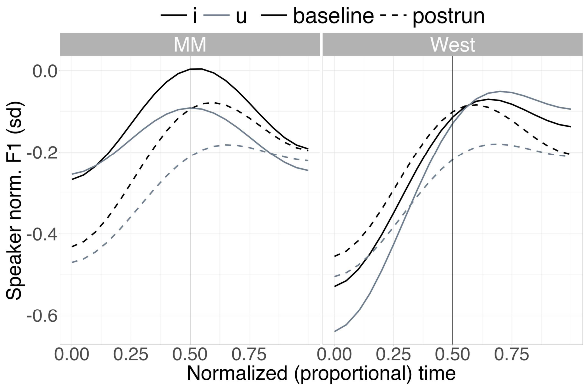

Figure 2: Aggregated DCT-smoothed F1 trajectories of the vowel in the baseline (solid) and post-run (dashed) in the context of the suffix /i/ (grey) and /u/ (dashed).

A comparison between MM and West on the solid trajectories shows greater diphthongization in an /i, u/ suffix vowel context for West in the baseline. This is especially so since these trajectories reach their peak at the vowel's temporal midpoint in MM (the vertical line at t = 0.5) but somewhat after this time marker in West. The extent of diphthongisation for MM was evidently greater in the post-run than in the baseline (compare black and grey trajectories).

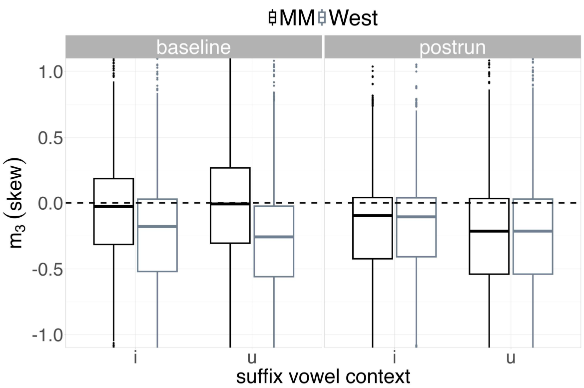

Figure 3: Distribution of the skew of F1 trajectories in the context of /i, u/ suffixes in the baseline and post-run for MM (black) and West (grey). The dashed line at m_3 = 0 denotes no skew.

Compatibly with Fig. 2, the distribution of m_3 (Fig. 3) shows greater skew in the post-run than the baseline for MM but minimal change for West. The results of the mixed model in (4) showed a significant effect of dialect on skew ($F_{1,26.8}$ = 14.648, p < 0.001) and of state ($F_{1,11.0}$ = 12.532, p < 0.01) and a significant ($F_{1, 11.0}$ = 32.689, p < 0.001) interaction between these factors. Post-hoc tests showed a significant difference on skew between the dialects in the baseline (p < 0.001) and between the baseline and post-run for MM (p < 0.001) but not for West. All these data are consistent, therefore, with the prediction of a shift such that MM stem vowels became more diphthongal following interaction but that those of West did not change (and did not become significantly more monophthongal in the direction of MM).

3.2. Phonologization

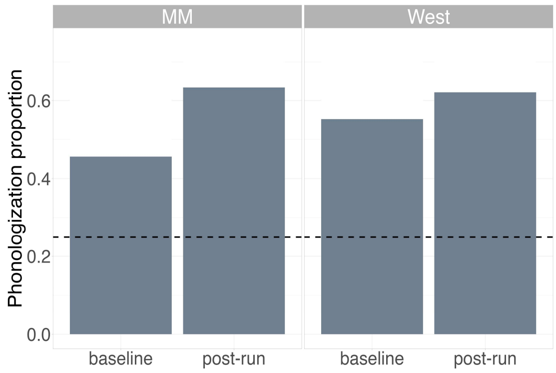

Figure 4: The phonologization score, p_{prop} (5), for the two dialects in the baseline and post-run. The horizontal dashed line corresponds to no phonologization.

