A richer model is not always more accurate: the case of French and Portuguese plurals*

Michael Becker michael@linguist.umass.edu
Lauren Eby Clemens lreby@fas.harvard.edu
Andrew Nevins a.nevins@ucl.ac.uk

Version of September 7, 2011. Your feedback welcome!

Abstract

We present two cases of morphophonological alternations in the plural of nouns, one from French and one from Brazilian Portuguese. In both of them, monosyllabic items are protected from alternations more than polysyllabic items are. Two large-scale nonce word tasks confirm the productivity of the trend in both languages. We analyze the results in terms of a discrete grammatical distinction between monosyllables and polysyllables, comparing this approach to continuous, more detail-rich models that rely on lexical statistics or phonetic duration. We show that the discretized approach is the most accurate predictor of participants' responses, as informationally richer models introduce irrelevant variance that diminishes their predictive force. We conclude that a relatively simple model of morphophonological alternations that describes the data in discrete, grammatical terms is closer to the one in fact employed by native speakers. A richer model is thus not always a more accurate representation of cognitive reality.

Keywords: morphophonological alternations, irregular plurals, initial syllable faithfulness, neighborhood density, phonetically-based phonology

1 Introduction

Word length, as measured by syllables, is one of the phonological properties that speakers find most fundamental and most readily accessible. In this paper, we examine morphophonological alternations found in the plural of French and Brazilian Portuguese nouns, and propose that both cases are governed by a discrete binary distinction between monosyllables and polysyllables. We contrast this discretized approach with continuous predictors that are grounded either in the phonetic realization of these alternations or in terms of interactions with word likelihood, as measured by neighborhood density or bigram frequency.

Much research has shown that speakers have detailed knowledge about the distribution of irregular morphophonological trends in their lexicon, and that this knowledge is applied to nonce words (Zuraw 2000; Ernestus & Baayen 2003; Albright & Hayes 2003; among many others). But what is the nature of this knowledge? How fine-grained are the mechanisms regulating grammatical generalization to novel forms?

One answer is that speakers use discrete formal primitives such as syllable count or feature-markedness constraints, as can be found in frameworks such Optimality Theory (Prince &

*Thanks to Diogo Almeida, Jill Beckman, Flávia Medina Cunha, Johanne Gaya, Peter Graff, Maximiliano Guimarães, John J. McCarhy, Clémence Jouët-Pastré, Joe Pater, Magda Oiry, Catherine O. Ringen, Cilene Rodrigues, Sverre Stausland Johnsen, Jason Shaw, Guillaume Pierre Yves Thomas, and Matt Wolf. Special thanks go to Filomena Sandalo, for her help with many aspects of this work. This paper is error-free.

Smolensky 1993/2004), and that they build a ranking to distill the lexical statistics in terms of a grammatical encoding that can be applied to nonce words. This characterizes the constraint cloning approach (Pater 2006, 2009; Coetzee 2008; Becker 2009; Becker, Ketrez & Nevins 2011a) and the UseListed approach (Zuraw 2000; Hayes & Londe 2006; Becker, Nevins & Levine 2011b). These authors differ in their commitment to the universality of the constraint set and its naturalness, but it is often implicitly assumed that universality and naturalness are guiding principles, if not ironclad rules. In this paper, we focus only on the ***discrete*** nature of these constraints.

In our case, we will show that in both French and Brazilian Portuguese, monosyllables are asymmetrically protected from plural alternations, whereas polysyllables are impacted more strongly. We attribute this size effect to ***initial syllable faithfulness*** constraints (Trubetzkoy 1939; Steriade 1994; Beckman 1997, 1998; Casali 1998; Barnes 2006; Jesney 2009; Becker 2009; Becker et al. 2011a).

In contrast to research that attempts to understand such phenomena in terms of positionally-relativized grammatical principles, another strand of investigation offers explanation in terms of the lexicon itself and properties derived from it, such as neighborhood density. Ussishkin & Wedel (2009) propose that words resist morphophonological alternation because they are dense lexical neighborhoods. This approach relies on continuous variables that are informationally rich, though also sensitive to irrelevant variance that may reduce their ability to hone in on the underlying pattern. Such variance may be the result of accidental distributional trends, such as those of particular onset clusters, which turn out to be irrelevant with respect to speakers' representations of morphophonological trends affecting elements in the rime.

Continuous, informationally-rich variables are also at the heart of approaches that look to derive the protection of shorter words from the phonetic basis of particular languages (Barnes 2006; Giavazzi 2010); usually the guiding principle is that phonetically salient elements (which have measurable properties such as greater duration, louder burst, etc.) will be protected from alternations. Again, this approach is vulnerable to potentially irrelevant variance.

The phonetic reality of the alternating elements is also relevant to usage-based approaches, which seek to ground morphophonological alternations in lenition processes (Hooper [Bybee] 1976; Pierrehumbert 2001). We will show that lenition is not involved in the Brazilian Portuguese data at hand, nor can lenition derive the size effect in French; these leave usage-based accounts in need of the same appeal to protection of monosyllables to drive the effect.

We will compare our discrete approach with the lexical and the phonetic continuous approaches, and show that the former is the best predictor of speakers' preferences, as determined by nested model comparison. Furthermore, continuous predictors make no significant improvement to the account, thereby rendering them otiose for the purpose of predicting morphophonological alternations. While neighborhood density may play a role in irregular alternations in French, we discuss its interaction with vocabulary size in §6, maintaining the assertion that syllable count is a fundamental perceptual property of words that is readily incorporable into learners' grammars.

The paper is structured as follows: We start with a study and analysis of the French lexicon in §2, and show how the analysis predicts participants' treatment of novel words in §3. Portuguese receives the same treatment in §4 and §5. In §6, we compare our discrete approach to the continuous approaches. We conclude in §7.

2 The French lexicon

We start by surveying French plural morphology, noting that in the real words of French, alternations impact polysyllables more than monosyllables. This size-based asymmetry is a novel observation, to our knowledge. We show that this trend is statistically significant, and we offer an analysis that captures the trend grammatically by appealing to initial syllable faithfulness constraints.

2.1 The French irregular plural

The plural normally has no distinct morphological marking in French nouns, e.g. ⟨nom⟩ [nõ] 'noun' is identical in the singular and plural, with number only marked on determiners and certain verbs. Some nouns that end in [l] or [j], however, have a phonologically distinct plural that impacts the stem, as shown in (1).

In addition to the nominal alternations seen in (1), the same [al] → [o] alternation is also fairly common for [al]-final adjectives, e.g. [lwajal ∼ lwajo] 'loyal'.

(1) Unfaithful, alternating French plurals[1]

	alternation	singular	plural	
a.	[al] → [o]	mal	mo	'evil'
		bokal	boko	'jar'
		ʒuʁnal	ʒuʁno	'newspaper'
b.	[aj] → [o]	baj	bo	'lease'
		tʁavaj	tʁavo	'work'
		supiʁaj	supiʁo	'basement window'

These alternations contrasts with near-minimal pairs of nouns that keep the same form in the singular and in the plural, as in (2). In addition, some 'vacillating' nouns (e.g. [val] 'valley', [beʁkaj] 'home') can take either kind of plural.

[1] All the alternating nouns are spelled with an ⟨l⟩ in the singular and ⟨aux⟩ in the plural, including those that end in [j], e.g. ⟨journal ∼ journaux⟩, ⟨soupirail ∼ soupiraux⟩.

(2) Faithful, non-alternating French plurals

	singular	plural	
a.	bal	bal	'ball'
	ʃakal	ʃakal	'jackal'
	kaʁnaval	kaʁnaval	'carnival'
b.	maj	maj	'hammer'
	detaj	detaj	'detail'
	evɑ̃taj	evɑ̃taj	'fan'

Historically, plural alternations were part of a larger pattern of l-vocalization (Pope 1952). French had a plural suffix [s], and general vocalization of velarized laterals, including the preconsonantal laterals which were uniformly velarized. Thus the paradigm [mal ∼ małs] 'evil, sg./pl.' of Gallo-Roman (7th century) turned to [mal ∼ maus] and then monophthongized to [mal ∼ mos] by the end of Early Old French (12th century). Soon after, the plural suffix was lost in all but liaison context, producing the modern [mal ∼ mo]. The [j]-final nouns in (1b-2b) had a palatal lateral that velarized preconsonantally, and thus followed a similar path, from e.g. [baʎ ∼ baʎs] 'lease' to [baʎ ∼ bałs] to [baʎ ∼ baus] to [baʎ ∼ bos] and with the loss of the plural suffix, [baʎ ∼ bo]. The simplification of the palatal lateral in the 18th century produced the modern [baj ∼ bo] (Pope 1952).

Up until Early Old French (12th century), all lateral-final nouns were affected without exception, including, e.g. [bal ∼ bos] 'ball', now [bal ∼ bal] (Pope 1952). Subsequently, some lexical items started losing the alternation, either keeping the lateral in both forms, like [bal], or losing it in both, like [ʃəvø] 'hair', originally [ʃəvɛl ∼ ʃəvø]. Why some items kept the alternations and others did not is unknown. It is often suggested that loss of alternations starts with infrequent items (see e.g. Bybee 1995, 2001). However, as we will see in §2.2, the alternations were mostly lost in short words, and those are generally more frequent.

