The Syllable in Domain Generalization: Evidence from Artificial Language Learning

Maya C. Wax Cavallaro
University of California, Santa Cruz

1 Introduction

Domain generalization is a term coined by Myers & Padgett (2014, from here on M&P) for a fairly widespread theory (e.g., Hyman 1978, Westbury & Keating 1986, Hock 1991, Barnes 2006, Padgett 2015) of the diachronic origins of certain domain-edge phenomena, such as word-final obstruent devoicing. While devoicing does not seem to have any phonetic grounding or motivation in word-final position, the phenomenon is often attributed to phonetic tendencies at the ends of longer phrases or utterances.[1]

In phrase-final position, phonetic pressures, such as a decline in the subglottal air pressure necessary for voicing over the course of an utterance, or a spreading of the vocal folds in anticipation of non-speech breathing, may cause variable final voicelessness (Ohala 1983, Blevins 2006). A domain-generalization-type account posits that this phonetic tendency can become phonologized and generalized from the phrase domain to the word domain. In this case, learners may interpret a tendency toward final voicelessness as a categorical phonological restriction. They then generalize this restriction from a ban on final voiced obstruents in utterance-final words to a ban on final voiced obstruents in all words.

M&P tested domain generalization empirically, through two artificial learning experiments. They show that, even given relatively short exposure, participants can learn an utterance-final restriction and generalize utterance-final devoicing to a word-level rule, applying the rule to utterance-medial words. While M&P and others (e.g., Blevins 2006) claim that phonological rules can be generalized from the phrase/utterance domain to word edges, they tend to leave for future work the question of whether there is something special about words that allows for this generalization.

M&P specifically acknowledge the possibility that utterance-final devoicing might be expected to generalize to syllable-final devoicing, though, "whether this is a good prediction [they] leave as an open question" (427). They also raise the possibility of a general preference for word-based phonology, citing patterns such as stress and vowel harmony that are typically associated with the word domain, rather than larger domains like the phrase. As discussed in Section 2 below, however, a number of phonological phenomena seem to show sensitivity to syllable structure.

This paper explores whether syllable-edge phenomena can, like word-edge phenomena, result from domain generalization. In the next section, I discuss similarities and differences between the word and the syllable as potential domains of generalization. In Section 3, I briefly summarize previous studies by M&P and myself. In Section 4, I present an artificial learning experiment, which shows evidence of generalization to the syllable, and in Section 5, I discuss the implications of these findings and potential future directions.

2 Generalization beyond the word

If generalization by analogy from larger prosodic domains (the phrase) to the word level does, in fact, occur, it seems that generalization beyond the word to smaller domains might also be possible. This is predicted by Blevins (2006) in an evolutionary account of phonological patterns, which states, "For sound changes originating with phrase-final laryngeal gestures and lengthening…the direction of final-devoicing sound patterns is predicted to be utterance > phrase > word > syllable" (140).

* I would like to thank Jaye Padgett, Ryan Bennett, Grant McGuire, Rachel Walker, Eric Baković, and the reviewers and participants of AMP2023 for their helpful feedback and guidance.

[1] In this paper, 'phrase' and 'utterance' will be used interchangeably.

Proceedings of *AMP 2023/2024*

Though Blevins (2006) does not give examples of generalization to the syllable, Zec & Zsiga (2022) do provide an example of generalization beyond the word to a smaller domain. They cite domain generalization as an account of patterns of High tone retraction in Štokavian varieties of Serbian. They claim that High tone retraction is motivated at the phrase level, in order to avoid tone crowding due to a phrase-final boundary tone. This phrase-edge process, they explain, has been generalized to word-final position and then, in some varieties, to foot-final position (80).

The syllable seems to be the domain of application of various phonological phenomena. In many languages, the consonants which may occur in coda position are restricted to a small subset. In languages such as Mandarin, for example, only nasal consonants may occur in codas (Faytak et al. 2020). "Lenition" patterns, such as /s/-debuccalization in many varieties of Spanish (e.g., /espaɳol/ → [e**h**paɳol] 'Spanish') may apply to syllable—rather than just word—edges. Different types of syllable-final laryngeal neutralization patterns—both diachronic and synchronic—are found in languages such as German, Korean, Maidu, and Thai (Lombardi 1995).Though German is often cited as a case of word-final obstruent devoicing, it has been claimed that word-internal, syllable-final obstruent devoicing occurs in both German and Dutch (e.g., Wetzels & Mascaró 2001).

A clear example of a syllable-level pattern comes from Tz'utujil (Mayan), which presents a typologically rare case of syllable-final approximant devoicing (Wax Cavallaro 2021, 2022).

(1) Approximant devoicing in Tz'utujil

a.	*jay*	[ja:**j̊**]	'house'	d.	*alnaq*	[a**l̥**naq^{h}]	'quickly'
b.	*uleew*	[ule:**w̥**]	'land'	e.	*q'oor*	[qʔɔ:**r̥**]	'masa/dough'
c.	*umul*	[umu**l̥**]	'rabbit'	f.	*warnaq*	[wa**r̥**naq^{h}]	'(s)he has gone to sleep'

Though this devoicing occurs word-internally, it is clearly a positional (domain-final) phenomenon, rather than an assimilatory one. In examples such as (1d) and (1f), sonorant consonants devoice in codas while adjacent on both sides to voiced segments. It is plausible that the same phonetic tendency towards phrase/utterance-final voicelessness typically cited as the motivation for word-final devoicing could also lead to the phonologization and generalization of final sonorant devoicing. If this is the case, however, Tz'utujil would show generalization beyond the word level, to the syllable.

