

Morphological influences on categorical perception of stop voicing in Welsh

Skye Anderson & Elise Bell*

Abstract. Listeners integrate a wide variety of cues when categorizing speech sounds, including lexical, syntactic, semantic and pragmatic information. We investigate the influence of Initial Consonant Mutation, a morphosyntactically-triggered alternation in the modern Celtic languages, on the categorical perception of stop voicing in Welsh. Once sandhi processes, Celtic mutations are now lexically and morphosyntactically triggered; in particular, Welsh Soft Mutation causes word-initial voiceless stops to become voiced when they are preceded by a triggering word or construction. This paper reports the results of a two-alternative forced choice task that tests the hypothesis that Welsh listeners integrate their knowledge of mutation-triggering environments during speech perception, accepting more ambiguous segments as voiced when preceded by a Soft-Mutation-triggering word relative to a non-triggering word. While the results of the experiment demonstrate categorical perception of stop voicing, no robust effect of mutation environment was found. Several hypotheses as to why the predicted result was not found are considered.

Keywords. categorical perception; voice onset time; morphology; initial consonant mutation; Celtic; Welsh

1. Introduction. Natural speech is characterized by extreme variability (c.f. Warner & Tucker 2011), requiring listeners to integrate a wide variety of cues when categorizing speech sounds in context. In addition to phonetic cues, listeners recruit lexical, syntactic, semantic and pragmatic information to adjust their perceptual boundaries. We extend this line of research by investigating the influence of Initial Consonant Mutation (ICM) on the categorical perception of stop voicing in Welsh. Consonant mutation in Celtic languages is a remnant of phonological sandhi processes such as intervocalic voicing. The conditioning environments have since disappeared and mutation is now lexically and morphosyntactically triggered. Welsh Soft Mutation in particular causes word-initial [p, t, k] to become [b, d, g] respectively, following a trigger-word. We hypothesize that Welsh listeners will exploit their knowledge of Soft Mutation-triggering words to adjust their VOT boundary when categorizing word-initial stops, judging more voiceless segments as voiced when preceded by a Soft Mutation trigger relative to a non-trigger. Section 1.1 provides an overview of Celtic ICM, focusing on Welsh Soft Mutation, and section 1.2 briefly reviews previous work on contextual factors influencing categorical perception.

1.1. INITIAL CONSONANT MUTATION. Celtic languages exhibit a rich system of morphosyntactically and lexically triggered word-initial consonant alternations (see for example Hannahs 2011 for an overview of the phenomenon in Celtic and Iosad 2010 for a broader survey of ICM in the languages of the world). ICM in Celtic is a remnant of sandhi processes which originated as phonologically predictable alternations, e.g. voicing between voiced segments, nasal place assimilation, etc. The phonological conditioning environments for these processes have largely

* The authors would like to thank Heddwen Brooks for serving as a native speaker consultant, all of the participants who were generous enough to donate their time, members of the University of Arizona Celtic Linguistics group, Dr. Natasha Warner, Jonathan Geary, and attendees of the 2018 Linguistic Society of America Annual meeting for useful discussion of this work. Authors are listed in alphabetical order. Authors: Skye Anderson, University of Arizona (skanderson@email.arizona.edu) & Elise Bell, University of Arizona (elisebell@email.arizona.edu).

disappeared in the modern Celtic languages and the initial consonant alternations are now lexically and morphosyntactically triggered.

Soft Mutation (SM) in Welsh comprises a set of initial consonant alternations triggered by some lexical items and grammatical constructions. In Soft Mutation environments, word-initial voiceless stops become voiced.¹ Table 1 shows example lemmas and their Soft Mutation forms.

Lemma		Mutated	Gloss
pabell	/pabɛl/	[i b a bɛl]	‘his tent’
tŷ	/ti/	[i d i]	‘his house’
cath	/kaθ/	[i g a θ]	‘his cat’

Table 1: Lemmas and their Soft Mutation forms following the trigger word *ei* [i] ‘his’.

The next section surveys previous work showing the wide range of contextual factors that can influence phonetic categorization and outlines the interest of ICM for this domain of inquiry.

