INTERSPEECH 2021
30 August – 3 September, 2021, Brno, Czechia

Articulatory characteristics of Icelandic voiced fricative lenition: gradience, categoricity, and speaker/gesture-specific effects

Brynhildur Stefansdottir, Francesco Burroni*, Sam Tilsen*

Cornell University, USA
bs724@cornell.edu, fb279@cornell.edu, tilsen@cornell.edu

Abstract

Icelandic voiced fricatives frequently reduce in connected speech. However, systematic investigations of the phenomenon from acoustic and articulatory perspectives are lacking. To further the understanding of this lenition process, we present electromagnetic articulography and acoustic data from four speakers concerning the intervocalic realization of the dental and velar fricatives. The results show that lenition is mostly gradient, but some speakers and places of articulation exhibit two distinct modes suggesting a categorical distinction. Moreover, in some tokens, the fricative constriction is absent from the articulatory trajectories. Finally, the relation between lenition and speech rate, style, and stress is also subject to speaker- and gesture-specific effects. We conclude by evaluating how our findings challenge the common assumptions, made in the literature, that lenition is a change in gestural target or a perceptually driven phenomenon.

Index Terms: articulatory phonology, EMA, Icelandic, lenition, voiced fricatives

1. Introduction

Icelandic voiced fricatives are described as frequently weakening and even deleting in connected speech [1], [2]. For example, the adverb *nefnilega* 'namely' has [nɛpnɪlɛɣa] as its citation pronunciation, but [nɛpnɪlɛa], with the fricative seemingly deleted, is considered more natural [1]. The apparent deletion of a fricative can lead to further reduction, e.g., [nɛpnɪlɛa]>[nɛplɛa]>[nɛpla]. In fact, the corresponding written forms *nebbla* and *nefla* appears in informal writing, indicating lexicalization [3]. Previous literature on lenition processes in Icelandic has mainly focused on identifying which processes occur and describing them, often either without experimental support or without a full statistical analysis, e.g., [1], [2]. One systematic investigation of the phenomenon by [3] looked at reduction in adverbs with the suffix *-lega*, but no articulatory data were discussed. The present work aims to offer further insight on the nature of Icelandic voiced fricative lenition by looking at the articulation of intervocalic dental [ð] and velar [ɣ] fricatives. We collected articulatory data since they are necessary to determine whether fricative constrictions are still formed, even when they are inaudible, and to get a more fine-grained perspective on speakers' production of lenited tokens.

1.1. Research questions, hypotheses, and predictions

The four main questions regarding Icelandic voiced fricative lenition (henceforth lenition) we aim to answer are:

1) Is lenition a categorical or gradient phenomenon?
2) Is lenition best characterized as gestural undershoot, or are there cases where no fricative gesture is present in the articulatory trajectories?
3) How does lenition interact with other (para)linguistic factors, namely, speech rate, speech style, and stress?
4) What is the relation between lenition as an articulatory process and its acoustic output, in particular intensity and duration?

We make the following hypotheses and predictions for each research question. For (1), if lenition is categorical, we expect speakers to exhibit two (or more) distinct categories, corresponding to unlenited and lenited categories, Figure 1 left. If lenition is gradient, however, we expect speakers to display either a single distribution or an unlenited to lenited continuum in constriction degree (CD) space, Figure 1 right.

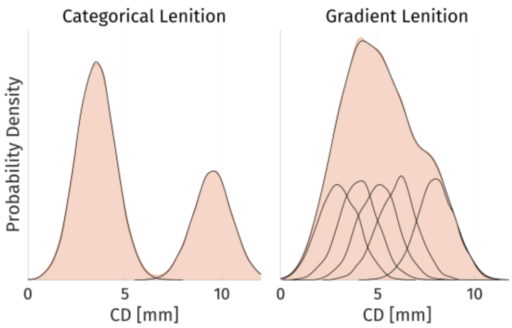

Figure 1: *Predictions of categorical, left, and gradient lenition, right, in CD space.*

For (2), if lenition happens online, we expect to see articulatory "traces" of the planned fricative gestures, even in cases of apparent acoustic deletion. However, if tokens have no traces of the fricative gesture, that could suggest either that speakers have lexicalized word plans without a fricative gesture; or that the endpoint of the lenition continuum is indistinguishable from lexical absence of a fricative gesture. For (3), we expect faster speech rates, casual speech, and unstressed syllables to induce stronger lenition effects since these conditions are known to correlate with less "extreme" articulatory movements. For (4), based on the assumption that the driving force behind lenition is a listener-oriented increase in intensity word-internally [4], we may expect a positive correlation between more or less extreme forms of lenition in CD space and changes in intensity. Similarly, if lenition is mostly a reduction in duration [5], we expect a correlation between the degree of lenition and acoustic segment duration.