For learners to generalize from the phrase level to the word level and beyond, it seems they must somehow interpret a pattern applying to segments at the phrase edge as a pattern applying to a different, smaller domain (e.g., word or syllable edges). If a pattern like final devoicing originates at the phrase level, learners must hear examples of words with and without final devoicing at the early stages of the change. M&P note that learner-directed speech tends to consist largely of shorter, even single-word, utterances. Utterance-final tokens may be, they suggest, overrepresented in the learner's lexicon or memory. If a high enough proportion of words encountered are utterance-final, the restrictions/alternations applying to utterance-final words may be interpreted as patterns applying to all words.

This type of frequency-based calculation fits well within something like Exemplar Theory (Johnson 1996, Pierrehumbert 2000). Exemplar Theory proposes that learners store experienced tokens of language in memory with an amount of phonetic detail that captures a great deal of variability. Decisions in the identification and production of sounds, words, etc. can be based on calculations over these stored forms. An alternative to the stored exemplar approach would be something like the Gradual Learning Algorithm (Boersma 1997), where experienced tokens of language do affect and update a learner's grammar (by promoting/demoting constraints) without being encoded in long-term memory. I will remain as theory-neutral as possible, while acknowledging that the grammar needs access to a unit over which to generalize.

If generalization to the syllable is possible, this requires information about syllable structure to be either stored (perhaps in separate exemplar clouds for syllables, or as part of the structure within stored words/utterances) or, minimally, accessible to the grammar for frequency calculations and derivation of phonological rules/patterns. In Section 4 below, I show experimental evidence that this could, indeed, be the case, but there are also differences between the word domain and the syllable domain that merit discussion.

2.1 *The syllable as different from the word* Words exist as both morphosyntactic and phonological units, though prosodic word and morphosyntactic word boundaries do not always align perfectly (see, e.g.,

Ito & Mester 2009). The syllable, on the other hand, is a purely phonological domain. This presents a potential point of departure in two different ways.

The first has to do with the structures over which frequency-type calculations can be made. To generalize over words or syllables, the grammar must have access to those units. The fact that words are somehow stored in the mind is relatively uncontroversial. Whether one believes in an entirely abstract mental lexicon, a complex map of exemplar clouds, or a combination or variation on these types of representations, words are a unit that is stored in memory and accessible to the grammar.

The status of the syllable, on the other hand, is more contentious. Many phonologists have argued that syllable structure is not part of phonological representation. A great deal of work in Optimality Theory (Prince & Smolensky 1993), for example, claims that syllable structure is determined by phonological constraints but not contained in underlying representations. Ohala & Kawasaki-Fukumori (1997) and Steriade (1999) propose that apparent syllable-level patterns can be inferred from phonetic factors such as coarticulation and licensing by cue, and that reference to the syllable is not necessary.

Others, however, defend the syllable. Vaux (2003), for example, argues for "a theory of the lexicon in which some, but not all, predictable information is stored." Providing evidence from Armenian, he claims some information about syllable structure must be present underlyingly. To assertions that syllables do not exist in the grammar, Vaux replies, "My response is that we do not yet fully understand the psychological mechanisms involved in accessing information stored in the brain, nor do we understand how these mechanisms interact with the performance abilities required to execute tasks such as marking parts of words or dividing them in parts…" (118).

M&P point out that, given an exemplar-type model where experienced tokens are stored in memory, there are more words than utterances available to the learner because all utterances consist of words. If generalizations are inferred over encountered items, this may lead to a preference for word-level phonology over phrase-level phonology. The same logic could apply to syllables. All words are made up of syllables (but not vice-versa). This means, theoretically, that the learner has access to more syllables than words, assuming that syllable structure is encoded within representations of words.

Another possibility is that, whether or not the syllable is accessible to the grammar, it is not a potential domain to which generalization may occur. In work on "the life cycle of phonological processes," Bermudez-Otero & Trousdale (2012) appeal to phonologization and domain narrowing in a way that is both quite similar to and significantly different from that discussed in M&P. Work of Bermudez-Otero and colleagues (see also Bermudez-Otero & McMahon 2006) takes a stratal-cyclic approach based in Lexical Phonology to explain phonological changes such as post-nasal /g/-deletion in English.

While this approach does allow for—and even posits—generalization from the phrase level to the word level and beyond to a smaller domain, the domains in question are more morphosyntactic, rather than purely prosodic. Burmúdez-Otero and colleagues propose that phonological rules, even those affecting syllable codas, may apply at the phrase stratum, the word stratum, or the morphological stem stratum. According to this approach, a pattern like phrase-final devoicing would not generalize from phrase-final position to syllable-final position, though it might generalize to the morphological stem level.

Given evidence of generalization from the phrase level to the word level (M&P, Zec & Zsiga 2022, etc.) and the existence of syllable-level phenomena (including patterns like syllable-final devoicing in Dutch, German, and Tz'utujil, that could plausibly result from the generalization of phrase-level phonetic tendencies), it seems likely that learners may generalize beyond the word domain to the syllable. At the same time, the fact that the syllable is a purely phonological domain means that syllables may be stored differently in memory or may not be accessible to the grammar in the same ways as words. The experiment presented below tests whether this is the case, presenting evidence that generalization to the syllable is, in fact, possible.

3. Domain generalization in artificial language learning

3.1 *Myers & Padgett's (2014) study* M&P present two artificial grammar learning (AGL, also known as miniature grammar learning or artificial language learning) experiments. The AGL paradigm allows researchers to control the structure of the language material to which learners are exposed in order to observe the learning that occurs. Participants are exposed to a constructed mini-language during a learning phase. The linguistic knowledge acquired during this phase is measured during a testing phase.

This type of experiment has been employed to observe both syntactic and phonological learning. While the timescale of the learning phase may vary from a few minutes to a few weeks, it is definitely a shorter exposure than in natural language learning. Despite the short duration of this type of experiment, as well as the artificial nature of the "language" and learning task, AGL has been able to provide convincing insight into how humans acquire linguistic knowledge.