1.2. CONTEXTUAL INFLUENCES ON SPEECH PERCEPTION. Listeners integrate information from wide range of linguistic and contextual domains during phonetic categorization, including lexical knowledge, with listeners categorizing an ambiguous token differently depending on whether the resulting percept is a real word or not. For example, Ganong (1980) presented listeners with VOT continua ranging between pairs like *kiss* : *giss* and *cush* : *gush*, where *kiss* and *gush* are real words. When presented with the former continuum (*kiss* : *giss*), listeners categorized more of the ambiguous stimuli as voiceless, while the latter continuum (*cush* : *gush*) elicited more voiced responses to ambiguous stimuli. This result indicates that listeners are biased to perceive real words (see for example Fox 1984, Connine & Clifton 1987, McQueen 1991 and Pitt 1995 for a discussion of the nature of the mechanism by which lexical information influences categorization and the implications of such a mechanism for models of spoken word recognition).

Syntactic restrictions can also influence phonetic categorization; Isenberg et al. (1980) found that English listeners presented with ambiguous tokens drawn from a continuum ranging between two function words (*to* : *the* [tə : ðə]) are sensitive to the phrase type indicated by the sentential context, e.g. the same ambiguous stimulus is more often categorized as *the* in the context *We tried X gold* and *to* in the context *We tried X go* because the noun *gold* is syntactically congruent with the definite article and the verb *go* with the preposition *to*. A similar effect was found for function word identification in Dutch (van Alphen & McQueen 2001, see Alphen).

Semantic and pragmatic congruity effects have been found as well: Miller et al. (1984) examined phonetic categorization of a *path* : *bath* continuum in three sentential contexts – (a) *She needs hot water for the X*, (b) *She likes to jog along the X*, and (c) *She is not thinking of the X*. The first is most congruent with *bath*, the second with *path*, and the third is semantically and pragmatically felicitous with either *path* or *bath*. When participants were required to respond by providing both the word which they had heard (*path* or *bath*) and the sentential context in which they had heard it, a semantic congruity effect arose, i.e. participants categorized more tokens as *bath* in context (a) and more as *path* in context (b). However, when participants were only required to respond with the word they had heard, the semantic congruity effect was neutralized,

¹ Other initial consonant alternations are triggered in Soft Mutation environments: Voiced stops are spirantized except for the voiced velar stop [g], which is deleted. Additionally, the voiceless alveolar trill [ʀ] becomes voiced, the voiceless lateral fricative [ɬ] becomes a lateral approximant [l], and the bilabial nasal [m] becomes a voiced labiodental fricative [v].

suggesting that the task itself must highlight the salience of the context in order for it to influence categorization.

More recently, Rhode & Ettliger (2012) asked whether pragmatic inferences about coreference are integrated when categorizing tokens along a *he* : *she* [hi : ʃi] continuum in English. In particular, they compared perception of ambiguous tokens from that continuum in sentences containing implicit causality verbs (e.g. *annoy*, *admire*, *impress*) that elicit listener inferences about coreference. Stimuli in the experiment contained sentences like *Malcom aggravated Natalie because X was too impatient*, where *aggravated* biases listeners to expect X to refer back to the subject (*he*) rather than the object (*she*) of the verb. They tested categorization of ambiguous tokens in *he*- and *she*-biasing sentences with both subject- and object-biasing verbs and found that sentences pragmatically congruent with *she* did in fact elicit more *she* responses and vice versa. This indicates that knowledge of typical patterns of coreference elicited by specific verbs influences phonetic categorization.