From the time of Modern French (16th century) and onwards, neologisms and loanwords making their way into the language do not alternate, e.g. [ʃakal] 'jackal' from Turkish, [mistral] 'the mistral wind' from Provençal. Similarly, all modern loanwords are non-alternators. The alternation became effectively frozen, or unproductive.

Completing our review of French plural morphology, we note that in addition to the [al/aj]-final nouns discussed above, only five other nouns have plurals that are different from their singulars: [sjɛl] 'sky', which in addition to the regular plural [sjɛl] 'skies', may still be associated with the plural [sjø] 'heavens', [œj] 'eye', which has the suppletive plural [(z)jø], and the three fricative-final nouns [œf ∼ ø] 'egg', [bœf ∼ bø] 'bull', and [ɔs ∼ o] 'bone', which lose their final fricative. We see, then, that the plural morphology of French is overall rather regular, and all plurals that are audibly different from their singular end in [o] or [ø]. The situation is even simpler in feminine nouns, which never change in the plural.

We should also mention that [l/j] alternations affect five adjectives that acquire a final [l] or [j] in the feminine and before vowel-initial singular nouns (liaison context; for a recent review of liaison,

see Côté 2011): [fu ∼ fɔl] 'crazy', [mu ∼ mɔl] 'soft', [bo ∼ bɛl] 'beautiful', [nuvo ∼ nuvɛl] 'new', and [vjø ∼ vjɛj] 'old'. Before a vowel-initial noun like [ami] 'friend', these adjectives give rise to what looks like plural alternations, e.g. [vjɛj ami] 'old friend' ∼ [vjø z ami] 'old friends'. Before a consonant-initial noun, these adjective stay unchanged in the plural.

2.2 Trends in the French irregular plural

To assess the prevalence of the alternations among the real words of French, we extracted all 1057 of the masculine [al/aj/ɛl/ɛj]-final nouns and adjectives from ***Lexique*** (New et al. 2001), an electronic dictionary of French that lists 143,000 words. Since alternations with [ɛ] are limited to the single word [sjɛl], we focused on the [a]-final items. A native speaker of French went over the [a] items and marked all the familiar ones as alternating, non-alternating, or variable. Unknown items were discarded, leaving 118 masculine nouns and adjectives. The results are shown in Figure 1, where we see that the alternations impact polysyllabic [al/aj]-final items more than monosyllabic ones. The dataset is available at [link redacted].

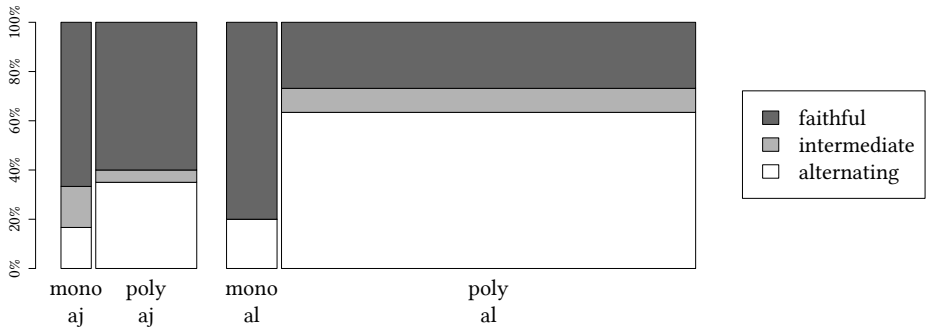

Figure 1: French [al/aj]-final singular nouns and adjectives (***n***=118), by monosyllabicity and final consonant. Monosyllables are less likely to take the alternating [o] plural.

To measure the strength and reliability of the patterns seen in Figure 1, and to make predictions about nonce words, we fitted a logistic regression model using the ***glm()*** function in R (R Development Core Team 2011). The dependent variable was a binary distinction between alternating and non-alternating plurals. Variable words, which can take either kind of plural (15 of the 118 items), were counted as non-alternators; we also tried counting them as alternators, which led to nearly identical results. The predictors in the model were ***consonant***, a binary factor that contrasted [l] and [j], and ***monosyllabic***, a binary factor that contrasted monosyllables with polysyllables. Model comparison confirms that ***monosyllabic*** makes a significant improvement to the model beyond that of ***consonant*** (χ^2***(1)*** $= 7.43$, $\boldsymbol{p} < .01$). The interaction does not make a significant improvement. The final model, in Table 1, enjoys low collinearity (κ=1.17).

The model in Table 1 shows that final [l] is conductive to significantly more alternations than final [j], and that polysyllables are conductive to significantly more alternations than monosyllables. In other words, monosyllables are protected from alternations, and this effect is predicted to apply to

	β	SE(β)	z	p-value
(Intercept)	0.07	0.2	0.34	
monosyllabic	−0.58	0.23	−2.46	<.05
consonant=l	0.44	0.2	2.16	<.05

Table 1: Regression model for the French lexicon

novel items. While the effects are not particularly strong, barely reaching significance, we will see in §3 below that French speakers extend the ***monosyllabicity*** effect very strongly, and also extend the ***consonant*** effect, though not as strongly.

To make sure that the trends in Table 1 are not confounded by token frequency, we added token frequency, as recorded in Lexique, to the model. While this made a significant improvement to the model (χ^2***(1)*** $= 7.51$, $p < .01$), token frequency itself did not reach significance; it merely served to increase the reliability of ***monosyllabicity***. More importantly, token frequency has no bearing on the predicted behavior of nonce words, which are all equally new to speakers.

2.3 Initial syllable faithfulness protects monosyllables

We have seen that the [al/aj → o] alternation impacts a larger proportion of polysyllables than monosyllables in the lexicon; in §3, we will also see that speakers represent this trend in their grammar, as evidenced by the fact that they extend the trend to novel items. We propose that initial syllable faithfulness (Trubetzkoy 1939; Steriade 1994; Beckman 1997, 1998; Casali 1998; Barnes 2006; Jesney 2009; Becker 2009; Becker et al. 2011a,b) is responsible for the effect. In polysyllables like [bo.kal ∼ bo.ko] 'jar', the initial syllable [bo] stays intact; thus, the alternation only violates general faithfulness, and there is no violation of initial syllable faithfulness. In a monosyllable like [mal ∼ mo] 'evil', however, the alternation impacts the initial syllable, and thus violates both general faithfulness and initial syllable faithfulness.

As discussed in §2.1 above, the alternation arose from a series of natural steps, including l-vocalization, monophthongization of [au] to [o], and loss of coda [s]. The alternation is no longer natural, since with the disappearance of the plural suffix, the environment conditioning the change is gone. The change itself is no longer natural either, especially in the case of [aj → o], where [aj] could plausibly fuse to [e], but not to [o]. It should be noted, however, that the alternation is not completely arbitrary in contemporary French, since it applies to the natural class of continuant sonorants. While French has the continuant sonorants [w] and [ɥ] in addition to [l] and [j], the former are not allowed word-finally (except for the interjection [waw] 'wow'). The French [ʁ] is a fricative, not a sonorant. Thus, all the continuant sonorants that are allowed word-finally in nouns participate in the alternation.

To offer a concrete analysis, we suggest that the plural affix is a floating [o], which docks on the stem's final [a]. In an autosegmental theory of mutation (Wolf 2007), the floating affix is required to

dock by MAX(float). This in turn causes the stem to surface unfaithfully, changing the lowness and roundness of the stem's final vowel, and thus creating violations of IDENT(round) and IDENT(low). Either of these IDENT constraints can be sensitive to being in the word initial syllable; we choose IDENT(low) here.

As seen in (3-4), MAX(float) is outranked by initial syllable faithfulness, preventing monosyllables from alternating. Polysyllables are allowed to alternate, since MAX(float) outranks general faithfulness.

(3) [bal ∼ bal] 'ball' is protected from alternation by IDENT-σ1(low)

/bal/ + /o/	IDENT-σ1(low)	MAX(float)	IDENT(low)
a. ☞ bal		*	
b. bo	*!		*

(4) [bo.kal ∼ bo.ko] 'jar' allows the floating /o/ to dock

/bokal/ + /o/	IDENT-σ1(low)	MAX(float)	IDENT(low)
a. bo.kal		*!	
b. ☞ bo.ko			*

The ranking in (3)-(4) is supplemented by two additional constraints involving right-alignment of the affixal [o] with the stem and deletion of the stem-final consonant. With ALIGN dominating MAX, the vowel-final output is selected (5).

(5) Root-final consonant deleted due to ALIGN ≫ MAX

/bokal/ + /o/	IDENT-σ1(low)	MAX (float)	IDENT (low)	ALIGN	MAX
a. bo.kal		*!			
b. bo.kol			*	*!	
c. ☞ bo.ko			*		*

The statistical model in Table 1 of §2.2 predicts that [l] is more deletable than [j]; we can incorporate this fact into our analysis by relativizing MAX to particular types of deleted segments. Additionally, [o] is only allowed to dock on a stem [a]; this restriction can be modeled with high-ranking Output-Output faithfulness constraints (Benua 1997 et seq.) that prohibit the change of any vowel feature other than [low] and [round].