M&P utilized the poverty of the stimulus methodology in AGL (Wilson 2003, 2006, Finley & Badecker 2009, White 2013) to test whether learners would generalize from a limited speech sample. In this type of experiment, there are classes of stimuli present in the test phase that were not included during the learning phase, though they are related to classes present in the learning phase. To test for generalization of final obstruent devoicing, for example, M&P included both voiced and voiceless obstruent onsets in their stimuli during both phases of their experiments. Utterance-finally, however, only voiceless obstruents (in addition to vowels and nasals) but not voiced obstruents could occur. This was the pattern participants were meant to learn. Utterance-medially, however, target words ended only in vowels or nasals. This is the poverty of the stimulus: Participants were not given any evidence about obstruents in utterance-medial, word-final position during the learning phase.

During the testing phase, M&P presented participants with new stimuli and asked them to judge whether each item belonged to the language to which they had been exposed. This method proved effective, showing that participants could learn the utterance-final pattern and generalize it to novel, non-final words.

In their first experiment, they trained one group on final devoicing (only voiceless obstruents allowed utterance-finally) and a second group on a less-natural final voicing pattern (only voiced obstruents allowed utterance-finally). The only obstruents in the experiment were sibilant [s] and [z]. Both groups demonstrated evidence of learning and generalization, though the final devoicing group did so more successfully, showing an effect of naturalness.

To capture more aspects of natural language learning, M&P incorporated a larger set of obstruents ([p,b,t,d,k,g]), as well as morphologically conditioned alternations in their second experiment. They again found evidence of both learning and generalization. They also found somewhat weaker evidence that learners have an overall bias against accepting more than one form for a given item, but that such alternating forms may assist in the learning of patterns such as final devoicing. While there is still work to be done on naturalness and alternations, the focus of my current work is on learning and generalization, and whether learners can generalize to the syllable domain as they generalized to the word domain in M&P.

3.2 *Building upon M&P* I have performed a series of AGL experiments (Wax Cavallaro 2023 and forthcoming). The first of these sought to replicate M&P's first experiment with a few changes. While M&P's experiment was run in a sound-attenuated booth in the UC Santa Cruz Phonetics Laboratory, using Superlab experiment presentation software, my experiments were carried out online using PCIbex (Zehr & Schwarz 2018). This was mainly due to safety concerns around the COVID-19 pandemic.

Rather than a voicing neutralization pattern, the stimuli presented a liquid [l~r] neutralization pattern. The concern was that generalization of a voicing neutralization pattern to word-internal syllable codas could be confounded by participants' L1 English phonology and a dispreference for voicing mismatches in adjacent obstruents. My experiment showed that, as with the devoicing and voicing patterns in M&P, participants were able to learn both [l]-to-[r] neutralization and [r]-to-[l] neutralization at the utterance level and generalize the pattern to the word level.

To test for generalization to the syllable level, the materials and procedure needed to be changed significantly. The experiment presented in the next section was designed to test for generalization to the syllable, and to re-test for generalization to the word level, using identical procedures and stimuli. This way, differences in performance between participants generalizing to the word level and those (potentially) generalizing to the syllable level could not be attributed to slight differences in experiment design and stimuli.

4. Experiment

To test for both types of generalization (utterance-to-word and word-to-syllable) in this experiment, participants were broken into two groups, which heard identical stimuli and performed the same learning and testing tasks. The only difference between the groups was in the instructions they were given and, thus, what they believed they were hearing and learning.

One group was told repeatedly that they were listening to *three-word sentences* in a made-up language. This group is referred to from here on as the utterance-to-word (UtW) group because it was hypothesized that they would generalize from utterance-final position to word-final position. The other group was told repeatedly that they were listening to *three-syllable words* in a made-up language. This group will be called the word-to-syllable (WtS) group because it was hypothesized that they would generalize from the word level to the syllable level. The procedure is outlined further in Section 4.3.

4.1 *Participants* Participants were all undergraduate students at UC Santa Cruz, who were granted course credit in exchange for their participation. Due to the nature of the subject pool, participants could not easily be prescreened or disqualified. While M&P had included only native English speakers in their study, more than a quarter of the participants recruited for this experiment were native speakers of languages besides English. Rather than exclude all their data, a decision was made to exclude data only from participants who reported that English was not their *dominant* language. This left 71 participants: 36 in the WtS group and 35 in the UtW group. Participants ranged in age from 18-29 years old and were mostly California natives.

4.2 *Materials* Stimuli consisted of tri-syllabic nonce utterances (to be interpreted as either words or sentences, depending on the group), of one of the two shapes in (2), where C belonged to the set [t,k,m,n,l,ɹ] and V belonged to the set [i,a,u]. Target items were not presented in carrier sentences, as in M&P, but they were grouped into two conditions, shown in (2), based on location of the coda consonant within the stimulus.

(2) a. *final* .CV.CV.CV**C**.
b. *medial* .CV.CV**C**.CV.

Target coda consonants are bolded and underlined for clarity.

There was a concern that some obstruent codas in the medial frame (2b) would be perceived as C1 in an onset cluster (when followed by liquids), rather than codas, due to participants' English phonotactics. It was decided that this affected a small enough percentage of the stimuli that participants should still be able to learn that codas are allowed in word/utterance-medial position. Obstruent codas are, otherwise, not part of the analysis, so this issue should not affect results.

Participants were presented with an /ɹ/ → [l] neutralization pattern in the learning phase. Each consonant occurred with equal frequency in onset position, but [ɹ] never occurred in codas, and [l] occurred only in final codas (2a). Participants, therefore, never heard liquid codas word/utterance-internally during learning. The stimuli for the learning stage are given in (3), and those for the testing stage are given in (4).