In summary, categorization of ambiguous speech sounds can be influenced by a number of factors, including lexical, syntactic, and pragmatic knowledge. The present research expands this line of questioning into an empirical domain which has not been explored, that of morphosyntactic alternations triggered across word boundaries. Specifically, the work presented here asks whether Welsh listeners use their morphosyntactic knowledge of mutation-triggering environments to adjust their perceptual boundaries when categorizing the voicing of following word-initial obstruents. Mutation phenomena provide an interesting empirical domain in which to test the bounds of non-phonetic factors that can influence speech perception because mutation allows for the creation of single-lexeme continua; for example, a [p : b] continuum can be constructed such that each end of the continuum represents a different surface realization of the same lexeme, e.g. *pabell* : *babell* ‘tent’. Furthermore, the existence of homophonous function words that trigger different surface realizations of the same lexeme allow for semantically equally well-formed phrases at both ends of the continuum that differ acoustically only in the target segment, e.g. *eu pabell* : *ei babell* [i pabɛl : i babɛl] ‘their tent : his tent’. Finally, the orthographic system of Welsh requires that such homophonous mutation triggers be distinguished orthographically, so homophonous stimuli can be presented with concomitant visual cues that should bias participants to process the identical auditory stimulus as either the mutation-trigger *ei* [i] ‘his’ or its non-mutation-inducing counterpart *eu* [i] ‘their’.² Based on the previous work surveyed above showing that listeners are biased to perceive sensible units of language, whether that means real words or syntactically or semantically felicitous phrases, we hypothesize that listeners will be biased to perceive morphologically well-formed utterances as well. Welsh listeners are predicted to categorize more ambiguous stops as voiced when they occur in a context that triggers Soft Mutation relative to a context that does not.

The remainder of the paper is structured as follows: Section 2 describes the experimental methods employed and section 3 presents the results. Finally, section 4 discusses possible interpretations of the results and outlines avenues for future research.

2. Methods

2.1. PARTICIPANTS. Twenty-six native speakers of Welsh (14 women, average age = 35.1 years) were recruited from online Welsh communities; 14 (9 women, average age = 34.7 years) were

² *Ei* ‘his’ and *eu* ‘their’ are pronounced homophonously as [i] in casual Welsh speech, although some speakers pronounce them differently, rendering *ei* ‘his’ as [ei] and *eu* ‘their’ as [ei]. King (2003) reports that this orthography-based pronunciation is non-traditional.

randomly assigned to the experimental condition reported here while the remaining 12 participated in an additional condition (not reported here) that examined Aspirate Mutation, another type of mutation in Welsh involving stop-fricative alternations.

Bilingualism with English is the norm for Welsh speakers and all participants in the present experiment were bilingual. All participants completed the Bilingual Language Profile (BLP) questionnaire (Birdsong et al. 2012), which consists of 19 questions about participants’ language history, patterns of use, proficiency, and attitudes about each language in question (Welsh and English in this case). Responses to the Bilingual Language Profile questionnaire are assigned weighted numerical values that are used to calculate a dominance score for each language. Calculating the difference between these values results in a composite numerical dominance score ranging from 218 (extremely dominant in Welsh) to -218 (extremely dominant in English). A score of 0 corresponds to balanced bilingualism in both languages. The Bilingual Language Profile provides a quantitative method of comparing language dominance (and factors contributing to language dominance) between and within populations of interest. It has been previously used in investigations of bilingual speech for language pairs including English, Spanish, and Catalan (Garraffa et al. 2017, Muxika-Loitzate 2017, Amengual 2016 and Simonet 2014, among others).

2.2. MATERIALS. The experiment comprised a two-alternative forced choice task in which listeners categorized the voicing of the initial stops of nonwords embedded in both an auditory and visual carrier phrase. In all trials, the auditory carrier phrase was the same, [i *blank*]. Half of the trials co-occurred with a visually-presented orthographic carrier phrase that triggers Soft Mutation (the possessive pronoun *ei* [i] ‘his’) and half were presented with a homophonous but orthographically distinct pronoun *eu* [i] ‘their’ that does not trigger Soft Mutation, as in Table 2.

	Carrier Phrase	
	Auditory	Visual
Soft Mutation	<i>i</i> _____	<i>ei</i> _____
No Mutation	<i>i</i> _____	<i>eu</i> _____

Table 2: Auditory and visual carrier phrases that accompanied the target words.

The nonword targets used in the experiment are shown in Table 3; they were recorded by a native Welsh speaker who read each nonword in both orthographic carrier phrases. Two tokens of each nonword were selected, one following the orthographic carrier *ei* [i] ‘his’ and one following *eu* [i] ‘their’. The recorded tokens were modified using the Praat Vocal Toolkit (Boersma & Weenink 2018, Corretge 2012) to fall along a 9-point continuum from -120 milliseconds of pre-voicing to 120 milliseconds of aspiration for each pair of homorganic stops, resulting in 54 unique stimuli (3 stop pairs × 2 words per stop pair × 9 continuum points). Each step of the continuum exhibited a 30 ms difference in VOT from the previous and subsequent steps.