While Fig. 4 shows a degree of phonologization which is greater than to be expected from varieties such as standard Italian with no phonologization (the horizontal line in Fig. 4), the phonologization in the baseline was even greater for West: this is to be expected, if metaphony, involving the phonologization of the suffix in the stem vowel, is more advanced in West than in MM (see Section 1.5). Compatibly with Fig. 4, the results of the application of (6) showed a significant increase in the phonologization score for MM from the baseline to the post-run (Wald z = 3.23, p < 0.01), but no other significant differences.

4. Discussion

According to the IP model [2,3], interaction between speakers who are conservative and innovative (as far as the uptake of a sound change is concerned) results in an asymmetric change such that conservative speakers shift in the direction of those who are more innovative as far as a sound change is concerned. The results from this study provide some support for such model. Firstly, vowels of the more conservative Mormanno variety adopted the diphthongal characteristics found in the more innovative variety of the West coast dialect, after the ABM was run. Secondly, suffix-dependent, morpho-phonological categorization based on acoustic processing of these vowels was greater prior to running the ABM for West, with the extent of categorization in Mormanno attaining that of the West after the ABM had run.

We are not suggesting, with this result, that all sound change is necessarily the result of dialect contact, but only that there has to be a group of individuals – i.e. a sociolect within the same dialect or younger speakers [30] – that have 'exaggerated' forms of phonetic variation that cause inherent phonetic biases due especially to coarticulation and undershoot to be further magnified through contact of such a group with the rest of the community.

The study is the first, to our knowledge, to have applied an agent-based model that makes explicit how real, dynamically changing speech signals, memorization, (morpho)phonological re-categorization, and changes to speech production are associated. This type of model is relevant not only for understanding the progression of metaphony that has been analysed here but more generally for so-called phonologization-based sound changes [6] such as tonogenesis [31] and the development of contrastive vowel nasalization [26] in which there is a re-weighting of cues [32] from a coarticulatory source (in the case of metaphony: the suffix vowel) to a coarticulatory effect (to the vowel of the lexical morpheme). However, a major limitation of this study is that, because we have not considered the acoustic properties of the suffix, we are not able with this analysis to provide a computational simulation of cue-trading i.e., the progressive shift in the listener's attention from the coarticulatory source to the effect that has been shown in various studies [33] to be one of the major factors in the progression of phonologization-based sound changes. The difficulties in incorporating cue-trading into this type of model stem not only from the increased complexity through the need to incorporate time-based, transitional probabilities between the coarticulatory effect and source, but also because there is not yet a sufficiently detailed understanding of the cognitive mechanisms by which cue-reweighting in perception influences speech production.

The present study showing how phonetic variation is converted into sound change raises the more general issue of whether sound change is inevitable in this type of model. This is an important consideration given that stability i.e., phonetic variation without sound change is far more likely than instability and change [34]. Previous studies have shown that the development of sound change in this type of model is by no means inevitable [1] if the coarticulatory cues that could lead to change are too weak or variably distributed across speakers [35]. In addition, the technique of automatically deriving (morpho)phonological classes from bottom-up derived acoustic clusters does not make phonologization (2.3.2) inevitable, for the reason that, if there is not sufficient acoustic information to distinguish e.g., the allomorphs of COTT, then they will in all likelihood remain in the same morpho-phonological class. We intend to investigate just this aspect further (and to test the prediction that standard Italian COTT variants remain within the same category) by applying this model to standard Italian, which does not have the types of metaphonic changes that characterise the Lausberg dialects investigated here.

Finally, another major issue to be re-considered is whether (and how) the two types of categorization in this study are connected. One of these was morpho-phonological and formed part of flexible (re)categorization during interaction. The other was *post-hoc* (and had no involvement in interaction), and was concerned with the extent of phonologization i.e., with the organization of same-suffixed words into mutually exclusive categories. Whether both these levels of abstractions are necessary and (if so) whether both should be involved in interaction will also be a question for future research.