The idealized grammar in (6), with MAX(float) sandwiched between general faithfulness and initial syllable faithfulness, uniformly impacts all polysyllables and protects all monosyllables. The actual grammar of French is more noisy, with some monosyllabic lexical items impacted (such as [mal]) and some polysyllables protected (such as [kaʁnaval] 'carnival').

(6) IDENT-σ1(low) ≫ MAX(float) ≫ IDENT(low)

The departure from the idealized grammar in (6) can be achieved with one of the available theories of exceptionality, either constraint cloning (Pater 2006, 2009; Coetzee 2008; Becker 2009; Becker, Ketrez & Nevins 2011a) or UseListed (Zuraw 2000; Hayes & Londe 2006; Becker, Nevins & Levine 2011b). Both approaches incorporate the lexical statistics in terms of the constraint rankings they acquire, and then project the learned statistics onto novel items.

The grammar in (7) exemplifies the constraint cloning approach: each faithfulness constraint has two copies, or clones. The clones of Ident-σ1(low) classify monosyllables into alternators with the low-ranking clone and non-alternators with the high-ranking clone, and polysyllables are similarly listed with clones of Ident(low).

(7) $\text{Ident-}\sigma1(\text{low})_{\text{bal, maj}} \gg \text{Ident(low)}_{\text{ʃakal, detaj}} \gg \text{Max(float)} \gg \text{Ident-}\sigma1(\text{low})_{\text{mal, baj}}, \text{Ident(low)}_{\text{bokal, ʒuʁnal, tʁavaj, supiʁaj}}$

When the grammar in (7) is applied to a novel word, clones that have more items listed with them exert greater influence. Since most polysyllabic lexical items alternate, the low-ranking clone of Ident(low) lists more items than the high-ranking clone, and thus a polysyllabic nonce word will likely alternate. The converse is true for Ident-σ1(low), so a monosyllabic nonce word will likely stay faithful.

To summarize, our analysis predicts that speakers construe the difference between short nouns and long nouns in terms of specific faithfulness constraints relativized to monosyllables. The analysis relies on a floating [o] plural suffix, which allows the change to occur synchronically even if it is no longer natural. In §3 below, we show that speakers extend the effect as predicted, from the words in their lexicon to nonce words.

3 Experiment 1: French nonce words

The irregular alternation in French contains many apparent eccentricities: the nature of the structural change is not synchronically natural, and the difference between [l]- and [j]-final words is similarly not expected given general typological principles. Our main interest in the current study is the nature of generalizations encoded by native speakers, and in particular whether they will demonstrate knowledge of a discrete monosyllabic vs. polysyllabic faithfulness effect, even if the lexical evidence is limited.

3.1 Materials

We created a total of 100 [l/j]-final items: 66 items with [a] in their final syllable (36 monosyllabic and 30 polysyllabic) and 34 items with [ɛ] (18 monosyllabic and 16 polysyllabic). All the items were made in pairs, with [l] in one member and [j] in the other, e.g. [snal, snaj]. The items' onsets were

designed to span a wide range of phonotactic wellformedness, from extremely common onsets like [d] and [ʁ] to extremely uncommon ones like [sn] and [spʁ] (see discussion of this point in §6). The full list of items, with by-item results, is in Appendix A. In addition, we chose 12 loanwords from English as fillers, e.g. [stɛk] 'steak', [dolaʁ] 'dollar'.

The items were recorded by a phonetically trained male native speaker of French in his twenties from the Ferté-Bernard region (∼100 miles from Paris) who was naive to the purpose of the experiment. The list included each noun in the faithful plural (which in French is identical to the singular) and unfaithful plural, which replaced [al, aj, ɛl] with [o] and [ɛj] with [ø]. The recording was made in a quiet room into a Macintosh computer. The full list was recorded in three randomly generated orders. The best token of each item was chosen and converted to mp3 format using the LAME converter. The recordings were not manipulated in any other way, other than normalizing the intensity with Praat (Boersma & Weenink 2011), and sounded quite natural.

3.2 Methods

The experiment was run in `Experigen` (Becker & Levine 2010). It was presented to the participants over the internet, using the web browser of the participant's choice. The web server executed a random selection of materials for each participant, choosing a total of 20 items: 12 target items and 8 fillers. In addition, the experiment started with the sample item [dal], which we include in the analysis.

In addition to [dal], the 12 selected target items consisted of six monosyllables and six polysyllables, where each size included 3 [al]-final, 1 [aj]-final, 1 [ɛl]-final, and 1 [ɛj]-final items. The 8 fillers were chosen randomly.

To make sure that the items were treated by speakers as masculine nouns, 20 frame sentences were created (recall that feminine nouns don't alternate). Each frame contained two phrases: the first included a placeholder for a singular noun, and the second included a placeholder for the plural form. Each frame contained at least one determiner or adjective that has phonologically distinct masculine and feminine forms, e.g. [gri] 'gray.MASC', which is distinct from [griz] 'gray.FEM'. These 20 frames were randomly paired with the 12 target items. In addition, frame sentences were created for each of the 12 fillers.

Before the experiment began, the participant was introduced to the made-up item [dal], and was asked to indicate their preference between a faithful plural and an alternating plural (presented in French as a choice between "pluriel en s" and "pluriel en x", referring to French orthography, which marks regular plurals of l/j-final nouns with ⟨s⟩ and vowel-changing irregular plurals with ⟨x⟩). Then, the item [dal] appeared in its frame sentence. A sound button played [dal] when pressed, and upon completion, another sound button appeared. When pressed, [do] was played, and then seven numbered buttons as in (8) appeared between the two sound buttons. Once one of the seven buttons was pressed, the next item appeared.

(8) Quel pluriel préférez-vous?
celui à gauche [1] [2] [3] [4] [5] [6] [7] celui à droite

Which plural do you prefer?
the one on the left [1] [2] [3] [4] [5] [6] [7] the one on the right

The order of presentation of the two plurals was randomized. After participants responded to the randomized list of 20 target items and fillers, they were asked to answer a few demographic questions (country of origin, current location, year of birth, sex, other languages spoken, suggestions).

3.3 Participants

The participants were recruited online using word of mouth. They all volunteered their time and effort. We gathered data from 185 people who completed the survey and self-identified as being born in France and being at least 18 years old, discarding the rest. The server logs indicate that these 185 participants took on average 4 minutes to complete the experiment (range 2–17 minutes, median 4). The participants reported an average age of 36 (range 18–74, median 32). Sex: 116 females, 65 males, 4 did not say. Other languages spoken: 83 participants listed no other language or indicated that they are monolingual. Of those who listed other languages, 76 indicated some knowledge of English, 18 Spanish, 15 German, and smaller numbers of other languages.

3.4 Results

On average, speakers demonstrated an asymmetry in the treatment of monosyllables, preferring faithful plurals for monosyllabic items and alternating [o/ø] plurals for polysyllabic ones (3.5 vs. 4.3 on the 1–7 scale). In addition, they preferred faithful plurals with [ɛ] and alternating plurals with [a] (2.7 vs. 4.4). Both of these effects can be seen in the bean-plots in Figure 2 (bean-plots are density plots with an added horizontal line to mark the mean). Figure 2 also shows an interaction: the monosyllabicity effect is much stronger for [a] than it is for [ɛ]. The raw results are available at [link redacted].

The difference between [j]-final and [l]-final items was overall rather small (3.6 vs. 4.0), with most of it concentrated in the polysyllabic [a] items (4.8 vs. 5.0).

To assess the statistical strength of these effects, we used a mixed-effects regression model using the ***lmer()*** function from the ***lme4*** package (Bates et al. 2011) in R. We used the following predictors: ***monosyllabic***, a binary predictor that contrasted monosyllables and polysyllables, ***vowel***, a binary predictor that contrasted [a] with [ɛ], and ***consonant***, a binary predictor that contrasted [l] with [j]. The effects that reached statistic significance are reported in Table 2.[2]

[2] We started with a base model that had ***item*** and ***participant*** as random effects, and the full crossing of ***monosyllabic***, ***vowel***, and ***consonant*** as fixed effects. To reduce the correlations between the predictors in the model, we first normalized the predictors using R's ***scale()*** function. We then removed the three-way interaction and the interaction of ***vowel*** and

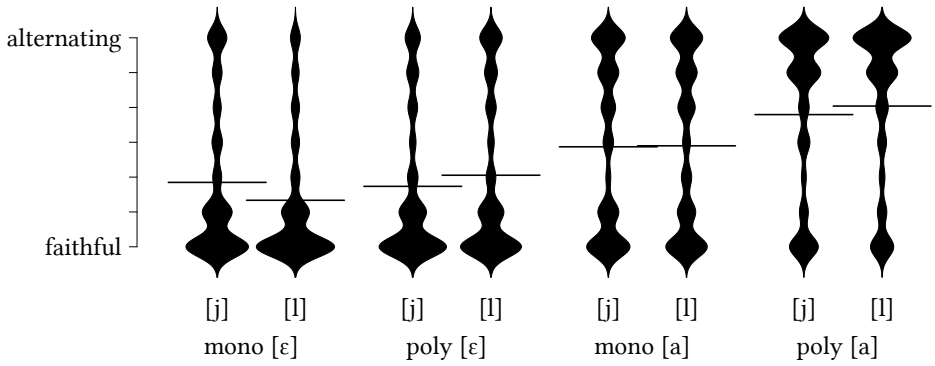

Figure 2: French nonce monosyllables protected from alternations

	β	SE(β)	***t***	***p*-value**[3]
(Intercept)	3.91	0.10	40.99	
monosyllabic	−0.39	0.06	−6.19	<.0001
vowel=[a]	0.76	0.06	11.84	<.0001
consonant=[l]	0.02	0.05	0.31	>.1
monosyllabic*:*vowel	−0.16	0.06	−2.80	<.01
monosyllabic*:*consonant	−0.10	0.04	−2.36	<.05

Table 2: Regression model for French nonce words

The model in Table 2 shows that alternating plurals (ending in [o] or [ø]) are highly significantly preferred in polysyllables and with [a]. Additionally, there are weaker (=smaller β) yet significant interactions showing a preference for more alternations with [l] and with [a] in the polysyllables.