	Voiceless	Voiced
<i>pabw</i>	[-abu]	<i>babw</i>
<i>pima</i>	[-ima]	<i>bima</i>
<i>tara</i>	[-ara]	<i>dara</i>
<i>tiro</i>	[-iro]	<i>diro</i>
<i>calo</i>	[-alo]	<i>galo</i>
<i>citw</i>	[-itu]	<i>gitw</i>

Table 3: Nonword target pairs used in the experiment.

Previous acoustic investigations of word onset stop voicing in Welsh have indicated that the fortis [p, t, k] and lenis [b, d, g] series of stops are distinguished by a difference in aspiration; the fortis series exhibits a greater period of aspiration than the lenis series (Ball 1984). The VOT continuum used here was designed both to satisfy the requirements of the present experiment and to establish a baseline for categorical perception of stop voicing in Welsh. As such, the edges of the continuum extended beyond the average VOT values found in the careful speech of a native Welsh speaker, and ceiling and floor effects were expected at the extreme points on the continuum.

A single recorded token of *ei* and another of *eu* were selected and appended to each item, yielding a total of 108 auditory stimuli (54 unique nonword tokens \times 2 frames). The visual stimulus displayed with each item always contained the same preceding token as the auditory stimulus (i.e. auditory stimuli beginning with the recorded token *ei* [i] ‘his’ were always paired with the *ei* visual stimulus). Though *ei* and *eu* are phonologically homophonous, this precaution was taken to ensure that any latent acoustic cues to the identity of the trigger word would be consistent with its visually-presented counterpart.

2.3. PROCEDURE. Responses were collected online using JSPsych (de Leeuw 2015, see Leeuw). Participants completed the experiment on their own computers at their leisure. We requested but could not enforce or verify that participants wear headphones during the task. Following completion of the language background questionnaire, a short training period took place in which feedback was provided after each trial. In each practice trial, participants heard a Welsh phrase (e.g. *eu pabell* [i pabeɫ] ‘their tent’ or *ei babell* [i babeɫ] ‘his tent’), and were instructed to use either the left or right arrow key to identify which of two written phrases (*ei pabell* or *ei babell*) they heard. Auditory stimuli used in the practice session were recorded by the same native Welsh speaker who recorded the test items. All training phrases consisted of common Welsh words (e.g. *pabell* [pabeɫ] ‘tent’, *tân* [ta:n] ‘fire’, *cath* [kaθ] ‘cat’) preceded either by the mutation trigger *ei* [i] ‘his’ or the homophonous word *eu* ‘their’, which does not trigger mutation.

Following completion of the practice trials, the experiment began. In experimental trials, onset of the auditory stimulus preceded the onset of the visual stimulus and, as in the practice task, subjects were provided with two written phrases as possible answers. They responded using the left and right arrow keys to indicate whether the phrase on the left or right better represented what they had heard. There was no time limit on participants’ responses and the next trial began only after a response had been provided.

3. Results and analysis. Because different places of articulation can have inherently different VOT boundaries due to the nature of the articulators used to produce them (Cho & Ladefoged 1999), the data were split by place of articulation (bilabial, alveolar and velar). Smoothed plots of the percent voiced (mutated) responses by place of articulation are shown in Figure 1 below. For each place of articulation, a Generalized Linear Mixed Model (GLMM) was fit to the data in R (R Core Team 2017) using the lme4 package (Bates et al. 2015). Response (voiced, voiceless, coded as 1 and 0 respectively) served as the dependent variable; the predictors were CONTINUUM STEP (1-9), visual FRAME (mutation-trigger *ei*, non-mutation trigger *eu*, reference level *ei*) and DOMINANCE (a value derived by centering and scaling the composite dominance score computed from the Bilingual Language Profile responses). Models included the interaction of CONTINUUM STEP and FRAME and random intercepts by subject. The more conservative alpha criterion of $\alpha = 0.01$ was selected to mitigate against Type I errors given that multiple comparisons were made.