5. Acknowledgements

The research was funded by the European Union (ERC, SoundAct, project n° 101053194). Views and opinions expressed are, however, those of the authors only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

6. References

[1] M. Jochim and F. Kleber, "Reconstructing the timeline of a consonantal change in a German dialect: evidence from agent-based modeling," in *Speech Dynamics: Synchronic Variation and Diachronic Change*, F. Kleber and T. Rathcke, Eds. Berlin, Boston: De Gruyter Mouton, 2024, pp. 307–340. [Online]. Available: https://doi.org/10.1515/9783110765328-011

[2] M. Gubian, J. Cronenberg, and J. Harrington, "Phonetic and phonological changes in an agent-based model," *Speech Commun.*, vol. 147, pp. 93–115, Feb. 2023.

[3] J. Harrington, F. Kleber, U. Reubold, F. Schiel, and M. Stevens, "Linking cognitive and social aspects of sound change using agent-based modeling," *Top. Cogn. Sci.*, vol. 10, no. 4, pp. 707–728, Mar. 2018.

[4] M. Maiden, *Interactive Morphonology: Metaphony in Italy*. London, New York: Routledge, 1991.

[5] M. Loporcaro, "Metaphony and diphthongization in Southern Italy: Reconstructive implications for sound change in early Romance," in *Approaches to Metaphony in the Languages of Italy*, F. Torres-Tamarit, K. Linke, and M. v. Oostendorp, Eds. Berlin, Boston: De Gruyter Mouton, 2016, pp. 55–87.

[6] A. Yu, "Toward an individual-difference perspective on phonologization," *Glossa: J. Gen. Linguist.*, vol. 6, no. 14, Feb. 2021. [Online]. Available: https://doi.org/10.5334/gjgl.661

[7] C. Beckner et al., "Language is a complex adaptive system: Position paper," *Lang. Learn.*, vol. 59, pp. 1–26, Dec. 2009. [Online]. Available: http://dx.doi.org/10.1111/j.1467-9922.2009.00533.x

[8] K. Johnson, "Speech perception without speaker normalization: An exemplar model," in *Talker Variability in Speech Processing*, K. Johnson and J. W. Mullennix, Eds. San Diego: Academic Press, 1997, pp. 145–165.

[9] J. B. Pierrehumbert, "Phonetic diversity, statistical learning, and acquisition of phonology," *Lang. Speech*, vol. 46, no. 2–3, pp. 115–154, 2003.

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

[11] S. Chang, M. Plauché, and J. Ohala, "Markedness and consonant confusion asymmetries," in *The Role of Speech Perception in Phonology*, E. Hume and K. Johnson, Eds. San Diego: Academic Press, 2001, pp. 79–101.

[12] M. Loporcaro, "Metaphony and diphthongization in Southern Italy: Reconstructive implications for sound change in early Romance," in *Approaches to Metaphony in the Languages of Italy*, F. Torres-Tamarit, K. Linke, and M. v. Oostendorp, Eds. Berlin: De Gruyter Mouton, 2016, pp. 55–87.

[13] P. Greca, M. Gubian, and J. Harrington, "The relationship between the coarticulatory source and effect in sound change: Evidence from Italo-Romance metaphony in the Lausberg area," *Lab. Phonol.*, vol. 15, no. 1, pp. 1–54, Jun. 2024. [Online]. Available: https://doi.org/10.16995/labphon.9228

[14] D. Recasens, *Coarticulation and Sound Change in Romance*. Amsterdam: John Benjamins, 2014.

[15] S. E. Öhman, "Coarticulation in VCV utterances: Spectrographic measurements," *J. Acoust. Soc. Am.*, vol. 39, no. 1, pp. 151–168, Jan. 1966.

