AMP 2024
November 2, 2024
Rutgers University

Scan to obtain PDF copy

An experimental study of Catalan consonant alternations

Kevin Liang, Victoria Mateu, and Bruce Hayes
University of California, Los Angeles

(1) Goals for this talk

- Report on two experiments that assess the productivity of four well-studied phonological alternations of Catalan: stem-final /n/- and /r/-deletion, /nt/ simplification, and [ʒ] ~ [tʃ] alternation.
- Discuss the implications of our experimental results for phonological theory questions relating to the learnability of exceptionality, opacity, and saltation.

BACKGROUND

(2) Typical morphological context of the four alternations we examined

- They occur at the right edge of stems.
- We focus on alternations between feminine forms, with suffix [-ə], and unsuffixed masculines.
- Normally, the feminine form preserves the UR intact, with phonological changes occurring in the masculine.

(3) /n/-deletion and /r/-deletion

- /n/ and /r/ are deleted in final position, following very similar patterns.
  - Deletion of /n/ /san/ [ˈsa**n**-ə] ~ [sa] ‘healthy fem./masc.’
  - Deletion of /r/ /dur/ [ˈdu**r**-ə] ~ [du] ‘hard fem./masc.’
- Both processes are *lexically specific*: applicability to individual items must be memorized, in some way.
- But the exceptionality is *patterned* (Zuraw, 2000): various factors influence deletion rates across the lexicon.
- The patterning is essentially the same for both /n/-deletion and /r/-deletion.
- Here are the patterns we study:

- **Penultimately-stressed stems** virtually never undergo deletion (e.g., [əwˈtɔk.tu.**n**ə] ~ [əwˈtɔk.tu**n**] ‘autochthonous-fem./masc’; [ˈprɔs.pə.**r**ə] ~ [ˈprɔs.pə**r**] ‘prosperous’).
- **Monosyllabic stems**: deletion is more frequent, about half the time ([ˈsa.**n**ə] ~ [ˈsa] ‘healthy’, but [ˈnɛ.**n**ə] ~ [ˈnɛ**n**] ‘child’; [ˈkla.**r**ə] ~ [ˈkla] ‘clear’, but [ˈpu.**r**ə] ~ [ˈpu**r**] ‘pure’).
- **Frequent suffixes**: deletion is exceptionless with *-i(**n**a)* ‘related to’ and *-do**r**(a)* ‘agentive’ ([ər.ʒənˈti.**n**ə] ~ [ər.ʒənˈti] ‘Argentine’, [əd.mi.nis.trəˈdo.**r**ə] ~ [əd.mi.nis.trəˈdo] ‘administrator’).
- In all **other cases**, it is *near*-exceptionless ([kə.təˈla.**n**ə] ~ [kə.təˈla] ‘Catalan’, [sə.ˈgu.**r**ə] ~ [sə.ˈgu] ‘safe’).

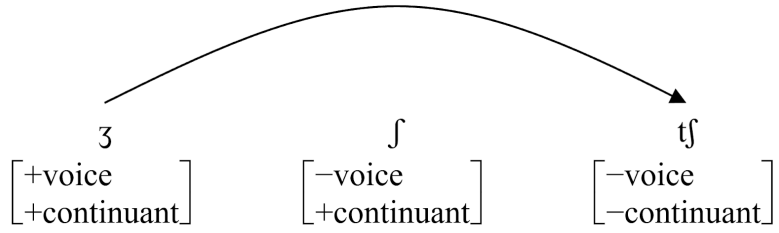
(4) /nt/ cluster simplification

- /t/ is deleted finally after /n/, as in [ˈsan.**t**ə] ~ [ˈsan] ‘saint’.
- Note that /nt/ cluster simplification and /n/-deletion show *counterfeeding opacity*:
  - Word-final [n] resulting from cluster simplification is never deleted (no cases like [ˈsa**nt**-ə] ~ *[ˈsa]).
- Here is an illustration using (for brevity) classical rule-based phonology:

‘holy-m.’	‘holy-f.’	‘healthy-m.’	‘healthy-f.’	
/sant/	/sant-ə/	/san/	/san-ə/	URs
—	—	sa	—	*Final /n/-deletion*: n → ∅ / ___]$_{word}$
san	—	—	—	*Cluster Simplification*: t → ∅ / n ___]$_{word}$
[san]	[santə]	[sa]	[sanə]	SRs

(5) [ʒ] ~ [tʃ] alternation

- This is a **saltatory** alternation, in the sense of Hayes & White (2015).
  - All voiced obstruents undergo devoicing in final position (e.g. [ˈgriz-ə] ~ [ˈgris] ‘gray-fem./masc.’).
  - But [ʒ] devoices not to the expected [ʃ] but [tʃ], as in [ˈbɔ.**ʒ**ə] ~ [ˈbɔ**t͡ʃ**] ‘crazy’.
  - Thus [ʒ] “saltates,” jumping over intermediate [ʃ] in arriving at surface [tʃ]:



- We can't simply turn final [ʃ] into [tʃ]: final [ʃ] is well-formed, e.g. [baʃ] 'short'.