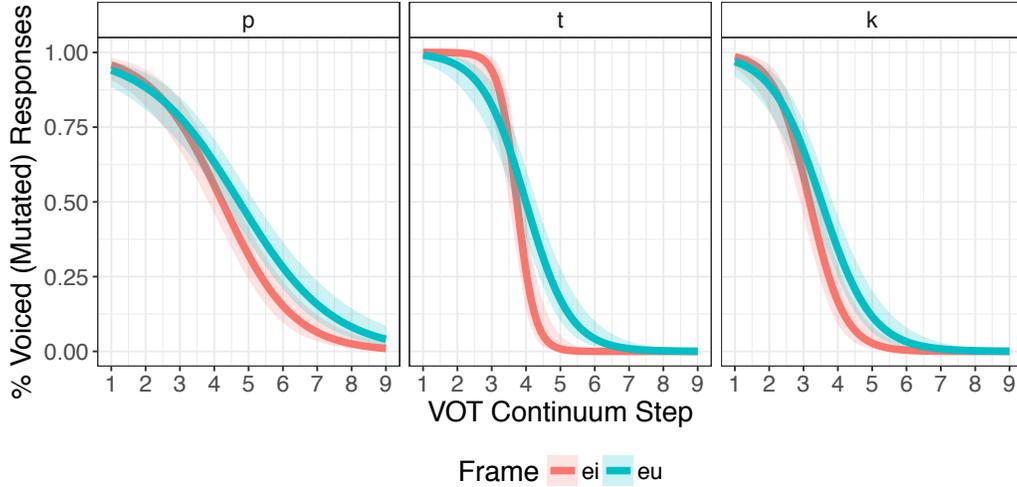


Figure 1. Smoothed plots of percent mutated (voiced) responses in the mutation and non-mutation contexts *ei* and *eu* for the bilabial, alveolar, and velar places of articulation.

The bilabial place of articulation exhibited a main effect of CONTINUUM STEP ($z = 8.51, p < 0.0001$), with participants categorizing more tokens as voiceless as VOT increased (see Table 4). LANGUAGE DOMINANCE exhibited a trend towards significance, with more Welsh dominant participants categorizing fewer stimuli as voiced overall, but the effect failed to reach significance at the specified alpha level of 0.01 ($z = -2.28, p < 0.025$; see the left pane of Figure 2 for a scatterplot of participants’ language dominance versus the percent of stimuli they categorized as voiced). Finally, there was no main effect of FRAME ($z = -0.86, p > 0.01$) and no interaction of FRAME and CONTINUUM STEP ($z = 1.63, p > 0.01$), indicating that participants’ VOT boundary for bilabial stops was not influenced by the orthographically-presented mutation context.

	Estimate	Std. Error	z	Pr(> z)
Frame (ref. level <i>ei</i>)	-0.62	0.71	-0.86	0.388
Continuum Step	-0.04	0.12	-8.51	0.000 ***
Language Dominance	-0.42	0.19	-2.28	0.023
Frame x Continuum Step	0.24	0.15	1.63	0.104

Table 4: Model summary for the bilabial dataset.

Similarly, the velar place of articulation exhibited no main effect of FRAME ($z = -1.15, p > 0.01$) and no interaction of FRAME and CONTINUUM STEP ($z = 1.69, p > 0.01$) (see Table 5). There was a main effect of CONTINUUM STEP as expected ($z = -6.65, p < 0.0001$), but unlike in the bilabial dataset, there was no main effect of DOMINANCE ($z = 0.25, p > 0.01$). Inspection of the far right pane of Figure 2 reveals that all participants rated the majority of velar stimuli as voiceless.

	Estimate	Std. Error	z	Pr(> z)
Frame (ref. level <i>ei</i>)	-1.36	1.19	-1.15	0.251
Continuum Step	-1.95	0.29	-6.65	0.000 ***
Language Dominance	-0.04	0.17	0.25	0.804
Frame x Continuum Step	0.58	0.34	1.69	0.092

Table 5: Model summary for the velar dataset.