(3) Stimuli used during the learning phase

Block 1 .CV.CV.CVC.			Block 2 .CV.CVC.CV.	
nasal-coda	obs-coda	[l]-coda	nasal-coda	obs-coda
.ni.ni.man	.ti.nu.lut.	.ka.na.nil.	.ti.mam.ki.	.ra.muk.na.
.ki.la.kum.	.nu.ti.rak.	.ta.mu.mul.	.na.kin.ka.	.ru.kik.mu.
.lu.ki.kan.	.ra.ru.lak.	.lu.ta.nal.	.la.lan.ra.	.li.nat.li.
.ta.ki.nim.	.ki.ma.rut.	.ra.lu.mil.	.ku.nim.li.	.mu.tat.nu.
.nu.nu.tin.	.mu.ku.lik.	.na.ti.nil.	.ti.kum.ka.	.li.nut.ku.
.tu.la.lan.	.na.la.tuk.	.ki.ri.rul.	.ka.run.ma.	.na.nak.ta.
.ri.ka.kim.	.mu.ti.tat.	.ti.mi.kil.	.ti.mun.li.	.na.rik.ta.
.ra.mi.kin.	.ra.lu.tak.	.ka.ma.tal.	.ma.lum.tu.	.mi.tat.ni.
.ri.li.kam.	.ti.ku.mit.	.mu.ra.lal.	.la.tun.mu.	.ki.lik.mu.
.nu.ku.rin.	.mi.tu.ruk.	.li.ni.nul.	.mu.run.ti.	.ru.mit.li.
.lu.na.lum.	.ka.ru.tut.	.mu.ta.mil.	.ki.ram.nu.	.tu.lit.ri.
.la.ra.num.	.mi.ri.mat.	.lu.mu.rul.	.nu.kam.ra.	.ru.tik.ru.

Each item in (3) and (4) was recorded by the author, a linguistics PhD student and native speaker of American English. Stimuli were intentionally pronounced with English-like vowels and liquids and English-like stress

(but without vowel reduction). All utterances were produced with declarative intonation and final stress that could be interpreted either as word-final stress (for the WtS group) or sentence stress (for UtW).

(4) a. Stimuli used during the testing phase – Block 1

Block 1 .CV.CV.CVC.			
nasal-coda	obst-coda	[l]-coda	[ɹ]-coda
.mu.lu.tim.	.ta.ti.kat.	.ma.nu.mul.	.ni.ki.mir.
.la.nu.run.	.tu.ri.lak.	.ku.la.nul.	.ki.li.tar.
.nu.ta.lan.	.ti.nu.tit.	.lu.mi.ril.	.li.ra.kar.
.ri.ki.tun.	.lu.ka.muk.	.ka.la.kul.	.ri.ma.tir.
.ti.ti.tam.	.mi.ni.nit.	.ta.ka.lil.	.lu.la.kir.
.li.ki.nam.	.ri.ta.mit.	.mu.mu.kul.	.na.na.tur.
.ma.mi.mun.	.ma.ta.kak.	.ra.tu.ril.	.ku.ra.lar.
.li.li.kum.	.na.ru.nak.	.ka.lu.kal.	.lu.ta.mur.
.ma.ra.rim.	.ra.mu.lut.	.ti.nu.kal.	.tu.mi.rir.
.li.ki.rin.	.ki.na.mik.	.ki.mu.kal.	.ru.ta.tar.
.na.ra.nin.	.nu.ri.tut.	.ti.ru.ral.	.ki.tu.mur.
.mu.nu.lam.	.tu.mi.mat.	.mi.mi.lil.	.ri.mi.lar.
.mi.la.nin.	.la.mu.mit.	.tu.ka.rul.	.nu.ka.tir.
.tu.nu.tum.	.ra.ri.kik.	.ma.mi.lul.	.ra.ru.lir.
.na.ru.nun.	.ru.la.nak.	.ku.li.nul.	.mi.ku.nar.
.ku.ni.ram.	.ra.na.lut.	.ti.lu.lil.	.lu.ki.tar.
.la.li.nam.	.na.ru.nuk.	.ki.ku.mil.	.na.tu.mur.
.ri.tu.run.	.na.ka.rak.	.nu.na.rul.	.ku.ti.kir.

b. Stimuli used during the testing phase – Block 2

Block 2 .CV.CVC.CV.			
nasal-coda	obs-coda	[l]-coda	[ɹ]-coda
.ka.tum.na.	.ti.ruk.tu.	.ri.tal.ti.	.tu.tar.ni.
.ni.tan.ru.	.ni.lak.mu.	.li.tul.mi.	.mi.mar.na.
.ra.kim.ku.	.ta.tik.ri.	.ki.tal.ni.	.mi.rur.ku.
.li.rin.ri.	.li.lut.ra.	.mu.ral.ma.	.ka.nar.nu.
.ru.mim.ta.	.ri.mik.ri.	.lu.tul.ti.	.nu.rir.ki.
.mu.lim.tu.	.ma.nat.mu.	.ru.nil.ma.	.ru.kir.tu.
.ni.mun.tu.	.na.mit.ki.	.na.tal.ni.	.mu.nur.ki.
.ru.rim.lu.	.ra.luk.ru.	.ta.lil.ta.	.mi.lur.nu.
.mu.man.ti.	.ti.nut.ku.	.ma.nil.ku.	.na.nur.ku.
.la.mun.ka.	.lu.kit.ku.	.li.lul.nu.	.ki.nar.la.
.ra.ram.lu.	.li.nak.li.	.nu.lal.ti.	.tu.mar.la.
.la.lin.ma.	.lu.rat.ma.	.li.mul.ku.	.na.kar.ni.
.ta.kim.la.	.nu.rik.li.	.ni.nal.na.	.ka.tur.la.
.ru.rin.ri.	.ru.rit.ki.	.ti.ril.ma.	.ta.nar.ma.
.ru.mam.ru.	.ka.muk.ta.	.mu.lul.ri.	.ka.kir.ti.
.na.man.la.	.ki.kut.li.	.ku.nul.mu.	.ki.kur.ma.
.ti.kum.lu.	.ta.tit.nu.	.ku.lal.ri.	.tu.kir.na.
.la.kun.la.	.ma.kak.ra.	.ku.tul.ri.	.mi.lir.ma.