(6) The productivity of saltation

- Hayes & White (2015) consider saltation to be marked, and document cases of diachronic breakdown.
- White's experimental and modeling work (artificial grammar learning) suggests a learning bias against saltation (White, 2014 in adult English-speakers; White & Sundara, 2014 in 12-month-old infants).

(7) Three theoretical issues that the work will address

a) What productive generalizations do learners make from exceptionful data?

- Current work suggest a two-part answer:
  - As a rough approximation (Zuraw, 2000 et seq.): when using their grammar productively, language learners *frequency-match the lexicon.*
  - But they *deviate* from frequency-matching due to UG biases (Becker et al., 2011 in Turkish; Becker et al., 2012 in English; Ernestus & Baayen, 2003 in Dutch; Hayes et al. 2009 in Hungarian).
- Do Catalan speakers' responses for /n/-deletion and /r/-deletion differ from the lexical pattern, and if so, why?

b) Can an opaque alternation be productive?

- As shown in (4), /nt/ cluster simplification interacts opaquely with /n/-deletion.
- Sanders (2003) argues that Polish counterbleeding opacity is not productive and is instead dealt with by memorization.
- What of the opaque pattern in Catalan?

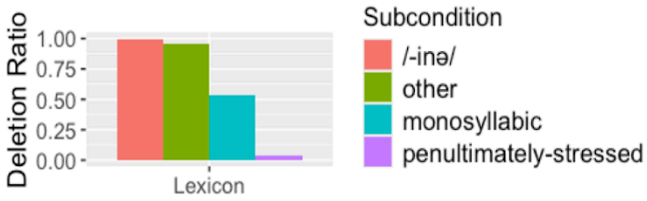
c) *Is the saltatory* [ʒ] ~ [t͡ʃ] *alternation of Catalan productive*?

- Do speakers commit "saltation repair?"
  - Example: [səˈlɔʃ] 'wug-masc.', instead of [səˈlɔt͡ʃ], when given feminine [səˈlɔʒ-ə].

(8) Evaluating the lexical generalizations quantitatively: our database

- The above generalizations are carefully covered in the analytical literature, notably Mascaró (1976) and Wheeler (2005).
- We reconfirmed the patterns and assessed them quantitatively by constructing and counting a corpus of 5,761 nominal and adjectival paradigms, compiled from Wiktionary.

Graph: rates of /n/-deletion for the four environments given above in the lexicon



- These data will appear in the graphs below as we compare the patterns seen in the wug test with the patterns of the lexicon.

OUR WUG-TEST EXPERIMENT

(9) Strategy

- A classical wug test (Berko, 1958).
- We gave the participants feminine forms, and designed the task to require them to construct the corresponding masculine, thus testing the productivity of the target phonological process.
  - **Experiment 1** (production task): given a feminine form, participants recorded themselves saying aloud the appropriate masculine form.
  - **Experiment 2** (rating task): participants rated the acceptability of two or three potential masculine forms on a scale from 1 to 7. Choices were as in table (12) below.
  - Example: asking for the masculine of [ˈfrun-ə] tests the productivity of /n/-deletion in monosyllables – will they respond with (Expt. 1) or prefer (Expt. 2) [ˈfru] or [ˈfrun]?

(10) Participants

- Adult, native speakers of Central Catalan who spoke Catalan at home and attended elementary school in Catalan, **37** per experiment after exclusions.
- Participated online.
- For exclusion criteria and other details, see written paper.