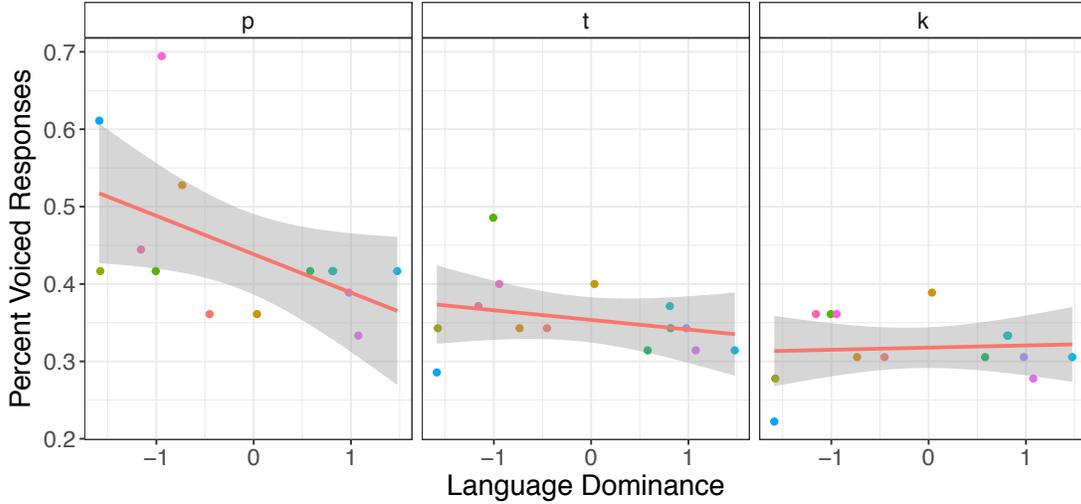


Figure 2. Scatter plots of subjects’ centered and scaled composite language dominance scores versus the percent of stimuli they categorized as voiced overlaid with a fitted linear model. A score of -1 reflects English dominance while 1 represents Welsh dominance; any given subject is represented by the same color in each pane.

For the alveolar place of articulation, on the other hand, there were significant main effects of FRAME ($z = 2.63, p < 0.01$) and CONTINUUM STEP ($z = -4.97, p < 0.0001$) (see Table 6). Additionally, the interaction of FRAME and CONTINUUM STEP was significant ($z = 2.86, p < 0.005$). Examination of the smoothed plot of percent voiced responses to alveolar stimuli in Figure 1 above indicates that perception was more categorical for identical auditory stimuli when they were accompanied by the mutation-triggering visual frame *ei* ‘his’, while stimuli accompanied by the non-mutation triggering *eu* ‘their’ were perceived as more ambiguous. As in the velar dataset, DOMINANCE had no significant effect ($z = -1.03, p > 0.01$) (see the middle pane of Figure 2 above).

	Estimate	Std. Error	z	Pr(> z)
Frame (ref. level <i>ei</i>)	-8.24	3.13	-2.63	0.009 *
Continuum Step	-3.97	0.80	-4.97	0.000 ***
Language Dominance	-0.26	0.25	-1.03	0.303
Frame x Continuum Step	2.33	0.81	2.86	0.004 **

Table 6: Model summary for the alveolar dataset.

Previous work on phonetic categorization has generally found contextual effects for mid-continuum ambiguous tokens, but not for unambiguous tokens at the tails of the investigated continua (see for example Ganong 1980 and Miller et al. 1984). This motivates analyzing the data further by removing unambiguous continuum steps and focusing on the continuum steps that participants perceived as ambiguous. Specifically, the step nearest the point at which 50% of responses were voiced was identified and two additional steps were retained on each side, discarding steps at the extremes of the continua. For the bilabial and alveolar datasets, this removed steps 1 and 7-9; for the velar dataset, steps 6-9 were removed.

After excluding those continuum steps, the interaction of FRAME and CONTINUUM STEP found in the full alveolar dataset failed to reach significance ($z = 1.15, p > 0.01$). Naturally, all

three places of articulation still exhibited a significant main effect of CONTINUUM STEP (bilabial: $z = -6.42, p < 0.01$; alveolar: $z = -4.82, p < 0.01$; velar: $-6.33, p < 0.01$), but no other comparisons reached significance at the alpha level of 0.01 (see Table 7 below for full summaries of the fitted models).