4.3 *Procedure* Materials including text instructions were presented to participants online on their own personal computers using PCIbex. Participants were briefed in groups of 1-4 via video conferencing on Zoom before they were given the link to the experiment. This allowed me to provide both print and verbal instructions, answer any questions, and check that participants had an appropriate setup in terms of equipment (computer and quality, wired headphones) and a quiet space in which they could hear clearly and focus for the duration of the activity (about 30-45 minutes). Because the participant were students, some information

about the experiment and what it was intended to observe was provided for educational purposes upon completion of the testing phase and the exit survey.

During the learning phase, participants were told they would be listening to either sentences (UtW group) or words (WtS group) in a made-up language. They were instructed to repeat each word/sentence aloud to get a feel for how the language sounds. This phase was self-paced—participants clicked a button on their screen to hear the next item, though they could only hear each utterance once per trial. Stimuli were presented in blocks by coda position, with the final-coda (2a) condition presented first and medial-coda (2b) condition second. Each block was repeated three times, with the items presented in a different randomized order each time. The learning phase comprised 180 total trials (three presentations of the 60-utterance learning set). Participants were given the opportunity to take a short break between blocks and between the learning phase and the testing phase.

During the testing phase, participants were presented with a new set of stimuli, shown in (4). Rather than repeat stimuli aloud, as in the learning phase, participants were instructed to listen to each word or sentence and determine whether it belongs to the language they had been learning. To answer 'yes' (i.e. to *accept* the item), participants could click a button on their screen or press the 'd' key on their keyboard. To answer 'no' (i.e. to *reject* the item), participants could click a button or press the 'k' key on their keyboard. The 'yes' button was always presented on the left, and the 'no' button on the right. As in the learning phase, this phase was self-paced. Participants could hear each item only once, and the next stimulus was presented after they selected 'yes' or 'no'.

Again, stimuli were presented in blocks by coda position. Following M&P, this was done so that participants would compare items within each coda position condition, to prevent them from making their acceptability judgments based on the position of the coda alone. Medial-coda (2b) items were less frequent in the learning set (because only nasal-coda and stop-coda—but not liquid-coda—target items were presented), so there was concern that participants would perceive the medial-coda condition as less well-formed than the final-coda condition on the basis of relative frequency. Each block was presented one time during this phase, with the final-coda block presented first. The order of the stimuli within the block was randomized. Participants were given the opportunity to take a short break between the two blocks.

After the testing phase, participants were given a debriefing questionnaire, which asked what they believed the experiment was investigating and what strategy they used to decide which sentences to accept/reject. There was a concern that the student participants, who all had some prior training in linguistics, might be more consciously aware of phonological patterns than average. Most of these students, however, had taken only introductory-level linguistics courses, and all were instructed not to attempt any analysis. Comments on their exit surveys did not explicitly mention a restriction on final liquids, suggesting that learning was relatively implicit, though a few participants did mention "R" or the right edge.

As often as possible, both verbal and written instructions referred to either '(three-word) sentences' or '(three-syllable) words,' depending on the learning group. The buttons participants clicked to hear the next item during the learning phase said either, 'Next Sentence,' for the UtW group, or, 'Next Word,' for the WtS group. The question on screen after each item in the testing phase was either 'Does this sentence sound like it belongs to the language you have been learning?' or 'Does this word sound like it belongs to the language you have been learning?' Judging by comments in the exit survey, the UtW group believed they were listening to three-word sentences, and the WtS group believed they were listening to three-syllable words.

To improve data quality and encourage participants to thoroughly read and follow instructions, three checks were included. After the first block of the learning phase, participants were asked whether they had been repeating the items (words or sentences) aloud and reminded that doing so is important. At two points during the testing phase, participants were presented with a simple yes-no question as an attention check. It was decided ahead of time that only failure of *both* attention checks would result in a participant's responses being excluded from analysis. Seven participants failed a single attention check, but nobody failed both.

4.4 *Hypotheses* All participants were exposed to the same pattern of liquid neutralization where [l] but not [ɹ] occurred in final position (and neither liquid occurred in word-medial codas) during the learning phase.

Hypothesis 1: Learning. If the participants in this experiment learn the pattern they are exposed to during the learning phase, their 'accept' responses should reflect the pattern in the final condition (2a) in the testing phase. This means that participants in the UtW group will accept test items with sentence-final [l] more frequently than those with sentence-final [ɹ], and participants in the WtS group will accept test items with

word-final [l] more frequently than those with word-final [ɹ], showing an effect of the learning set on participants' responses during the testing phase.

Hypothesis 2: Generalization. The learning phase does not provide participants with evidence about which liquids are allowed in word/sentence-medial codas. If participants generalize the pattern they learned from the right edge of the larger domain (word or sentence) to the smaller domain (syllable or word), then the pattern should hold for test items in the medial condition (2b). This means participants in the UtW group will accept utterance-medial, word-final [l] more frequently than they accept utterance-medial, word-final [ɹ], and participants in the WtS group will accept word-medial coda [l] more frequently than they accept word-medial coda [ɹ]. This would show generalization of the learned pattern to the smaller domain.

Hypothesis 3: Syllable as domain. If the syllable and the word are both domains to which phonological patterns can be generalized, then both groups should show evidence of learning and generalization. A difference in performance between the two groups could suggest differences in how utterances, words, and syllables are treated in the grammar. Specifically, if both groups show evidence of learning but only the UtW group shows generalization to medial position, that would suggest that generalization to the syllable is not possible (at least in this AGL experiment) or involves a different process than generalization to the word.