---

* Equal contribution.

http://dx.doi.org/10.21437/Interspeech.2021-1903

2. Methods

Data were collected in Ithaca, NY from four native Icelandic speakers (SP01-SP04), one of which was the first author. Synchronized acoustic and articulatory data were collected using an NDI Wave electromagnetic articulometer (EMA) and a shotgun microphone positioned about 1.5 m from the participant. Articulator sensors were placed on the upper lip and lower lip, on the gumline below the lower right incisor, on the tongue blade around 2 cm from the tongue tip (TT), and on the tongue body (TB) around 5 cm from the tip of the tongue. Three reference sensors were placed on the left and right mastoid processes and the nasion to correct for head movement. Target words appeared in randomized order on a screen and were produced within the carrier phrase [seiɣðʏ __ fɪrɪr mɪx] *segðu __fyrir mig* 'say__for me'. Participants alternated blocks where they were instructed to speak either in a formal and clear manner or in a casual and informal manner. Participants completed a total of 20 blocks, except SP01, who completed 24. Stimuli consisted of 20 unique tokens, Table 1. [ð] and [ɣ] appeared as syllable onsets in intervocalic position in two different stress environments: after primary stress and after non-primary stress. Non-primary stress here means after an unstressed syllable or after a syllable with secondary stress. The condition after secondary stress was elicited only for [ɣ] due to constraints on experiment duration.

Table 1 *Stimuli.*

Post primary stress	Post non-primary stress
baða ['paː.**ð**a]	*döbbaða* ['tœp.pa.**ð**a]
staða ['staː.**ð**a]	*aðstaða* ['að.sta.**ð**a]
beða ['pɛː.**ð**a]	*rauðbeða* ['rœið.pɛ.**ð**a]
sleða ['stl̥ɛː.**ð**a]	*snjósleða* ['stnjou(ː).stl̥ɛ.**ð**a]
daga ['taː.**ɣ**a]	*bardaga* ['par.ta.**ɣ**a]
laga ['laː.**ɣ**a]	*tillaga* ['tʰɪl.la.**ɣ**a]
vega ['vɛː.**ɣ**a]	*farvega* ['far.vɛ.**ɣ**a]
lega ['lɛː.**ɣ**a]	*mannlega* ['man.lɛ.**ɣ**a]
	æskudaga ['ais.kʏ.ˌta.**ɣ**a]
	ferðalaga ['fɛr.ða.ˌla.**ɣ**a]
	jeppavega ['jɛh.pa.ˌvɛ.**ɣ**a]
	lagalega ['laː.ɣa.ˌlɛ.**ɣ**a]

Acoustic data were hand-segmented in PRAAT [6] using visual analysis of the waveform, the spectrogram, and changes in intensity and formant trajectories. Fricative duration was defined from the segmentation. In articulatory landmarking, the identification of a point of maximal constriction associated with the fricative gesture was constrained to a target region, which was defined as the acoustic interval of the fricative. In cases where no fricative was audible, the target region was constrained to a fixed period centered on the midpoint of the vowel sequence. The point of maximum constriction was identified using the vertical position of TT for [ð] and TB for [ɣ]. The degree of lenition was operationalized by recentering and reversing the vertical position of TT/TB sensors such that 0 represents maximum constriction and larger values indicate a greater opening in millimeters [7]. When no peak was present, the token was counted as an instance in which the constriction gesture is absent. No CD was estimated for such tokens. We further collected an acoustic proxy of CD from intensity to evaluate whether speakers manipulate CD and acoustic intensity in tandem. This was operationalized as the maximum RMS amplitude during the consonant divided by the minimum RMS measured during the following vowel (V_2) , $\frac{max\ (RMS_C)}{min\ (RMS_{V2})}$ [8]. Speech rate was calculated by dividing the number of syllables per utterance by the total utterance duration in seconds. Speech rate, thus, represents a measure of syllables produced per second. A few caveats regarding our measure of CD are in order: (i) our metric was based only on vertical sensor position, but some aspect of the constriction may involve horizontal motion as well; (ii) the sensors may have not be placed on the tongue locations which maximally represent the constriction targets. Some caution in interpreting our measure of CD measure is thus warranted.