	Estimate	Std. Error	z	Pr(> z)
Bilabial				
Frame (ref. level <i>ei</i>)	-0.24	1.49	-0.16	0.873
Continuum Step	-1.69	0.26	-6.42	0.000 ***
Language Dominance	-0.40	0.21	-1.90	0.058
Frame x Continuum Step	0.06	0.35	0.18	0.104
Alveolar				
Frame (ref. level <i>ei</i>)	-3.46	3.54	-0.98	0.329
Continuum Step	-3.85	0.80	-4.82	0.000 ***
Language Dominance	-0.21	0.25	-0.83	0.405
Frame x Continuum Step	1.05	0.92	1.15	0.251
Velar				
Frame (ref. level <i>ei</i>)	-0.18	1.38	-0.13	0.898
Continuum Step	-1.92	0.30	-6.33	0.000 ***
Language Dominance	0.10	0.18	0.58	0.564
Frame x Continuum Step	0.21	0.41	-0.53	0.597

Table 7: Model summaries for the trimmed bilabial, alveolar, and velar datasets.

This result suggests that only the extreme ends of the VOT continua elicited differential categorization of stimuli depending on the orthographically presented (non-)mutation context, and only for the alveolar place of articulation. The next section summarizes and discusses these results and outlines avenues for future research.

4. Discussion and future directions. While this task successfully demonstrated categorical perception of stop voicing in Welsh using the online recruitment and experiment-presentation methods described in Section 2, it did not demonstrate any robust influence of the orthographically-presented morphological environment on phonetic categorization of the following target stop consonant. Contrary to previous work, the only effect of the contextual manipulation was found at the extreme ends of the VOT continuum, and only for the alveolar place of articulation.

Given the extensive evidence of adaptation of perceptual boundaries in the presence of informative lexical, syntactic and semantic cues, the failure to find an robust effect of mutation context here is surprising. However, some aspects of the experimental design may have mitigated against the appearance of such an effect. First, the task itself did not require participants to actively attend to the visually-presented cues; as such, the visual carrier phrases may simply have been ignored by participants. This would be consistent with Miller et al. (1984)’s finding that semantic congruity effects disappear if the task itself does not highlight the salience of the sentential context in which the target word occurs. Second, the items containing the target stop consonants consisted of nonwords; this decision was made to circumvent frequency-related confounds that can arise when using real words. However, borrowed words do not consistently undergo mutation in Welsh (Jones 2013, King 2003); as a result, nonword phrases exhibiting a technically ungrammatical lack of mutation following a mutation trigger, e.g. *ei citw* [i kitu] ‘his (nonword)’, may not have elicited a strong ungrammaticality judgment from participants.

Finally, the application of mutation in spoken Welsh is variable, meaning that Welsh speakers are occasionally exposed to examples of, for instance, the mutation trigger *ei* [i] ‘his’ followed by an unmutated word, e.g. **ei pabell* [i pabɛɭ] ‘his tent’. However, Soft Mutation is the most robust of the mutations in Welsh, and while mutation does not always carry functional load, in the case of the possessive pronouns, mutation is the only overt encoding of possessor gender and number, making the specific trigger used in this experiment amongst the most consistent available in the language.

Future work will investigate the role of mutation context on phonetic categorization by implementing a variant of the task reported here in which participants, rather than explicitly categorizing the voicing of the auditory stimulus, identify the possessor of the target noun, e.g. they will hear an auditory stimulus consisting of [i blank] and will be asked “Who does *blank* belong to?”, with written choices of the nominal pronouns *fo* [fo] ‘he’ and *nhw* [ɲu] ‘they’. Additionally, frequency-matched real word targets will be used in lieu of nonwords. We anticipate that these alterations to the experimental design may be sufficient to elicit evidence of the effect of morphosyntactic information on the categorization of stop voicing in Welsh.