(11) Materials

- We employed 100 different feminine wug forms, such as [səˈða.n-ə], for this study; any one participant saw a balanced selection of 20.
- In designing the wug forms we sought to achieve:
  - **Phonotactic acceptability** (wugs sound natural to a native speaker)
  - **Novelty** (wugs and their inflected forms were not real words of Catalan)
  - **Variegation** (they contained a wide range of distinct consonants and vowels)

(12) Sample wug forms

- There were 4 conditions and 10 subconditions as exemplified in the table below.
- The table entries correspond directly to the phonological phenomena and environments described earlier.
- Comment on possible outcomes for the masculine form:
  - First outcome: process applies
  - Second outcome: process does not apply

Condition	*Subcondition*	*Feminine form (presented to participants)*	*Anticipated masculine responses*
/n/-deletion	frequent affix /-inə/	[bəlunˈtrin-ə]	[bəlunˈtri], [bəlunˈtrin]
	monosyllabic	[ˈfrun-ə]	[ˈfru], [ˈfrun]
	penultimately-stressed	[ˈdɔstun-ə]	[ˈdɔstu], [ˈdɔstun]
	other	[gəˈmɛn-ə]	[gəˈmɛ], [gəˈmɛn]
/r/-deletion	frequent affix /-dorə/	[gruəˈdor-ə]	[gruəˈdo], [gruəˈdor]
	monosyllabic	[ˈlɛr-ə]	[ˈlɛ], [ˈlɛr]
	penultimately-stressed	[ˈsɔlir-ə]	[ˈsɔli], [ˈsɔlir]
	other	[kəˈnar-ə]	[kəˈna], [kəˈnar]
/nt/ final cluster reduction (opacity)	—	[mirˈbunt-ə]	[mirˈbun], [mirbunt], [mirˈbu] (feeding order)
/ʒ/ final obstruent devoicing (saltation)	—	[səˈlɔʒ-ə]	[səˈlɔt͡ʃ], [səˈlɔʃ] (final devoicing only)

(13) Frame paragraphs

- The feminine wug items were first presented once in isolation, and then embedded in frame paragraphs read by a female native speaker.
- Sample paragraph:

  WUG-fem.

  Una obra __WUG-fem__ era una peça d'art on s'havien aplicat tècniques mixtes amb ornaments de metalls i pedres precioses. Al segle XV, un artista català va crear la primera escultura ____, feta de marbre, pedres precioses, i or. El primer quadre ___ no es va crear a Espanya fins al segle XVII.

  'A ____ work was a work of art where they applied mixed media with precious metals and stone ornaments. In the 15th century, a Catalan artist created the first ___ sculpture, made of marble, precious stones and gold. The first ___ painting was not created in Spain until the 17th century.'

- The frame paragraphs were constructed with the goal of encouraging participants to interpret the stimuli as authentic Catalan words.
- The paragraphs were recorded such that there was a pause where a response was requested.
- The grammatical context was always one which would force the use of a masculine form of the wug word to fill the pause.
- The frames were presented both in spoken form and as text. However, the wugs never appeared in written form.

KEY RESULTS AND THEORETICAL INFERENCES

(14) How we report the findings

- Expt. 1 and Expt. 2 yielded very similar results, so we report them together.
- We are not reporting statistical testing in this talk; generally, differences we report here test as significant; please ask us for the written paper to see full details.

(15) Cases not included in the totals here

- Didn't repeat the wug form correctly (see (13)) (6%).
- Isolated, hard-to-interpret forms, like [ʎuˈdaʒə] → [ʎuˈda] (9%).
- "Avoidant" responses (18%): use the rare masculine endings [-u] and [-ə] – these let you to avoid having to make a commitment about phonology.
  - See Do (2018) for the same behaviour in Korean children.

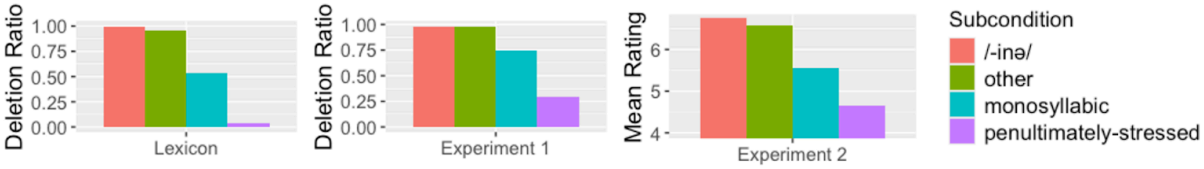
(16) General findings

- All of the processes we investigated were productive at least to some degree.
- In detail, the findings shed light on various theoretical questions.