3. Results

3.1. Categorical or gradient?

To answer the question of whether lenited and unlenited fricatives correspond to separate categories in CD space, we analyzed each participant's realization of [ɣ] and [ð] using Gaussian mixture models (GMMs). Each mixture follows a normal distribution $\mathcal{N}(\mu,\sigma^2)$ and contributes a proportion α to the overall distribution of fricatives in CD space. We fit GMMs with up to 6 components and chose the optimal number of components for each fricative and speaker separately by selecting the model with the lowest Bayesian Information Criterion (BIC).

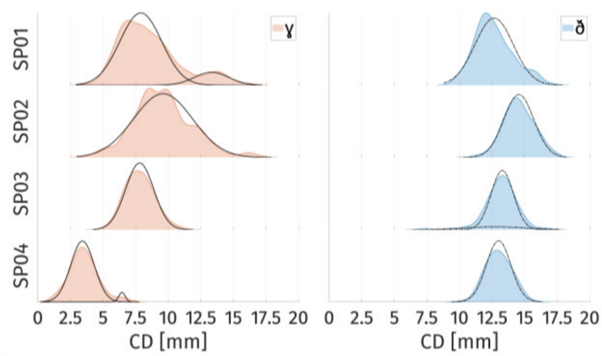

Figure 2 *GMMs of fricatives in CD space, higher CD values indicates lower maximal constriction.*

In Figure 2, the rows show each speaker's CD values in mm on the x-axis, with higher values representing larger distances from the maximum constriction point; the y-axis represents probability density. Overall, the GMMs suggest a gradient lenition process. Only for [ɣ] of SP01 do the GMMs offer evidence for the existence of two non-overlapping mixtures, suggesting categorical lenition. In this case the lenited mixture (larger distance from 0) accounted for around one fourth of the total tokens. Another potential case of categorical lenition, [ɣ] of SP04, cannot be interpreted as such since the second mixture represents only two tokens. In all other cases, the selected model involved just a single Gaussian distribution or, for [ð] of SP03, two distributions with nearly identical modes but differing variance and probability density. In sum, Icelandic voiced fricative lenition seems to mostly be a gradient process, however, instances of categorical lenition seem possible depending both on the speaker and gesture in question.

3.2. Are fricative gestures always present?

Cases with no trace of a fricative constriction gesture only account for around 8% of the data. Some examples of trajectories with and without fricative gestures are plotted in Figure 3. It shows SP01 producing [ˈaiskʏˌtaɣa]. The lines indicate the vertical movement of TB, centered at the onset time of the fricative as estimated from the acoustic signal. In most of the casual speech examples (purple lines) there is no indication of a TB movement towards the fricative target, unlike the examples from formal speech (yellow lines).

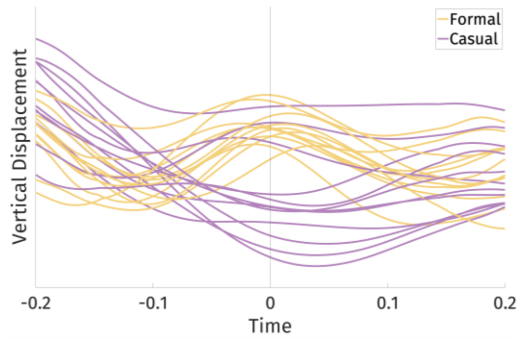

Figure 3: *Example trajectories with* [ɣ] *gesture absent.*

We fit a logistic regression (R^2 = .44, p < .001) to determine relevant factors driving the presence or absence of a fricative gesture. Both word and speaker are significant factors. Figure 4 shows, for instance, that ['stnjoustl̥ɛða], where [ð] appears in an unstressed environment, has the highest proportion, 0.27, of tokens with no identifiable fricative constriction gesture.

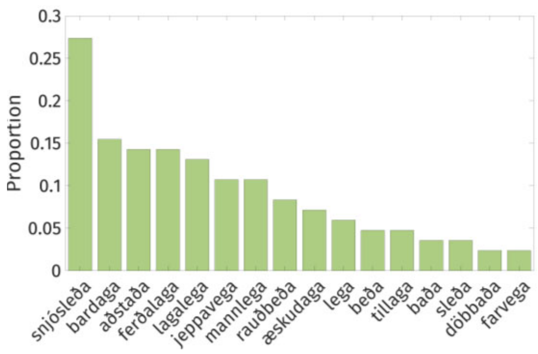

Figure 4: *Proportion of tokens with absent fricative constriction gesture by word.*

Figure 5 shows that some speakers are more likely to produce tokens without a fricative gesture, with SP01 having by far the highest proportion of tokens with no identifiable fricative constriction gesture, 0.15. Factors that decrease the likelihood of producing a token without a fricative constriction gesture are an increase in target word duration and formal speech style. In sum, the results indicate that lenition of Icelandic voiced fricatives is best described as gestural undershoot. While there are examples where no gesture is observed, these are uncommon. The factors driving the absence of fricative constriction gestures are speaker, word identity, word duration, speech style, and fricative gesture type.