4.5 *Results* After excluding only non-English-dominant participants, there were a total of 71 participants (36 WtS and 35 UtW). Only the testing phase utterances with coda liquids were included in the analysis. For each participant, there were 18 observations per condition (18 final [ɹ] codas, 18 final [l] codas, 18 medial [ɹ] codas, and 18 medial [l] codas). The percentage of 'accept' responses is presented in Figure 1, by liquid coda consonant, learning group, and position of the target coda in the utterance/word.

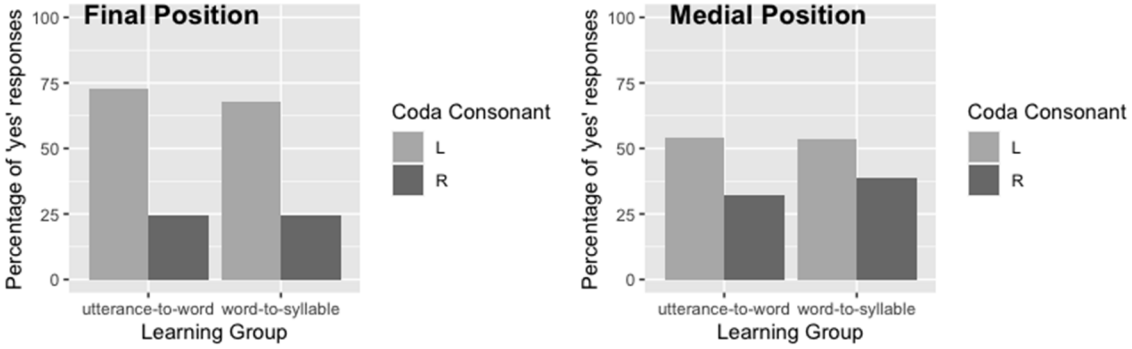

Figure 1: Percentage of 'yes,' or 'accept,' responses by liquid consonant and learning group: (left) final position; (right) medial position (all participants).

Qualitatively, there appears to be clear evidence of learning in both groups, with participants accepting substantially more [l] codas than [ɹ] codas in final position. In medial position, there is still a preference for [l] over [ɹ] codas in both groups, though this preference is weaker than the final condition.

Not all participants, however, showed evidence of having learned the desired pattern. It was determined that participants who did not have a difference of at least four (out of a possible 18) between the number of [l] codas they accepted and the number of [ɹ] codas they accepted in the final condition did not show a preference for [l] over [ɹ] and, therefore, did not learn the intended pattern from the learning phase. While this could happen for many reasons and is not particularly concerning, it would not be worthwhile to include these participants in any analysis of generalization, as they cannot generalize a pattern they did not learn. After excluding the non-learners, 30 participants remained in the UtW group and 31 remained in the WtS group. The percentage of 'accept' responses for this smaller set of participants is presented in Figure 2, by liquid coda consonant, learning group, and position of the target coda in the utterance/word.

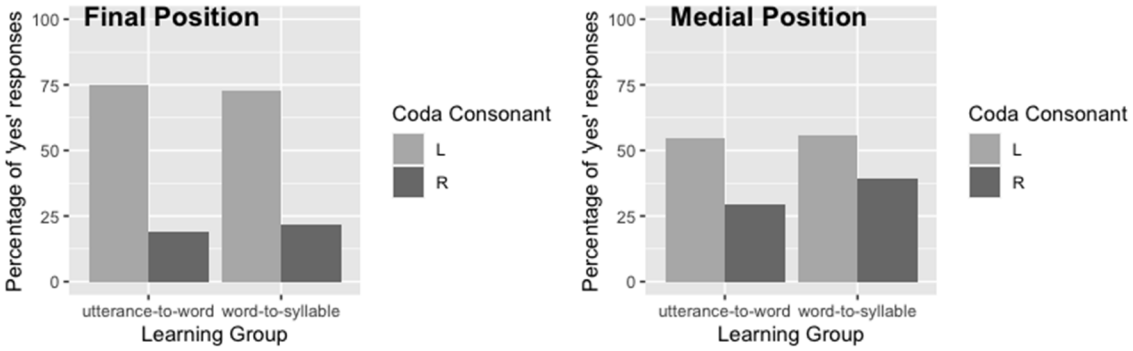

Figure 2: Percentage of 'yes,' or 'accept,' responses by liquid consonant and learning group: (left) final position; (right) medial position (learners).

The data were modeled by means of a mixed effects logistic regression analysis. The dependent variable was the response ('no' or 'yes') to the question, *Does this sentence/word sound like it belongs to the language you have been learning?* 'Yes' was treated as the marked value of the response variable (coded as 1), and 'no' as the default (coded as 0). The fixed effects were the liquid coda consonant (L or *R*), utterance/word position of the target coda (final or *medial*), and learning group (UtW or *WtS*). In each of these factors, the first level was treated as the default value and the second (in italics) as the marked value. Random effects included individual participants and items. Intercepts were included in the model for both of these random effects, as well as a slope for Coda consonant for each participant. The results for the fixed effects are given in Table 1.

	b	*Z*	*p*
Intercept	1.2457	6.307	2.84e-10*
Coda consonant ([ɹ])	-3.0106	-10.518	< 2e-16*
Coda position (medial)	-1.0295	-5.222	1.77e-07*
Learning group (WtS)	-0.1021	-0.427	0.669
Coda consonant × Coda position	1.7204	5.980	2.23e-09*
Coda consonant × Learning group	0.1912	0.550	0.582
Coda position × Learning group	0.1454	0.743	0.458
Coda consonant × Coda position × Learning group	0.2198	0.754	0.451

Table 1: Fixed effects in a logistic regression model of responses.
Significant values ($p < 0.05$) are indicated by *.

There were significant main effects of Coda consonant and Coda position but not Learning group. There was also significant interaction between Coda consonant and Coda position but there was no significant interaction involving Learning group, suggesting that the UtW group and the WtS group behaved similarly. To further explore these main effects, the data were broken into subsets according to Coda position and Learning group. Table 2 gives the results for four tests, with the same random effects structure as in the previous analysis, and the single fixed effect of Coda consonant.