(17) Frequency-matching in /n/-deletion

- We obtained clear evidence of frequency-matching for both experiments:

Graph: /n/-deletion in the lexicon, Experiment 1 (production), and Experiment 2 (ratings)

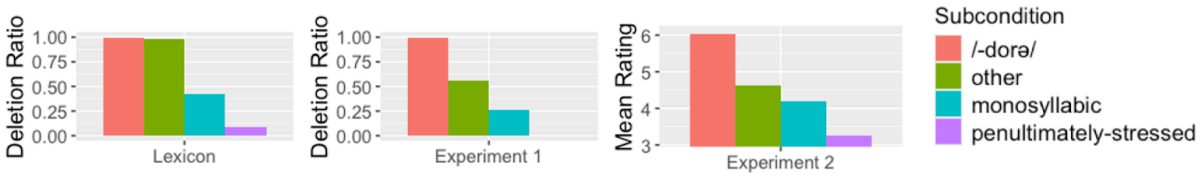


  - Experiment 1 (production): Across four environments, the environments where /n/-deletion applies most often in the lexicon match the environments where speakers most often applied /n/-deletion: *frequent affix* > *other* > *monosyllabic stems* > *penultimately-stressed stems.*
  - Experiment 2 (ratings): Same pattern (most to least acceptable).
  - *Not* "dialect mix": although there were participants who consistently deleted and others who consistently produced /n/ or /r/, most participants provided both types of answers (ditto for all other phenomena).

(18) Participants also frequency-matched for /r/-deletion

- The four contexts for /r/-deletion have similar *relative* frequencies in the lexicon and experiments.

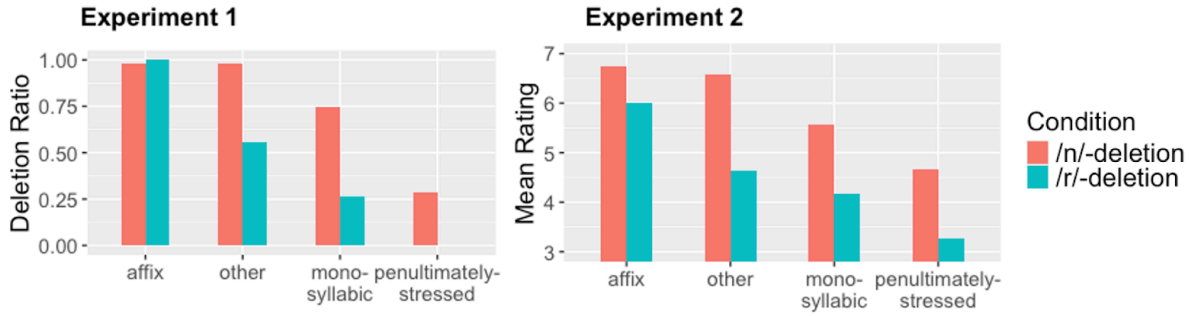
Graph: /r/-deletion in the lexicon, Experiment 1 (production), and Experiment 2 (ratings)



(19) A frequency-matching puzzle: why does /n/ delete far more often than /r/?

- [n]-deletion closely matched the lexical frequencies.
- But [r]-deletion matched only in *relative* terms:
  - Speakers consistently disfavored [r]-deletion, relative to [n]-deletion.

Graph: /n/- and /r/-deletion in Experiment 1 (production) and Experiment 2 (ratings)



(20) Why the /n/ - /r/ difference? Hypothesis I: dialect variation

- Speakers of Central Catalan encounter speakers of another major dialect, Valencian, which lacks /r/-deletion.
- /n/-deletion is pan-dialectal (Wheeler, 2005).

(21) Why the /n/ - /r/ difference? Hypothesis II: orthographic influence

- Previous work argues that phonological intuitions are often influenced by orthography (see Kawahara, 2018; Daland, Oh & Kim, 2015).
- In Catalan orthography, /n/-deletion is spelt out:
  - [ˈsan-ə] ~ [ˈsa] is spelt *sana ~ sa*
- /r/-deletion is not spelt out:
  - [ˈklar-ə] ~ [ˈkla] is spelt *clara ~ clar*
- Rough idea: Our participants may have been constructing appropriate orthographic representations for what they heard, preferring to pronounce these representations faithfully.
- We are exploring models that could express and incorporate this influence on participant responses.