Figure 5: *Proportion of tokens with absent fricative constriction gesture by speaker.*

3.3. How do speech rate, speech style, and stress affect lenition?

We analyzed the effects of speech rate, speech style, and stress, on lenition by fitting multiple linear regressions by speaker and gesture separately. The main effects observed are summarized in Table 2, which reports the estimated effect of speech rate, style, and stress on CD, when such factors were found to have a statistically significant effect (i.e., p < .05). To clarify how contrasts were coded for the multiple linear regression, an increase in speech rate corresponds to an increase of one syllable per second, i.e., a faster speech rate, and the effects of style and stress indicate a change to casual speech and to a post non-primary stressed environment. A negative sign indicates a more constricted target (closer to the maximum point of constriction, i.e., 0), while a positive sign indicates a less constricted target (further away from the maximum point of constriction).

Table 2 *Estimated effects of speech rate (SR), style (STY), and stress (STR) on constriction degree. Only significant coefficients are shown.*

		SP01	**SP02**	**SP03**	**SP04**
ɣ	SR	0.6	-	-	0.3
	STY	1.9	-	-	0.4
	STR	1.0	-1.2	-0.3	1.1
ð	SR	-	-	0.8	-0.2
	STY	-	-	-1.1	-0.9
	STR	-	-	-	-0.7

For [ɣ], less constricted targets were observed in faster speech, casual style, and after non-primary stress for SP01 and SP04. However, for SP02 and SP03, we observe more constricted targets following non-primary stress. For [ð] the picture is less clear. Faster speech rate corresponds to less constricted targets for SP03, but more constricted targets for SP04. Similarly, both SP03 and SP04 have more constricted targets in casual speech. Finally, for SP04, tokens are more constricted following non-primary stress. These observations that speech rate, style, and stress interact differently with lenition, depending on the speaker and gesture, is also evident when we fit a linear mixed effect model. A loglikelihood ratio test supports a model that includes effects of speech rate, style, and stress as fixed effects, as well as random intercepts and slopes for each speaker and word (Adj. R^2 = .75, $\chi^2_{(9)}$ = 20.38, p < .01). None of the main effects, however, has a coefficient whose 95% confidence intervals lie entirely above or below 0. This suggests that the magnitude and direction of the main effects vary by speaker and gesture.

3.4. What is the relationship between lenition in articulation and acoustics?

Finally, do changes in CD correlate with changes in duration and intensity? These two dimensions of the acoustic signal are hypothesized to be relevant for listeners' perception of lenited tokens. To answer this question, we measured, separately by speaker and fricative type, the strength of the correlation between CD and changes in RMS amplitude and duration. We found little evidence for a correlation between articulatory lenition and acoustic proxies of lenition based on RMS amplitude. A positive correlation was observed only for [ɣ] of SP03 (r = .18, p < .01) and SP04 (r = .26, p = .0001). Interestingly, a *negative* correlation was observed for [ð] of SP04 (r = -.33, p = .0001). For other speakers' [ð] no correlation was observed. Overall, the strength of the correlation between CD and RMS amplitude is weak, even when significant. A negative correlation for SP04's [ð], as well as a lack of correlations for the other speakers, are problematic findings for the idea that articulatory lenition is being manipulated in tandem with its acoustic output. Regarding the correlation between CD and acoustic duration of the target fricative, we found that three of the four speakers display the expected negative correlation for [ɣ] (SP01 r = -0.24, p < 0.001; SP03 r = -0.20, p < 0.01; SP04 r = -0.49, p < 0.001). Recall again that this sign of the correlation coefficient indicates more constricted targets at longer durations, see Figure 6, where [ɣ] targets are in orange and [ð] in blue.

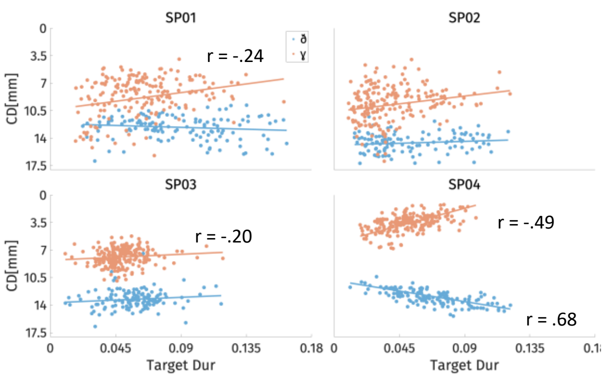

Figure 6 : *Correlation between CD and duration*

On the other hand, SP04's [ð] shows a positive correlation with fricative duration (r = .68, p = 0), indicating more lenition for longer fricative duration. Overall, these patterns indicate that less lenited articulatory targets correlate with longer fricative duration, at least for [ɣ]. An opposite trend for SP04's [ð], which is more lenited at longer durations, suggests that speaker and gesture specific effects come into play and that lenition is not a unitary phenomenon, even when it comes to the relationship with its acoustic output.