3.5 Comparison with the lexicon model

The effects in Table 2 largely reflect those in the lexicon, where alternations are preferred with polysyllabic items, and more alternations with [l]. There are some small differences, e.g. the difference between [l] and [j] was a main effect in the lexicon, but in the experiment, it only showed up in the polysyllables.

Comparing the predictions of the lexicon model in Table 1 to the participants' responses per item, we see a strong and highly significant correlation, shown in both panels of Figure 3 (Spearman's rank correlation, ρ=.75, $\boldsymbol{p} < .0001$). In other words, there is an excellent match between the predictions of our lexicon model and the observed treatment of nonce words.

consonant, each time making sure that there is no significant change to the model using ANOVA model comparison. Then, we added the predictors shown in Table 2 as random slopes to ***item*** and ***participant***. The full model did not converge; we report in Table 2 the model with random slopes for ***participant*** only. The final model has low collinearity measures (VIF≤1.07, κ=1.59), calculated using ***mer-utils***, by Austin Frank, available at `https://hlplab.wordpress.com/2011/02/24/diagnosing-collinearity-in-lme4/`.

[3]Exact p-values could not be obtained with ***lme4*** or ***pvals.fnc()***, due to disagreement in the community about the best

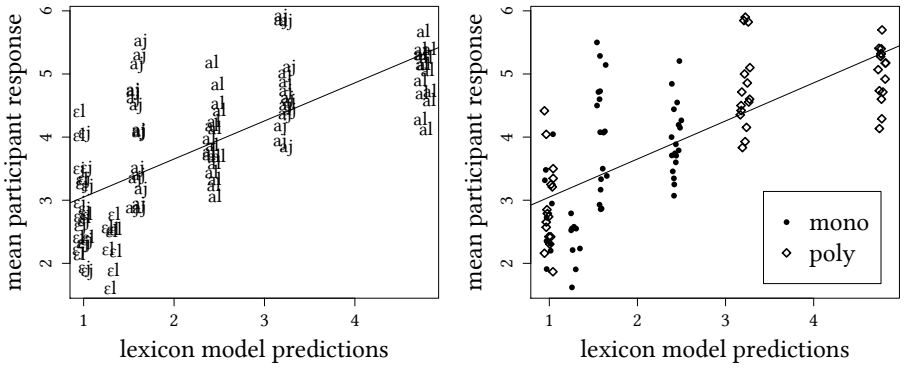

Figure 3: French lexicon predictions (from Table 1) vs. human responses to nonce items, shown by final vowel and consonant (left) and by prosodic shape (right)

3.6 Discussion

The experiment confirms that speakers have access to the lexical trends that we described in §2, and that they extend these trends from the real words of their language to novel words. In particular, speakers prefer the alternation in polysyllabic words, keeping monosyllables relatively protected.

The lexical trends are projected from the lexicon even though the alternation is no longer productive or natural. We take this pattern as evidence that speakers organize their lexicon using grammatical principles even when the alternation is frozen. While it seems that there is no "benefit" to encoding the size-based faithfulness effect for synchronic speakers, the discrete and grammatical treatment of this pattern is apparently irresistible.

4 The Brazilian Portuguese lexicon

In this section we turn to a related study in another language, where monosyllabicity again plays a striking explanatory role, despite the absence of a specific origin in terms of monosyllabicity in the historical development of the process. The laterals that gave rise to the [al/aj → o] alternation in French also play a role in Portuguese, where we have a [w → j] alternation. While the languages have a common Latin root, the changes occurred after the languages separated, and the details of historical development are different. In particular, while the alternation is no longer natural in either language, the Portuguese alternation is very much alive, extending productively to polysyllabic words and loanwords.

Anticipating the discussion of phonetic duration in §6, we note that while the French alternation involves the process of shortening (a single [o] is shorter than [al] or [aj]), the Portuguese alternation keeps duration constant, with the back and front glides being the same length.

way to calculate them. We estimated p-values based on the ***t*** statistic.

4.1 The Brazilian Portuguese regular plural

For most nouns in Portuguese, the plural is completely regular and predictable. Nouns that end in [ɦ] or [s][4] take the suffix form [-is], e.g. ['floɦ ∼ 'floɾis] 'flower', ['vɔ(j)s ∼ 'vɔzis] 'voice'. No suffix is added on nouns that already end in an unstressed vowel and [s] in the singular, e.g. ['lapis] 'pencil', ['õnibus] 'bus'. Nouns that end in an oral vowel or [j], or any nasal vowel other than [ɐ̃w̃], take [-s], e.g. ['baɦku ∼ 'baɦkus] 'ship', [e'ɾɔj ∼ e'ɾɔjs] 'hero', ['ifẽ ∼ 'ifẽs] 'hyphen'.

These regular cases cover the vast majority of Portuguese nouns, which overwhelmingly end in a vowel, a glide, or one of the consonants [ɦ, s]. We see that two plural suffixes are available, in complementary distribution: [-s] after vowels and glides, otherwise [-is] after consonants. This leaves the nouns that end in [w] or in [ɐ̃w̃]. The rest of this paper focuses on [w]; for the similar yet distinct behavior of [ɐ̃w̃], see Huback (2007).

4.2 The plural of [w]-final nouns

Most of the [w]-final nouns of the language change the [w] to [j] in the plural, e.g. [a'nɛw ∼ a'nɛjs] 'ring', as in (9). We analyze this alternation pattern as taking the plural suffix /-is/, with concomitant fusion of the stem's [w] and the suffix's [i].

When the stem's final [w] is preceded by [i], the [w] is lost completely, with the outcome depending on the position of stress. With final stress, the stem's [w] is simply lost, as in [ba'hiw ∼ ba'his] 'barrel'. With penultimate stress, the stem's preceding [i] changes to [e], as in ['hɛpʧiw ∼ 'hɛpʧejs] 'reptile'. See Huback (2007) for further detail on these changes.

(9) Brazilian Portuguese alternating, unfaithful plurals

shape	singular	plural	spelling	
mono	'saw	'sajs	⟨sal⟩	'salt'
	'mɛw	'mɛjs	⟨mel⟩	'honey'
	'pɾɔw	'pɾɔjs	⟨prol⟩	'advantage'
iamb	ʒoɦ.'naw	ʒoɦ.'najs	⟨jornal⟩	'newspaper'
	a.'nɛw	a.'nɛjs	⟨anel⟩	'ring'
	ba.'hiw	ba.'his	⟨barril⟩	'barrel'
trochee	'ni.vew	'ni.vejs	⟨nível⟩	'level'
	'hɛp.ʧiw	'hɛp.ʧejs	⟨réptil⟩	'reptile'

In contrast to nouns that take the unfaithful plural, other [w]-final nouns form their plural with a simple suffixation of [s] and no further change, e.g. ['gow ∼ 'gows] 'goal', as in (10). We analyze these as taking the plural suffix /-is/ with concomitant deletion of the suffixal [i].

[4] The word-final pronunciation of these two consonants varies greatly between dialects. The former can be heard as [x], [χ], [h], or [ɻ], the latter as [ʃ]. These sounds usually assimilate in voicing to a following consonant. They always give rise to [ɾ] and [z], respectively, when suffixed with the plural.

(10) Brazilian Portuguese non-alternating, faithful plurals

shape	singular	plural	spelling	
mono	ˈpaw	ˈpaws	⟨pau⟩	‘stick’
	ˈpnɛw	ˈpnɛws	⟨pneu⟩	‘tire’
	ˈgow	ˈgows	⟨gol⟩	‘goal’
iamb	ka.ˈkaw	ka.ˈkaws	⟨cacau⟩	‘cocoa’
	mu.ˈzew	mu.ˈzews	⟨museu⟩	‘museum’
	fu.ʧi.ˈbow	fu.ʧi.ˈbows	⟨futbol⟩	‘football’
trochee	ˈaw.kow	ˈaw.kows	⟨alcool⟩	‘alcohol’

A very small number of nouns, most of them monosyllabic, take the plural suffix [-is] with the stem’s [w] surfacing as [ɬ], e.g. [ˈmaw ∼ ˈmaɬis] ‘evil’. We consider these few items to be faithful plurals as well, in that they keep the backness of the stem. The surfacing lateral is not the front lateral [ʎ] that Portuguese has in words like [fiʎu] ‘son’, but rather the lightly velarized alveolar [ɬ]. Thus, [ˈmaw ∼ ˈmaɬis] involves a violation of [±lateral], but not a violation of [±back]. The experiment we report in §5 below does not examine this highly infrequent type of plural, but we predict that it will pattern like other faithful plurals.