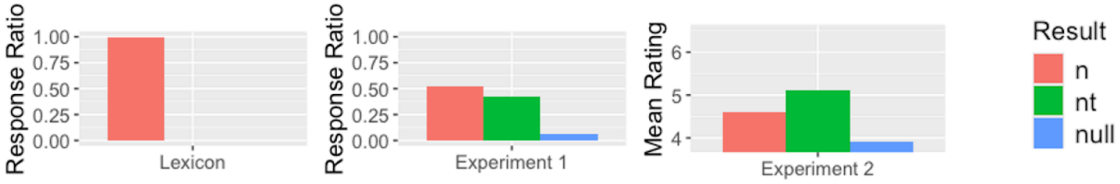
(22) Contexts where /n/ deletion is *over*applied relative to the lexicon

- This occurred in both monosyllables and penultimately-stressed stems ((17) above).
- Overapplication to monosyllables is particularly interesting, in light of Becker et al.'s (2012) experimental evidence for a UG bias *against* alternation in monosyllables.
- We conjecture that here, two countering effects pair up to override this bias:
  - **Simplicity bias** (Moreton & Pater 2012): plain, uncontextual deletion is favored; monosyllables and penultimately-stressed forms often just follow the simplest available generalization.
  - **Attestation bias** (Albright & Hayes, 2003; Siah, 2024): you need enough data to take the generalization seriously; and there aren't enough cases of these two types (see Appendix for lexical counts).

(23) /nt/ Cluster Simplification

- Despite this process being exceptionless in the lexicon, 42% of the responses in Experiment 1 had final [nt].
- Such forms were also rated higher than expected, better than forms undergoing cluster simplification.

Graph: /nt/ cluster simplification in the lexicon, Expt. 1 (production), and Expt. 2 (ratings)



- We conjecture three possibilities:
  - Exposure to other languages or other dialects of Catalan that allow final [nt] (e.g., Wheeler, 2005:221) weakens the native-language phonotactic constraint banning final [nt].
  - Orthographic influence, as above: /nt/ cluster simplification is *not* spelt out, e.g. [san] 'saint-masc.' is spelt *sant*.
  - Opacity repair: see immediately below.

(24) Results for opacity (/nt/ cluster simplification and /n/-deletion)

- We examined the tokens provided for the conditions /nt/ and /n/-other (the prosodically matched subcondition) within each speaker, and found the following proportion of paired answers:[1]

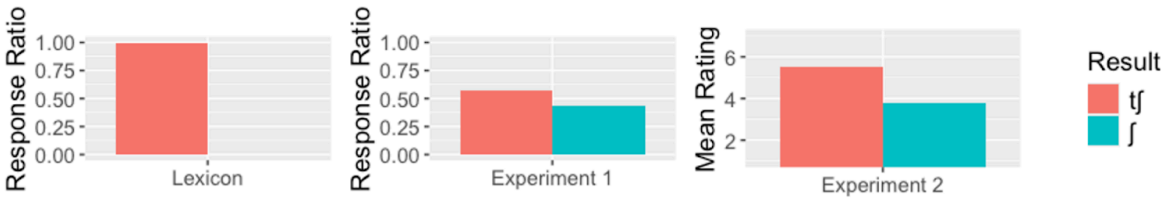
/n/-deletion result	*/nt/ cluster simplification result*	*Fraction of total*
[gəˈmɛn-ə] → [gəˈmɛ]	[mirˈbuntə] → [mirˈbun]	67%
[gəˈmɛn-ə] → [gəˈmɛ]	[mirˈbuntə] → [mirˈbunt]	33%
[gəˈmɛn-ə] → [gəˈmɛn]	[mirˈbuntə] → [mirˈbun]	0%
[gəˈmɛn-ə] → [gəˈmɛn]	[mirˈbuntə] → [mirˈbunt]	0%

[1] Participants received two wug words for each subcondition; where their responses treated both pairs identically, we counted the data as two response patterns.

- This examination suggests that *counterfeeding opacity can be quite productive*: it is found for two thirds of the response patterns.
- Of the remaining cases, all were of the type [gəˈmɛn-ə] → [gəˈmɛ], [mirˈbuntə] → [mirˈbunt].
  - Opacity-related? – if you don't apply /nt/ deletion, the resulting output keeps /n/-deletion transparent.

(25) The saltatory [ʒ] ~ [tʃ] alternation

Graph: [ʒ] ~ [tʃ] alternation in the lexicon, Expt. 1 (production), and Expt. 2 (ratings)



- Many speakers produced forms that repaired saltation (e.g., [ʎuˈdaʒə] → [ʎuˈdaʃ]) and rated such forms highly.
- These saltation repairs have [ʃ], not [ʒ], because Final Devoicing remains a powerful phonotactic principle.
- Lexical basis: Forms with [ʒ] ~ [ʃ] are *not attested* in the lexicon, nor in any other dialect of Catalan.
- White (2014) and Hayes & White (2015) argue that saltation is a form of "unnatural phonology," liable to repair – we may be seeing such a case here.
- However, attestation bias (few saltatory forms in lexicon) may also explain these results.