4. Discussion

Our results show that, at least as far as Icelandic voiced fricative lenition is concerned, lenition can be appropriately characterized as a reduction in CD, as hypothesized in previous work [9], and to a lesser extent, as a shortening of fricative duration, as it can be estimated from the acoustic signal [5]. Interestingly, our results also suggest that lenition generally does not correspond to the emergence of a new phonological target for production purposes, as is often assumed, e.g., [7]. Clear cases where lenition seems to reveal a more categorical nature are limited to one speaker and a single fricative. SP01's [ɣ], but, even in this case, the more lenited category accounts for a small proportion of the total tokens. Our data also reveal that speakers can produce tokens where the fricative gesture is fully absent. This finding is best interpreted as evidence that different exemplars, rich in articulatory detail, may coexist for the lexical representation of a single word [10]. Lenition is also modulated by speech rate, style, and stress. However, somewhat surprisingly, the magnitude, and even the direction of the effects, depend on the speaker and gesture in question. Furthermore, our data suggests that lenition as an articulatory process does not seem to correlate in a particularly strong fashion with changes in intensity that may facilitate listener perception, as hypothesized by some, e.g., [4]. This lack of correlation between articulatory lenition and its acoustic output does not necessarily imply that listener-oriented accounts of lenition are to be rejected. It is possible that the intensity measurement we took does not appropriately characterize changes in intensity for fricatives like it does for stops. Finally, lenition as an articulatory process correlates better with acoustic duration of the target segment, as hypothesized by [5]. Whether this reduction in duration is the rationale behind lenition or a consequence of it is a question that we leave open for a full-fledged model of voiced fricative lenition that we hope to develop in future work.

5. Conclusion

In this paper we have presented evidence that Icelandic voiced fricative lenition corresponds to a reduction in CD or, sometimes, to a full absence of the fricative constriction gesture. Furthermore, the process is suggestive of coexisting lexical representations, that display word-, speaker-, and gesture-specific variation. This is evident both in terms of the largely gradient nature of lenition and in the way it is modulated by other factors, such as speech rate, style, and stress. Finally, while lenition seems to correspond to a shorter acoustic duration of the target fricative, speaker- and gesture-specific effects seem to determine the extent to which lenition correlates with this dimension of the acoustic output.

6. References

[1] K. Árnason, *The phonology of Icelandic and Faroese*. Oxford University Press, 2011.

[2] P. Helgason, *On Coarticulation and Connected Speech Processes in Icelandic*. Reykjavík: Málvísindastofnun Háskóla Íslands.

[3] M. Schäfer, "Phonetic Reduction of Adverbs in Icelandic," *Promot. Am GRK*, vol. 1624, 2014.

[4] J. Kingston, "Lenition," in *3rd Conference on Laboratory Approaches to Spanish Phonology*, 2008, pp. 1–31.

[5] U. Cohen Priva and E. Gleason, "The causal structure of lenition: A case for the causal precedence of durational shortening," *Language*, vol. 96, no. 2, pp. 413–448, 2020.

[6] P. Boersma and D. Weenink, *Praat: doing phonetics by computer*. 2021. [Online]. Available: http://www.praat.org/

[7] B. Parrell, "Dynamical account of how/b, d, g/differ from/p, t, k/in Spanish: Evidence from labials," *Lab. Phonol.*, vol. 2, no. 2, pp. 423–449, 2011.

[8] B. Parrell, "Articulation from acoustics: Estimating constriction degree from the acoustic signal.," *J. Acoust. Soc. Am.*, vol. 128, no. 4, p. 2289, 2010.

[9] R. Kirchner, "Consonant lenition," in *Phonetically Based Phonology*, B. Hayes, R. Kirchner, and D. Steriade, Eds. Cambridge University Press, 2004, pp. 313–345. doi: 10.1017/CBO9780511486401.010.

[10] J. B. Pierrehumbert, "Word-specific phonetics," in *Laboratory Phonology 7*, C. Gussenhoven and N. Warner, Eds. De Gruyter Mouton, 2002, pp. 101–140. doi: 10.1515/9783110197105.1.101.