Coda position	Learning group	*b*	*z*	*p*
final	UtW	-2.8886	-10.387	< 2e-16*
final	WtS	-2.9561	-7.583	3.37e-14*
medial	UtW	-1.3102	-4.856	1.2e-06*
medial	WtS	-0.9029	-4.076	4.57e-05*

Table 2: Effects of Coda consonant in data subsets defined by Coda position and Learning group.
Significant values ($p < 0.05$) are indicated by *.

There was a significant effect of liquid coda consonant ([l] vs. [ɹ]) on responses for both learning groups in the final condition, showing that participants learned the pattern presented in learning phase (that [l] but not [ɹ] may occur in codas). The significant effect of Coda consonant in the medial condition in both groups shows that this pattern was generalized to the smaller domain (the word domain in the UtW group and the syllable domain in the WtS group).

To test whether one of the groups was more successful at learning or generalizing the liquid neutralization pattern, the responses were recoded in terms of correctness, where a response was treated as 'correct' if it corresponded to the learning set pattern, and incorrect otherwise. In other words, it was 'correct' to accept coda [l] and to reject coda [ɹ].

	final	medial
UtW	78%	63%
WtS	75%	58%

Table 3: Percentage of correct responses by Learning group and Coda position.

In both groups, there was a higher percentage of correct responses in the final condition. There was also a higher percentage of correct responses in the UtW group than in the WtS group.

The effect of Coda position was found to be significant in a mixed model regression analysis with the correctness of the response as the dependent variable. Random effects intercepts were included in the model for both participant and item, with a slope factor for Coda consonant within participant. The fixed effects were Learning group and Coda position. Only Coda position had a significant main effect. Neither Learning group, nor the interaction between Learning group and Coda position was significant, as can be seen in the summary in Table 4.

	b	*Z*	*p*
Intercept	1.43847	10.275	< 2e-16*
Coda position (medial)	-0.87895	-6.072	1.27e-09*
Learning group (WtS)	-0.07397	-0.453	0.651
Coda position × Learning group	-0.08530	-0.587	0.557

*Table 4: Fixed effects in a logistic regression model of response correctness. Significant values ($p < 0.05$) are indicated by *.*

The significant effect of Coda position suggests imperfect generalization to the utterance/word-medial position. Participants in both groups learned the distribution pattern more successfully in the position for which they had direct evidence of the pattern in their learning sets. While the higher percentage of correct responses in the UtW group than in the WtS group may suggest that generalization to the syllable is more difficult, less natural, or a slower process than generalization to the word level, this effect did not reach significance.

5. Discussion

This experiment shows that generalization to the syllable can occur in the same type of experiment as generalization to the word domain.

Results from the UtW group confirm that the learning of an utterance-final distribution pattern and generalization to utterance-medial, word-final position found in earlier experiments can be replicated with slightly different experiment design. Despite differences in materials and procedure between this experiment, M&P, and my earlier domain generalization experiment (discussed briefly in Section 3), the results still demonstrated both learning and generalization.

Results from the WtS group showed that the syllable is also a domain to which such distributional patterns may be generalized. Participants accepted word-internal coda [l] more frequently than word-internal coda [ɹ], despite a lack of direct evidence about word-internal liquid coda goodness during the learning phase of the experiment (due to the poverty of the stimulus design).

Because this group was presented with single-word utterances, the pattern to which they were exposed during the learning phase was both word-final and utterance-final. Future research may want to investigate whether word-to-syllable generalization relies upon learning at the utterance edge, and whether generalization must be stepwise (utterance-to-word-to-syllable) or can go directly form an utterance-final pattern to a syllable-level generalization. Importantly, however, the domain to which the restriction on final liquids was generalized was the syllable. Participants believed they were listening to single-word utterances, yet their learned preference for coda [l] over coda [ɹ] carried over to word-internal codas in the testing phase.

The lack of a main effect of Learning group, or any interaction involving Learning group in either responses (Table 1) or correctness (Table 4) shows that the two groups behaved quite similarly in terms of both learning and generalization. It would not be surprising if generalization to the word and generalization to the syllable were slightly different processes. The syllable is, after all, a smaller domain than the word (contained within the word). The results of this experiment, however, do not show a significant difference between the two groups.

It may be possible to analyze a syllable-final pattern, like Tz'utujil devoicing in (1), without direct reference to syllable structure, by characterizing it as word-final and pre-consonantal, or non-prevocalic. It is then, theoretically, possible that participants in this experiment generalized from word/utterance-final liquid neutralization to non-prevocalic liquid neutralization. This would still be a type of generalization to a smaller-than-word environment, though it seems less plausible than generalization by analogy from one prosodic domain to another—from word/utterance-final syllables to all syllables.

This experiment may also raise questions of ecological validity and learning mechanisms. Adults do not explicitly inform child learners that they will be hearing a series of individual trisyllabic words or three-word sentences. This element of the AGL experiment is quite artificial and may raise concern about whether the two groups—which performed so similarly—really believed they were doing different tasks. Exit survey responses, however, suggest that participants were, indeed, thinking of items as multi-word sentences in the UtW group and individual words in the WtS group. Future work will seek to address this concern further.

This study presents empirical evidence of phonological generalization to the syllable domain. This shows that the syllable is a relevant structure at some levels of memory and grammar, contra to claims that the syllable is not a linguistic unit.

References

Barnes, Jonathan. 2006. *Strength and weakness at the interface: Positional neutralization in phonetics and phonology*. Berlin, New York: Mouton de Gruyter.

Bennett, Ryan. 2016. Mayan phonology. *Language and Linguistics Compass* 10.469-514.