Historically, unfaithful plurals originate from nouns with a final lateral. Deletion of intervocalic laterals, which happened across the board in Galician-Portuguese (9th century), affected the plural, taking [ˈsal ∼ ˈsalis] ‘salt’ to [ˈsal ∼ ˈsais]. Much later (18th century), coda [l] vocalized to [w], giving [ˈsaw ∼ ˈsais], with hiatus resolution creating the glide in [ˈsaw ∼ ˈsajs]. The faithful plurals are traced back to nouns that originally ended in [u] or [w], which simply took the plural [s], e.g. [muˈzew ∼ muˈzews] ‘museum’. Some nouns with faithful plurals had an intervocalic lateral that deleted in both singular and plural, e.g. [ˈpalo ∼ ˈpalos] ‘stick’ leading to [ˈpao ∼ ˈpaos] and from there via raising of unstressed mid vowels and hiatus resolution to the modern [ˈpaw ∼ ˈpaws].

With the vocalization of final laterals, the historical distinction between final [l] and final [w] can only be heard as a distinction between alternating and non-alternating plurals. This naturally leads to some fluctuation: some nouns that have normative faithful plurals have developed an innovative unfaithful plural, e.g. [deˈgraw] ‘step’ is often heard pluralized colloquially as [deˈgrajs]. Similar variation is seen in a number of nouns, e.g. [ʃaˈpɛw] ‘hat’, [tɾoˈfɛw] ‘trophy’, as documented in Huback (2007) and Gomes & Manoel (2010). These nouns are all polysyllabic with a lax stressed vowel; as we will see in §4.3, both polysyllabicity and lax vowels are conducive to unfaithful plurals. The innovation, then, is driven by extension of the existing trends in the lexicon. On the other hand, [ˈsaw] ‘salt’, which has the normative unfaithful plural [ˈsajs], has the innovative faithful plural [ˈsaws]. Here, monosyllabicity is the factor that causes this noun, and other monosyllabic nouns, to develop a faithful plural in the other direction, nonetheless extending the trend for monosyllables.

These same trends are also extended to loanwords, with monosyllables such as [ˈgow], from English ‘goal’, receiving faithful plurals, and polysyllables such as [kokeˈtɛw], from English ‘cocktail’, receiving unfaithful plurals.

To summarize, the [w → j] change was created by a series of natural steps that resulted in a synchronically unnatural alternation. While the change happens in the environment of the overt plural suffix, it can hardly be described as assimilation, lenition, or any other such natural process. Thus, the plurals in both French and Portuguese represent synchronically unnatural patterns, which are nevertheless extended productively to nonce forms. In French, the extension is limited to experimental settings, whereas in Portuguese, the pattern is strongly productive outside the lab as well.

4.3 Trends in the irregular plurals

We collected the [w]-final nouns and adjectives and their plurals from two word lists.[5] We then presented them to three native speakers of Brazilian Portuguese, and kept only those items which were familiar and had a distinct plural form for at least one speaker. This left 387 w-final items: 32 monosyllables, 47 trochees (polysyllables with penultimate stress) and 308 iambs (polysyllables with final stress). The plural of each item was marked as either unfaithful, faithful, or variable (=50% faithful). When our three speakers did not agree, we averaged their preferences. The dataset is available at [link redacted]. Each item was coded for its number of syllables and position of stress, as well as the stressed and final vowel sounds.

The mosaic plot in Figure 4 shows all 387 nouns in our list, plotted by prosodic shape. We see that only about 29% of the monosyllables take unfaithful plurals, while 88% of polysyllables do, i.e. monosyllables are protected from alternation. We note that the rate of unfaithful plurals is lower for iambs (87%) than it is for trochees (96%).

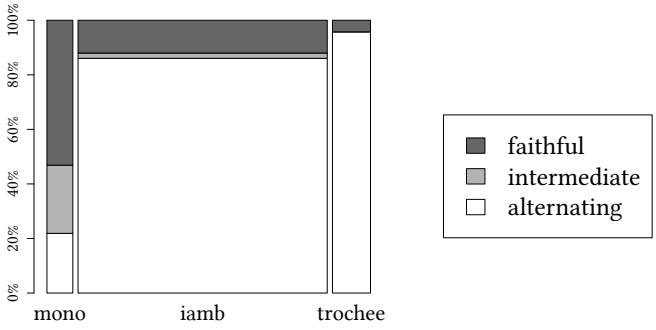

Figure 4: Brazilian Portuguese w-final words (***n***=387), by prosodic shape. Monosyllables are less likely to take the alternating plural.

Another factor that correlates with the choice of plural is the quality of the stem's final vowel, particularly its tenseness, as seen in Figure 5. Among the monosyllables, the lax vowels [a, ɛ, ɔ] are most conducive to alternating plurals, while the tense mid vowels [e, o] are most likely to have

[5]One is LABEL-LEX, available at `http://label.ist.utl.pt/pt/apresentacao.php`. The other is available at `http://artsci.wustl.edu/~tcpollo/palavras_explic.html`.

faithful plurals. The difference between [ɛ] and [e] is also seen with iambs (although there are no items with [o] to compare to).

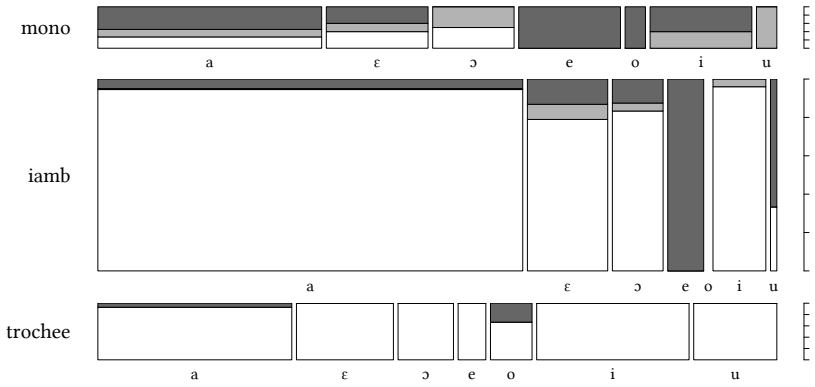

Figure 5: Brazilian Portuguese w-final words (***n***=387), by prosodic shape and stressed vowel. Monosyllables and tense vowels are less likely to take the alternating plural.

We used a regression analysis with R's ***glm()*** function to determine the strength and reliability of the correlation between the properties of nouns and the kind of plural they take, and at the same time, make predictions about the treatment of novel nouns. The independent variable was a binary distinction between faithful vs. alternating plurals. Since 20 of the 387 items had an intermediate faithfulness level, we had to choose a cutoff level. We tried a few different levels, none of which made any difference to speak of, so we used 50% as the cutoff faithfulness value in the regressions we report below.

We started with a model that had ***prosodic shape*** as a predictor, backward-difference coded as one factor that contrasts monosyllables and iambs, and one that contrasts iambs and trochees. We also included the following factors that encoded the stressed vowel's features: ***back*** (a binary factor), ***high*** (a binary factor), ***lax*** (a three-level factor contrasting the lax [a, ɛ, ɔ] with the tense mid [e, o], and high vowels at zero), ***low*** (a three-level factor contrasting [a] with [ɛ, ɔ], and the other vowels at zero), and the interaction of ***back*** with ***high*** and with ***low***. The vowel predictors were so defined as to minimize collinearity between them. Other interactions were either completely or nearly nonidentified (see §5.8 of Gelman & Hill 2007). We then pared the model down, each time removing predictors that do not improve the model, arriving at the model in Table 3. This model has acceptably low collinearity (κ=13.8).

The model in Table 3 confirms that iambs are significantly more likely to take alternating plurals than monosyllables, and trochees are significantly more likely to take alternating plurals than iambs. The lax vowels [a, ɛ, ɔ] are significantly more likely to take alternating plurals than the tense mid vowels [e, o]. Additionally, there is a small (=small β) yet significant effect of more faithful plurals following the high back vowel [u].

	β	$SE(\beta)$	z	$p(>\|z\|)$
(Intercept)	0.10	0.38	0.27	
mono vs. iamb	3.21	0.49	6.60	$<.0001$
iamb vs. trochee	3.52	1.24	2.85	$<.005$
lax	2.67	0.57	4.64	$<.0001$
high	0.20	0.29	0.69	$>.1$
back	0.08	0.22	0.39	$>.1$
high*:*back	-0.56	0.20	-2.78	$<.01$

Table 3: Regression model for the Portuguese lexicon

4.4 Initial syllable faithfulness protects monosyllables

In this subsection we provide an analysis of the asymmetry in plural formation between monosyllables and polysyllables in terms of faithfulness constraints specific to the initial syllable. Assuming that the plural suffix is underlyingly /-is/, [w]-final nouns form their plurals either by deleting the [i], as in /gow + is/ → [gows] (violating MAX), or by fusing the [w] and [i] to [j], as in /aˈnɛw + is/ → [aˈnɛjs], violating IDENT(back) and the anti-fusion constraint UNIFORMITY.[6]

The difference between monosyllables and polylsyllables can be attributed to an initial syllable-specific version of any of the three faithfulness constraints involved; we use IDENT(back) here. The analysis is outlined in the tableaux in (11-12) and largely parallels our analysis of French in §2.3. Here, IDENT-σ1(back) protect the monosyllables such as [ˈgow] from changing its stem [w], but MAX forces the alternation in the polysyllables such as [aˈnɛw].