[16] P. Hoole and M. Pouplier, "Öhman returns: New horizons in the collection and analysis of imaging data in speech production research," *Comput. Speech Lang.*, vol. 45, pp. 253–277, Mar. 2017.

[17] T. Kisler, U. Reichel, and F. Schiel, "Multilingual processing of speech via web services," *Comput. Speech Lang.*, vol. 45, pp. 326–347, Jan. 2017.

[18] C. Watson and J. Harrington, "Acoustic evidence for dynamic formant trajectories in Australian English vowels," *J. Acoust. Soc. Am.*, vol. 106, pp. 458–468, Mar. 1999.

[19] B. M. Lobanov, "Classification of Russian vowels spoken by different speakers," *J. Acoust. Soc. Am.*, vol. 49, pp. 606–608, Feb. 1971.

[20] J. Cronenberg, M. Gubian, J. Harrington, and F. Schiel, "soundChangeR: An agent-based model of sound change," 2022, R package version 1.0.0. [Online]. Available: https://github.com/IPS-LMU/soundChangeR

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

[22] A. Wedel, "Exemplar models, evolution and language change," *Linguist. Rev.*, vol. 23, no. 3, pp. 247–274, Nov. 2006.

[23] N. Chawla, K. Bowyer, L. Hall, and W. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, Jun. 2002.

[24] D. Reynolds, "Gaussian mixture models," in *Encyclopedia of Biometrics*, S. Li and A. Jain, Eds. Boston: Springer, 2009, pp. 659–663.

[25] O. Mangin, D. Filliat, L. ten Bosch, and P.-Y. Oudeyer, "MCA-NMF: Multimodal concept acquisition with non-negative matrix factorization," *PLoS One*, vol. 10, no. 10, 2015. [Online]. Available: https://doi.org/10.1371/journal.pone.0140732

[26] P. Beddor, "Advancements of phonetics in the 21st century: Theoretical and empirical issues in the phonetics of sound change," *J. Phonetics*, vol. 97, 2023. [Online]. Available: https://doi.org/10.1016/j.wocn.2023.101228

[27] A. Kuznetsova, P. Brockhoff, and R. Christensen, “lmerTest package: Tests in linear mixed effects models,” *J. Stat. Softw.*, vol. 82, no. 13, pp. 1–26, 2017. [Online]. Available: https://cran.r-project.org/package=lmerTest

[28] R. Lenth, *emmeans: Estimated Marginal Means, aka Least-Squares Means*, version 1.10.7, 2025. [Online]. Available: https://cran.r-project.org/package=emmeans

[29] R Core Team, *R: A Language and Environment for Statistical Computing*, version 4.4.2, Vienna, Austria: R Foundation for Statistical Computing, 2024. [Online]. Available: https://www.R-project.org

[30] R. Dodsworth, “Bipartite network structures and individual differences in sound change,” *Glossa: J. Gen. Linguist.*, vol. 4, no. 61, pp. 1–29, Jun. 2019, doi: 10.5334/gjgl.647.

[31] J. Gao and J. Kirby, “Laryngeal contrast and sound change: The production and perception of plosive voicing and co-intrinsic pitch,” *Language*, vol. 100, no. 1, pp. 124–158, Mar. 2024.

[32] J. Schertz and E. Clare, “Phonetic cue weighting in perception and production,” *WIREs Cogn. Sci.*, vol. 11, no. 2, 2020, doi: 10.1002/wcs.1521.

[33] J. Kuang and A. Cui, “Relative cue weighting in production and perception of an ongoing sound change in Southern Yi,” *J. Phon.*, vol. 71, pp. 194–214, Oct. 2018.

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

[35] M. Stevens and J. Harrington, “Individual variation and the coarticulatory path to sound change: Agent-based modeling of /str/ in English and Italian,” *Glossa: J. Gen. Linguist.*, vol. 7, no. 1, Jul. 2022, doi: 10.16995/glossa.8869.