References

- Alphen, Petra van & James M. McQueen. 2001. The time-limited influence of sentential context of function word identification. *Journal of Experimental Psychology: Human Perception and Performance*, 27(5). 1057–1071. <https://doi.org/10.1037/0096-1523.27.5.1057>.
- Amengual, Mark. 2017. Type of early bilingualism and its effect on the acoustic realization of allophonic variants: Early sequential and simultaneous bilinguals. *International Journal of Bilingualism*. <https://doi.org/10.1177/1367006917741364>.
- Ball, Martin J. 1984. Phonetics for phonology. *Welsh phonology: Selected readings*. University of Wales Press, Cardiff, UK.
- Bates, Douglas, Martin Mächler, Benjamin Bolker & Steven Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1). 1-48.
- Birdsong, David, Libby M Gertken & Mark Amengual. 2012. Bilingual language profile: An easy-to-use instrument to assess bilingualism. *Center for Open Educational Resources and Language Learning*. University of Texas at Austin.
- Boersma, Paul & David Weenink. 2018. Praat: Doing phonetics by computer [computer program]. Version 6.0.36. <http://www.fon.hum.uva.nl/praat/> (August, 2017).
- Cho, Taehong & Peter Ladefoged. 1999. Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics*, 27. 207-229. <https://doi.org/10.1006/jpho.1999.0094>.
- Connine, Cynthia M & Charles Clifton Jr. 1987. Interactive use of lexical information in speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, 2. 291–299.
- Corrette, Ramon. 2012. Praat vocal toolkit: Overview. <http://www.praatvocaltoolkit.com> (August, 2017).
- Fox, Robert Allen. 1984. Effect of lexical status on phonetic categorization. *Journal of Experimental Psychology: Human Perception and Performance*, 10. 526–540.
- Garraffa, Maria, Mateo Obregon & Antonella Sorace. 2017. Linguistic and cognitive effects of bilingualism with regional minority languages: A study of Sardinian-Italian adult speakers. *Frontiers in Psychology*, 8. 1907. <https://doi.org/10.3389/fpsyg.2017.01907>.
- Ganong, William F. 1980. Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance*, 6(1). 110-125.

- Hannahs, S.J. 2011. Celtic mutations. In Oostendorp, Marc van, Colin Ewen, Beth Hume & Keren Rice (eds.), *Blackwell Companion to Phonology*, Volume V. Boston: Wiley-Blackwell, 2807-2830.
- Iosad, Pavel. 2010. Right at the left edge: Initial consonant mutations in the languages of the world. In Wohlgenuth, Jan & Michael Cysouw (eds.), *Rethinking Universals: How Rarities Affect Linguistic Theory*. Mouton de Gruyter. 105-138.
- Isenberg, David, Edward CT Walker & Joan M Ryder. 1980. A top-down effect in the identification of function words. *The Journal of the Acoustical Society of America*, 68. S48. <https://doi.org/10.1121/1.2004759>.
- Jones, Christine. 2013. *Welsh Grammar You Really Need To Know*. Hodder & Stoughton.
- King, Gareth. 2003. *Modern Welsh: A Comprehensive Grammar*. Routledge.
- Leeuw, Joshua R de. 2015. JSPsych: A Javascript library for creating behavioral experiments in a web browser. *Behavior Research Methods*, 47(1). 1-12. <https://dx.doi.org/10.3758/s13428-014-0458-y>.
- McQueen, James M. 1991. The influence of the lexicon on phonetic categorization: Stimulus quality in word-final ambiguity. *Journal of Experimental Psychology: Human Perception and Performance*, 17. 433–443.
- Miller, Joanne L, Kerry Green & Trude M Schermer. 1984. A distinction between the effects of sentential speaking rate and semantic congruity on word identification. *Perception & Psychophysics*, 36(4). 329–337.
- Muxika-Loitzate, Oihane. 2017. Sibilant merger in the variety of Basque spoken in Amorebieta-Etxano. *Languages*, 2. 25. <http://dx.doi.org/10.3390/languages2040025>.
- Pitt, Mark A. 1995. The locus of the lexical shift in phoneme identification. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 21. 1037–1052.
- R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Rohde, Hannah & Marc Ettliger,. 2012. Integration of pragmatic and phonetic cues in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(4). 967-983. <https://dx.doi.org/10.1037/a0026786>.
- Simonet, Miquel. 2014. Phonetic consequences of dynamic cross-linguistic interference in proficient bilinguals. *Journal of Phonetics*, 43. 26-37. <https://doi.org/10.1016/j.wocn.2014.01.004>.
- Warner, Natasha & Benjamin V Tucker. 2011. Phonetic variability of stops and flaps in spontaneous and careful speech. *The Journal of the Acoustical Society of America*, 130(3). 1606-1617. <https://doi.org/10.1121/1.3621306>.