(11) Monosyllabic [ˈgow] protected from alternation

/gow$_1$ + i$_2$s/	IDENT-σ1(back)	MAX	IDENT(back)
a. ☞ gow$_1$s		*	
b. goj$_{1,2}$s	*!		*

(12) Polysyllabic [aˈnɛw] fuses with the suffixal [i]

/aˈnɛw$_1$ + i$_2$s/	IDENT-σ1(back)	MAX	IDENT(back)
a. aˈnɛw$_1$s		*!	
b. ☞ aˈnɛj$_{1,2}$s			*

The fully faithful concatenation of the base and suffix is ungrammatical in Portuguese, e.g. *[gowis], *[aˈnɛwis]. We attribute this to a constraint against [w] in onset position, which holds without exception in the plural's derived environment, and also fairly generally in the language. Ranking of IDENT(lateral) below IDENT(back) will derive the [ˈmaw ∼ ˈmałis] type of plural.

The grammar in (11-12) is idealized; the actual grammar of Brazilian Portuguese will allow some monosyllabic items to alternate and some polysyllabic items to stay faithful. As we did for French

[6]We gloss over some of the issues involved with the definition of faithfulness constraints in cases of fusion; see Pater (1999) for discussion.

in §2.3, either constraint cloning or UseListed can be used to make the grammar derive existing lexical items correctly, and extend the trends in the lexicon to novel items.

To capture the vowel laxness effect, the constraint cloning approach would require constraints that prefer alternating plurals in the presence of lax vowels (i.e. *ɛw, *ɔw), or constraints that prefer faithful plurals in the presence of tense vowels (i.e. *ej, *oj). Since all four constraints are often violated in the language, their effect would have to be limited to derived environments. See Becker (2009) for technical discussion about combinations of lexical trends in the constraint cloning framework.

Finally, we note that our analysis always assumes that the singular's underlying representation is identical to its surface form, i.e. we do not encode the plural behavior of [w]-final nouns in their underlying representation. We need not assume an underlying lateral for nouns that take alternating plurals, despite the orthography, which is irrelevant to the trends at hand, as innovative plurals such as [ʃaˈpɛjs] 'hats' show, and clearly not at play in purely auditory experiments in which the ⟨w⟩/⟨l⟩ final distinction is neutralized. Since the plural behavior isn't encoded in the underlying representation, it must be captured by the grammar, which we argue, is exactly where it belongs, as it is needed to generalize to novel items.

To summarize, our analysis predicts that speakers monitor the distribution of faithful and unfaithful plurals in terms of monosyllabicity and vowel quality. In particular, the distinction between short words and long words is binary and discrete. In §5, we present speakers with a full crossing of prosodic shapes and vowel qualities and measure the effect of each factor.

5 Experiment 2: Brazilian Portuguese nonce words

5.1 Materials

We created a total of 89 [w]-final items: 47 monosyllabic, 21 trochaic, and 21 iambic. The full list of items, with by-item results, is in Appendix B. In addition, we constructed 21 [s]-final items as fillers, 7 each monosyllabic, trochaic, and iambic. As we did in French, the items' onsets were designed to span a wide range of phonotactic wellformedness, from extremely common onsets like [d] and [f] to extremely uncommon ones like [bɾ] and [dɾ] (see discussion in §6).

The items were recorded by a female native speaker of Brazilian Portuguese in her twenties from Rio de Janeiro. She received basic phonetic training, and was naive to the purpose of the task. The list included each noun in three forms: the singular, a faithful back plural, and an unfaithful front plural. The recording and processing was done as in §3.1.

5.2 Methods

The experiment was run in `Experigen` (Becker & Levine 2010), as in §3.2. The web server executed a random selection of materials for each participant, choosing a total of 24 items: 15 target items (5 of each shape) and 9 fillers (3 of each shape). In addition, the experiment started with the sample item [ʃaˈpɛw] 'hat', which is known to vary in the plural between the standard [ʃaˈpɛws] and the colloquial [ʃaˈpɛjs].

The items were presented in frame sentences that had a placeholder for a singular noun in a first phrase, followed by a second phrase with a placeholder for a plural noun. Upon pressing a first button, the singular and one of the plurals was played, and a second button appeared. When the second button was pressed, the singular and the other plural were played, and then seven numbered buttons appeared between the two play buttons, as in (13). Pressing one of the seven buttons moved the participant to the next item. The order of plurals was randomized.

(13) Qual plural você prefere?

Eu prefiro o da esquerda [1] [2] [3] [4] [5] [6] [7] Eu prefiro o da direita

Which plural do you prefer?

the one on the left [1] [2] [3] [4] [5] [6] [7] the one on the right

After the participant responded to all 24 target items and fillers, they were asked to answer a few demographic questions (country of origin, current location, year of birth, sex, other languages spoken, suggestions).

5.3 Participants

The participants were recruited online using word of mouth. They all volunteered their time and effort. We gathered data from 181 people who completed the survey and self-identified as being born in Brazil and being at least 18 years old; other data was discarded. The server logs indicate that these 181 participants took on average 6.4 minutes to complete the experiment (range 3–43 minutes, median 5). The participants reported an average age of 30 (range 18–71, median 27). Sex: 95 females, 51 males, 35 did not say.

When asked about the kind of Portuguese they speak, 97 participants listed São Paulo, 23 listed Rio de Janeiro, 21 Minas Gerais, 7 Porto Alegre, and a smaller number of other places in Brazil. Other languages spoken: 13 participants listed no other language or indicated that they are monolingual. Of those who listed other languages, 46 indicated some knowledge of Spanish, 43 French, 14 English, and a small number of other languages.

5.4 Results

On average, speakers preferred faithful plurals for monosyllabic items (3.9 on the 1–7 scale), and unfaithful plurals for polysyllabic ones, more strongly so for iambs (5.1) than for trochees (4.6), as seen in Figure 6. The raw results are available at [link redacted].

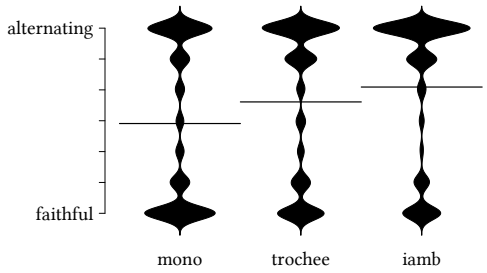

Figure 6: Monosyllables protected from alternations in Brazilian Portuguese nonce words.

As for the effect of the vowel that preceded the word-final [w], faithful plurals were chosen most often with the tense mid vowels [e,o] (4.3), and unfaithful plurals were chosen most often with the lax mid vowels [ɛ, ɔ] (5.0). This is seen in Figure 7.

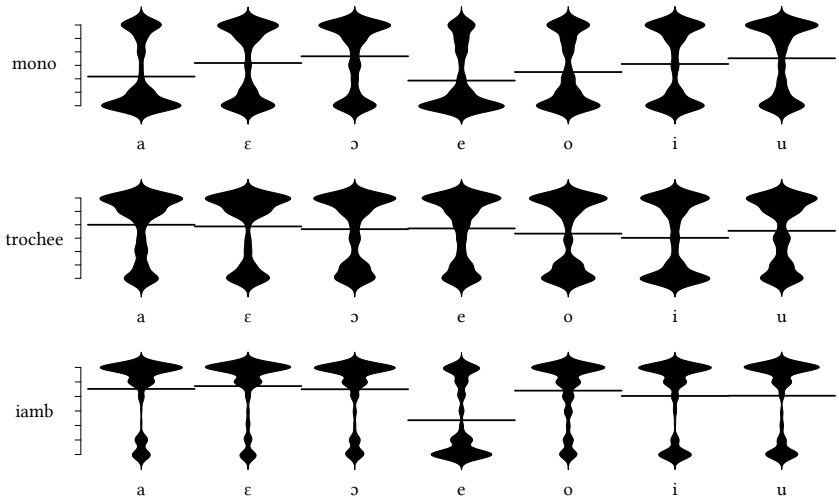

Figure 7: [ɛ, ɔ] most conducive to alternation, [e, o] least conducive in Brazilian Portuguese nonce words.