Bermudez-Otero, Ricardo, and April McMahon. 2006. English phonology and morphology. In Aarts, B. & A. McMahon (eds.), *The handbook of English linguistics* (pp. 82-410). Oxford: Blackwell.

Bermúdez-Otero, Ricardo, and Graeme Trousdale. 2012. Cycles and continua: On unidirectionality and gradualness in language change. In T. Nevalainen, & E. Closs Traugott (Eds.), *The Oxford handbook on the history of English* (pp. 691-720). [55] (Oxford Handbooks Online). Oxford University Press.

Blevins, Juliette. 2006. A Theoretical Synopsis of Evolutionary Phonology, *Theoretical Linguistics* 32:117-65.

Boersma, Paul. 1997. How we learn variation, optionality, and probability. *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam* 21.43–58.

Dayley, Jon. 1985. *Tzutujil grammar*, vol. 107 University of California Publications in Linguistics. Berkeley, CA: University of California Press.

Faytak, Matthew, Suyuan Liu, and Megha Sundara. 2020. Nasal coda neutralization in Shanghai Mandarin: Articulatory and perceptual evidence. *Laboratory Phonology* 11.1-29.

Finley, Sara, and William Badecker. 2009. Artificial language learning and feature-based generalization. *Journal of Memory and Language* 61.423–437.

Hock, Hans Henrich. 1991. *Principles of historical linguistics*. 2nd edn. Berlin & New York: Mouton de Gruyter.

Hyman, Larry. 1978. Word demarcation. In J. H. Greenberg (ed.) *Universals of Human Language*. Vol. 2. Phonology. Stanford University Press, pp. 447-470.

Ito, Junko, and Armin Mester. 2009. The extended prosodic word. In Kabak, Baris, and Jaent Grijzenhout, eds. *Phonological Domains: Universals and Derivations.* The Hague, The Netherlands: Mouton de Gruyter. 135-194.

Johnson, Keith. 1996. Speech perception without speaker normalization. in Johnson, K. and Mullennix (eds.) *Talker Variability in Speech Processing*. San Diego. Academic Press.

Lombardi, Linda. 1995. Laryngeal Neutralization and Syllable Wellformedness. Natural Language & Linguistic Theory, 13(1), 39–74.

Myers, Scott, and Jaye Padgett. 2014. Domain generalisation in artificial language learning. *Phonology* 31.399-433.

Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. In P. F. MacNeilage (ed.), *The Production of Speech*. New York: Springer. 189-216.

Ohala, John, and Haruko Kawasaki-Fukumori. 1997. Alternatives to the sonority hierarchy for explaining the shape of morphemes. In *Language and its ecology: Essays in memory of Einar Haugen*, ed. by Stig Eliasson and Ernst H. Jahr, 343-365. Berlin: Mouton de Gruyter.

Padgett, Jaye. 2015. Word-edge effects as overphonologization of phrase-edge effects. In U. Steindl, T. Borer, H. Fang, A. García Pardo, P. Guekguezian, B. Hsu, C. O'Hara, & I. Chuoying Ouyang (Eds.), *Proceedings of the 32nd West Coast Conference on Formal Linguistics* (pp. 41-53). Somerville, MA: Cascadilla Proceedings Project.

Pierrehumbert, Janet. 2000. Exemplar dynamics: Word frequency, lenition and contrast. In Bybee, J. & P. Hopper (eds.), *Frequency effects and the emergence of linguistic structure*. Amsterdam: John Benjamins.

Prince, Alan, and Paul Smolensky. 1993. *Optimality Theory: constraint interaction in generative grammar*. Ms, Rutgers University & University of Colorado, Boulder. Published 2004, Malden, Mass. & Oxford: Blackwell.

Steriade, Donca. 1999. Alternatives to the syllabic interpretation of consonantal phonotactics. In *Proceedings of the 1998 Linguistics and Phonetics Conference,* ed. by Osamu Fujimura, Brian Joseph, and Bohumil Palek, 205-242. Prague: The Karolinum Press.

Vaux, Bert. 2003. Syllabification in Armenian, Universal Grammar, and the Lexicon. *Linguistic Inquiry*, 34(1).91-125.

Wax Cavallaro, Maya. 2021. [SG] and final consonant allophony in Tz'utujil. In R. Bennett, R. Bibbs, M.L. Brinkerhoff, M.J. Kaplan, S. Rich, N. Van Handel & M. Wax Cavallaro (eds.), *Proceedings of the 2020 Annual Meeting on Phonology*. Washington, DC: Linguistic Society of America.

Wax Cavallaro, Maya. 2022. *Cross-linguistic patterns of sonorant devoicing as evidence for [spread glottis] and [voice] in sonorants*. Santa Cruz: University of California, Santa Cruz qualifying paper.

Wax Cavallaro, Maya. 2023. *The syllable in domain generalization: Evidence from artificial language learning*. Santa Cruz: University of California, Santa Cruz thesis.

Westbury, John R., and Patricia A. Keating. 1986. On the naturalness of stop consonant voicing. *JL* 22.145-166.

Wetzels, W. Leo, and Joan Mascaró. 2001. The Typology of Voicing and Devoicing. *Language* 77.207–244.

White, James. 2013. Evidence for a learning bias against saltatory phonological alternations. *Cognition* 130.96–115.

Wilson, Colin. 2003. Experimental investigation of phonological naturalness. *WCCFL* 22.533–546.

Wilson, Colin. 2006. Learning phonology with substantive bias: an experimental and computational study of velar palatalization. *Cognitive Science* 30.945–982.

Zec, Draga, and Elizabeth Zsiga. 2022. Tone and stress as agents of cross-dialectal variation: The case of Serbian. In Kubozono, H., Ito, J & A. Mester (eds), *Prosody and Prosodic Interfaces*.

Zehr, Jérémy., and Florian Schwarz. 2018. *PennController for Internet Based Experiments (IBEX).*