SUMMARY OF FINDINGS

(26) Tentative answers to our research questions (7a-c)

- Catalan speakers generally:
  - **frequency-match the lexicon**, with deviations resulting perhaps from orthography or dialect differences.
  - **can manage opacity**: Many participants gave the /nt/ → [n], /n/ → ∅ pattern.
  - **tend to repair saltation.**

Moltes gràcies!

(27) Thanks to …

- Marta Camps, Roger Castells-Graells, Anna Gavarró, Mireia Marimón, Joan Mascaró, Jaume Mateu, Benet Oriol Sabat, Francesc Reda Coll, Gemma Repiso-Puigdelliura, Aina Soley Mateu, Mireia Toda Cosi
- CU Política Linguística (Generalitat de Catalunya), Institut d'Estudis Catalans, Societat Catalana de Llengua i Literatura
- Our experimental participants
- UCLA Dean of Humanities for research funding
- Members of the UCLA Phonology Seminar

Appendix: Lexical attestation of the various phenomena in our corpus analysis (nouns and adjectives)

Condition	*Subcondition*	*Undergoers*	*Non-undergoers*	*Total*
/n/-deletion	frequent affix /-inə/	105	0	105
	monosyllabic	8	7	15
	penultimately-stressed	1	26	27
	other	390	20	410
/r/-deletion	frequent affix /-dorə/	205	0	205
	monosyllabic	3	4	7
	penultimately-stressed	2	22	24
	other	250	6	256
/nt/ final cluster reduction (opacity)	—	40	0	40
/ʒ/ final obstruent devoicing (saltation)	—	7	0	7[2]

References

Albright, Adam & Bruce Hayes. (2003). Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition, 90*(2), 119-161.

Becker, Michael, Nihan Ketrez, & Andrew Nevins. (2011). The surfeit of the stimulus: Analytic biases filter lexical statistics in Turkish laryngeal alternations. *Language*, 84-125.

Becker, Michael, Andrew Nevins, & Jonathan Levine. (2012). Asymmetries in generalizing alternations to and from initial syllables. *Language, 88*(2), 231–268.

[2] Just 7 cases in masculine-feminine paradigms, but more in other paradigms.

Berko, Jean (1958). The child's learning of English morphology. *Word, 14*(2-3), 150-177.

Daland, Robert, Mira Oh & Syejeong Kim. (2015). When in doubt, read the instructions: Orthographic effects in loanword adaptation. *Lingua, 159*, 70-92.

Do, Young-Ah. (2018). Paradigm uniformity bias in the learning of Korean verbal inflections. *Phonology, 35*(4), 547–575.

Ernestus, Miriam, & Harald R. Baayen. (2003). Predicting the Unpredictable: Interpreting Neutralized Segments in Dutch. *Language, 79*(1), 5–38.

Hayes, Bruce, Kie Zuraw, Peter Siptar, & Zsuzsa Londe (2009). Natural and unnatural constraints in Hungarian vowel harmony. *Language* 85: 822-863.

Hayes, Bruce & James White. (2015). Saltation and the P-map. *Phonology, 32*(2), 267–302.

Kawahara, Shigeto. (2018). Phonology and orthography: The orthographic characterization of rendaku and Lyman's Law. *Glossa, 3*(1).

Mascaró, Joan. (1976). *Catalan phonology and the phonological cycle*. [Doctoral dissertation, Massachusetts Institute of Technology].

Moreton, Elliott & Joe Pater. (2012) Structure and substance in artificial-phonology learning, part I: Structure. *Language and linguistics compass* 6:686-701.

Sanders, Nathan (2003). *Opacity and sound change in the Polish lexicon*. [Doctoral dissertation, University of California, Santa Cruz].

Siah, Jian-Leat. (2024). *Prosodic end-weight effects in Malay echo redpublication: the role of naturalness and attestedness*. [Master's thesis, University of California, Los Angeles]

Wheeler, Max. (2005). *The phonology of Catalan*. Oxford University Press.

White, James. (2014). Evidence for a learning bias against saltatory phonological alternations. *Cognition, 130*(1), 96–115.

White, James & Megha Sundara. (2014). Biased generalization of newly learned phonological alternations by 12-month-old infants. *Cognition, 133*(1), 85–90.

Zuraw, Kie. (2000). *Patterned exceptions in phonology*. [Doctoral dissertation, University of California, Los Angeles].