To assess the statistical strength of these effects, we used a mixed-effects regression model using the ***lmer()*** function from the ***lme4*** package in R. We Helmert-coded the ***prosodic shape*** (monosyllabic, trochaic, iambic) as two binary factors, one that constrasted monosyllables with trochees (since trochees were on average closer to monosyllables), and one that contrasted monosyllables and trochees with iambs. For the vowel effects, we coded the stressed vowels for all items as we did for the lexicon model in §4.3, and for the trochees, also the final unstressed [e, o]. This gave us five predictors: ***back***, ***high***, ***lax***, ***low***, all defined as in §4.3, and for the unstressed final vowels of

trochees, the binary predictor ***unstressed back***. The effects that reached statistic significance are reported in Table 4.[7]

	β	SE(β)	t	p-value[8]
(Intercept)	4.47	0.10	42.91	
mono vs. trochee	0.35	0.09	3.80	$<.0005$
mono & trochee vs. iamb	0.55	0.11	4.77	$<.0001$
lax	0.43	0.09	4.63	$<.0001$
unstressed back	-0.42	0.15	-2.86	$<.005$
low	-0.40	0.26	-1.53	$>.1$
mono vs. trochee*:*low	0.72	0.31	2.35	$<.05$
mono & trochee vs. iamb*:*low	0.26	0.38	0.69	$>.1$

Table 4: Regression model for Brazilian Portuguese nonce words

The model in Table 4 shows that unfaithful, alternating plurals are highly significantly preferred in polysyllables, both trochees and iambs, and significantly more so in iambs. Alternating plurals were also highly significantly preferred following the lax vowels [a, ɛ, ɔ]. Additionally, alternating plurals were preferred following unstressed front vowels, and dispreferred following [a] in trochees.

We used Helmert coding for ***prosodic shape*** because it reduces collinearity; backward-difference coding made collinearity in the model unacceptably high. However, Helmert coding does not allow a simle comparison of trochees and iambs. To confirm specifically that the difference between trochees and iambs is significant, we ran a separate model on the polysyllabic items only. This model showed a significant difference between iambs and trochees ($p < .05$), above and beyond the effect that vowel features have on polysyllables (ANOVA model comparison, $\chi^2(1) = 5.77$, $p < .05$). While the use of two different models to establish the difference between the different prosodic shapes is not ideal, it enabled the assessment of two highly trustworthy models (as measured by κ and VIF). The amount of data and the strength of the effects alleviate any concerns about Type I errors in this case.

5.5 Comparison with the lexicon model

The lexicon model offered in §4.3 made the prediction that monosyllables would be preferred with faithful plurals, and that polysyllables would be preferred with unfaithful plurals. Additionally, it predicted that trochees would be preferred with unfaithful plurals more strongly than iambs. In §5.4, we saw the lexical preference for faithful plurals on monosyllables extended to nonce words.

[7] We started with a base model that had ***item*** and ***participant*** as random effects. The fixed effects were prosodic shape and its interactions with ***back***, ***high***, ***lax***, and ***low***, and ***unstressed back***. To reduce the correlations between the predictors in the model, the binary factors were normalized using R's ***scale()*** function. We then removed the interactions of prosodic shape with ***back***, ***lax*** and ***high***, and then the predictors ***high*** and ***back***, each time making sure that there is no significant change to the model using ANOVA model comparison. No random slopes were added, as no model we found would both converge and improve on the base model. The final model enjoys low collinearity (κ=1.24, VIF$\leq$1.06).

[8] The ***pvals.fnc()*** from the ***languageR*** package (Baayen 2011) was used to calculate p-values.

However, for the polysyllables, the trend was reversed: iambs were preferred with unfaithful plurals more strongly than trochees.

This mismatch is shown in Figure 8, which plots the predictions of the lexicon model of Table 3 against the experimental responses. The predictions and the observations are correlated quite well (Spearman's rank correlation, ρ=.48, $\boldsymbol{p} < .0001$). Yet there are systematic mismatches between the predictions and the results: of the 21 trochees, 14 are less acceptable with alternating plurals than predicted, while 17 of the iambs are more acceptable with alternating plurals than predicted. This contrasts with the predictions for vowels, which are not as systematically skewed.

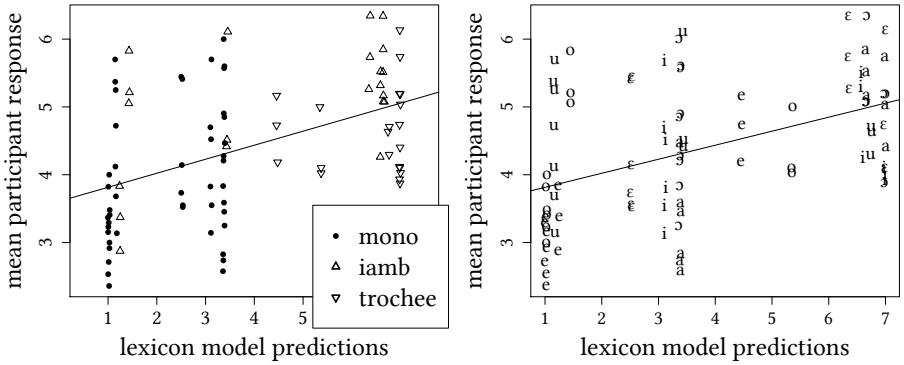

Figure 8: Brazilian Portuguese lexicon predictions (from Table 3) vs. the participants' responses to nonce items, shown by prosodic shape (left) and by stressed vowel (right)

We find it remarkable that speakers go directly against the evidence provided to them by the real words of their language, given the strong case for speakers as excellent statistical learners (e.g. Ernestus & Baayen 2003 and others). Additionally, faithfulness to stressed syllables protects iambs from alternation, but here we do not see the effect of this protection; rather we see its opposite. We speculate that participants notice the alternation more in the stressed syllables of iambs, and are thus more confident that the alternation applies productively in such words. This, however, decouples our learning mechanism from the principles of grammar itself. We hope that an answer to this puzzle will emerge as additional test cases come to light.

5.6 Discussion

The results confirm that Brazilian Portuguese speakers track the plurals of [w]-final nouns in terms of their prosodic shape and the stem's final vowel. In particular, they prefer alternations in polysyllables and following lax vowels. The two effects are highly significant, and independent of each other (showing no significant interaction).

The generalization of these grammatical patterns to nonce items that were presented auditorily confirms that the phonological factors we identified are independent of the effects of history and orthography. Additionally, while Huback (2007) makes a strong case for the effect of token

frequency in plural formation for real words of Portuguese, token frequency can say nothing about nonce items, which are all equally infrequent and equally unfamiliar to the listener.

While speakers' treatment of novel words largely follows the predictions of our lexicon model, we note a discrepancy within the polysyllables: trochees alternate more than iambs in the lexicon, but less than iambs in nonce words. This suggests that faithfulness to stressed positions is not playing a role here, and that the ultimate explanation for this effect is elsewhere.

6 Gradient approaches to word-size effects

We have established that monosyllables are protected from alternations in French and Brazilian Portuguese, and offered an analysis in terms of initial syllable faithfulness. While the details of implementation differ in the two languages, both analyses rely on a discrete, binary distinction between monosyllables and polysyllables.

Turning back to the stimuli in our two experiments, however, we note that in terms of continuous parameters, they offer a great deal more variance in length and word likelihood than we have thus far examined, which could potentially matter. For example, some monosyllables have simple onsets, while others have complex onsets, even going up to triconsonantal clusters in French. Our polysyllables also offer a fairly wide variety of consonants and clusters that differ substantially in length. In addition, we were careful to represent a wide range of phonotactics in our materials, from extremely common singleton onsets up to each language's least frequent clusters.

We generated our stimuli with the idea that a wide range of lengths and word likelihoods would give a fair chance to theories that use this kind of variance to explain size effects in alternations. In particular, we focus on phonetically-grounded explanations that rely on phonetic duration (Barnes 2006; Giavazzi 2010), and lexicon-based explanations that rely on neighborhood density (Ussishkin & Wedel 2009; Stausland Johnsen 2011). While these two approaches differ in their sources of explanation, both rely on informationally rich, continuous, language-specific predictors.

Rich, continuous predictors, however, are not always better than discrete predictors. In particular, if some of the variance in the continuous predictor is misaligned with participants' behavior, then the predictive power of including such factors is diminished. We examine duration-based explanations in §6.1, and lexicon-based explanations in §6.2, estimating the predictive power of each approach and comparing it to our discrete approach.

6.1 Duration-based predictors

The idea that phonological contrast is based on its perceptibility is very well established in the literature (Hayes 1999; Steriade 2001/2008, and many others since). In Optimality Theory, contrast is maintained thanks to faithfulness constraints, which are taken to protect prominent positions

more strongly (Smith 2002, 2010). Prominence can be established in several different ways, e.g. an intense burst for stops, high pitch on a vowel, etc. Duration, however, is taken to be one of the central correlates of prominence.

In particular, Barnes (2006) capitalizes on duration as the driving force of alternations in Turkish and a variety of other languages. Similar notions are pursued in Giavazzi (2010), although she focuses on the role of stress, not on phonological size. The idea in these works is that longer elements are protected from alternation. Since in a short word each individual segment is on average longer than the corresponding segment in a long word, it stands to reason that short words will resist alternations by virtue of having longer and thus more prominent segments.

To compare our monosyllabicity approach with the duration-based approach, we measured the duration of the alternating part and non-alternating part in our materials. For French, we labeled the following points on each stimulus in Praat: the onset of the word, the onset of the word's final vowel (which was followed by an [l] or [j] in the singular), and the offset of the word. This was repeated for the singular and plural of each stimulus, yielding a total of four duration-based predictors. For Portuguese, we again labeled the onset of the word, the onset of the word's last vowel (which was followed by a [w], [ws], or [js]), and the offset of the word. Word-offsets were marked when intensity dropped below 48dB. This was repeated for each of the three forms of each stimulus, giving us a total of six duration-based predictors.

Figure 9 shows the French nonce paradigm [vøzal ∼ vøzo] and the Portuguese nonce paradigm [pɾiˈzɛw ∼ pɾiˈzɛjʃ] (our speaker has ʃ in coda throughout). The segmentation shows the non-alternating parts [vøz] and [pɾiz], and the alternating parts [al, o, ɛw, ɛjʃ].

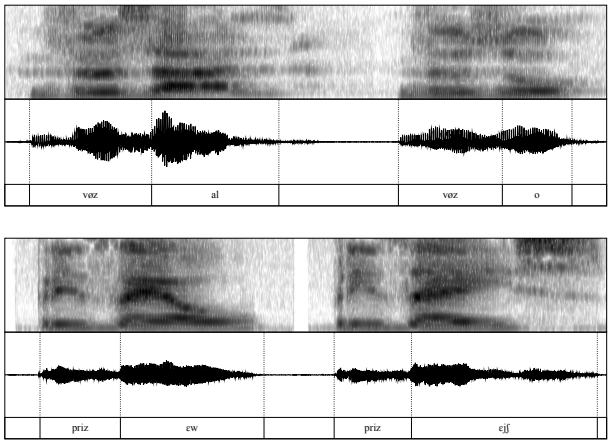

Figure 9: Sample nonce paradigms, segmented into alternating and non-alternating parts

A straightforward interpretation of the phonetically-based approach is that the duration of the potentially alternating portion, i.e. the [al] of French or the [w] of Portuguese, should be negatively correlated with alternation rates. Indeed, the alternating portions in French (al/aj/ɛl/ɛj) are

significantly longer in monosyllables than they are in polysyllables (427ms vs. 369ms, t-test, $t(73.0) = 4.48$, $p < .0001$). The same is true in Portuguese, where the stem's final vowel and following [w] were significantly longer in monosyllables (462ms vs. 384ms, t-test, $t(75.6) = 6.2$, $p < .0001$). Here, we chose not to try and separate out the glide from the preceding vowel, a rather arbitrary task that would be hard for us and equally hard for the speakers of the language. The non-alternating portions similarly distinguish monosyllables and polysyllables; this is shown in Figure 10, where the monosyllables are rather well separated from the polysyllables.

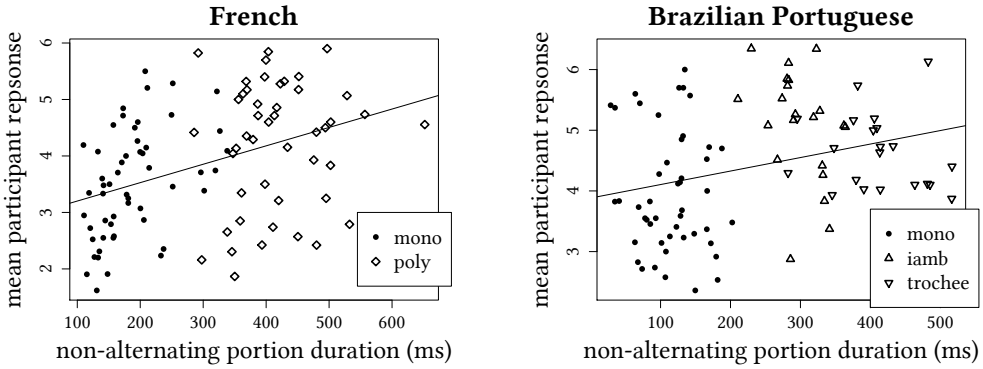

Figure 10: Phonetic duration as a predictor of participant responses

We see in both languages, then, that shorter words have significantly longer alternating portions and longer words have significantly shorter alternating portions. It is clear that the phonetic duration of the alternating portions distinguishes short words from long words; the question is again how well duration predicts participants' responses.

We compared our monosyllabicity criterion and the duration-based approached using nested models. Starting with French, we constructed three models: a model for monosyllabicity (identical to the one in Table 2), a model for phonetic duration, and a superset model that contains the predictors of both. We did the same for Portuguese, based on the model in Table 4. Comparing the superset models to each of their subset models allows us to measure the contribution of each approach. The results, reported in Table 5, is that monosyllabicity makes a significant difference, while the contribution of phonetic duration is too small to reach significance. In other words, monosyllabicity is the best predictor of the participants' responses, leaving the other predictors with no contribution to make beyond it. Monosyllabicity subsumes these other approaches.[9]

To be clear, we are not arguing from a null result. Phonetic duration is an ***excellent*** predictor of our participants' responses, and it is highly significant in the models that have it without monosyllabicity; it is simply the case that its effect is subsumed by monosyllabicity. Once monosyllabicity is present in the model, phonetic duration no longer has anything to add.

[9]We attempted this comparison in many different variations, with various combinations of the duration-based predictors, including with and without interactions, and log transformations. In all cases, the contribution of monosyllabicity was strong and significant, while the contribution of phonetic duration did not reach significance.

		χ^2	df	*p*-value
French	monosyllabicity	16.0	3	<.005
	phonetic duration	13.7	9	>.1
Portuguese	monosyllabicity	30.1	5	<.0001
	phonetic duration	.4	1	>.1

Table 5: Monosyllabicity is the best predictor of participants' responses to nonce words

Another perspective on the question of duration comes from the idea that the source of morphophonological alternations is lenition, as expressed in Hooper [Bybee] (1976) and Pierrehumbert (2001), among others. If lenition is understood primarily as a shortening force, then its primary correlate would be the change in duration that the alternation causes. In French, The [al → o] alternation makes the word shorter; in our materials, the mean difference is 184ms, which is highly significant (paired t-test, $t(98) = 29.0$, $p < .0001$). The shortening, however, is quite similar in monosyllables (173ms) and polysyllables (196ms), a non-significant difference (t-test, $t(87.3) = 1.77$, $p > .05$). If speakers are sensitive to the change in duration that the alternation causes, then duration does not help distinguish monosyllables from polysyllables in French.

The same point holds in Portuguese: here, the choice between the two plurals (i.e. [w] vs. [j]) has little to no bearing on duration. First, the alternating portions are of similar length in the two plurals (paired t-test, $t(88) = .43$, $p > .1$), and second, this holds in both monosyllables and polysyllables (t-test, $t(79.5) = .03$, $p > .1$). This should not come as a surprise, since the difference in backness between [w] and [j] is not usually expected to correlate with duration (unless there is a concomitant difference in height). Escudero et al. (2009) report duration measurements of Brazilian Portuguese vowels, finding that [i] and [u] are of nearly identical length, with the difference (1ms for women, 5ms for men) being only marginally significant. Thus, the Portuguese alternating plural cannot be reasonably described as a lenited plural.

In this sense, the alternations in French and Portuguese contrast with alternations that are quite plausibly described as lenition, as in Turkish laryngeal alternations (Inkelas & Orgun 1995; Inkelas et al. 1997; Petrova et al. 2006; Becker et al. 2011a; a.o.). In Turkish, alternations like [gurup^{h} ∼ guru<u>b</u>-u] 'group' are strongly preferred in polysyllables, which could be ascribed to lenition affecting shorter stops more strongly.

In the absence of lenition, usage-based theories (as in Hooper [Bybee] 1976; Pierrehumbert 2001) rely on analogy to regulate the extension of lexical patterns. As we will see in §6.2 below, this analogy mechanism will have to incorporate a notion of monosyllabicity to achieve optimal empirical coverage. Recall that token frequency is not relevant for predicting the behavior of nonce words, which by definition are all equally new to the participants.

To summarize, we have seen that in both French and Portuguese, alternating portions are longer in monosyllables, and their duration is highly predictive of participants' treatment of nonce words. Monosyllabicity, however, is a significantly better predictor, completely subsuming the predictive

power of phonetic duration in both languages. In addition, we saw that the plural alternations either involve no shortening at all (Portuguese), or the amount of shortening does not correlate with word-length (French), and thus viewing these alternations as cases of lenition will not derive the observed word-size effects.

6.2 Lexicon-based predictors

Lexical neighbors (Luce & Pisoni 1998) are words that are one sound away from a given word in terms of addition, deletion, or substitution, e.g. the neighbors of English [tɹɪp] include [tɪp], [tɹæp], [stɹɪp], etc. Neighborhood density is a measure of the similarity of a given word to other words in the language; more generally, it measures how representative a word is of a given language. Ussishkin & Wedel (2009) survey a range of languages where shorter words are protected from morphophonological alternations, and show a correlation between low alternation rates and high neighborhood density. They propose that both alternations and neighbors reduce the speaker's ability to retrieve lexical items, and thus the negative correlation maintains a balance between the two considerations. In a study of phrase-level phenomena in Norwegian, however, Stausland Johnsen (2011) finds a correlation between dense neighborhoods and more, not less, alternations for real words, but no effect for nonce words. Thus, the connection between neighborhood density and alternation rates remains somewhat mysterious.

Ussishkin & Wedel's (2009) approach has been questioned by Pycha et al. (2007), and later by Becker & Nevins (2009) and Becker et al. (2011a), who show that neighborhood density is a rather weak predictor of morphophonological laryngeal alternations in Turkish. Here, we examine the situation in French and Brazilian Portuguese, but this time with materials that were designed specifically to bring out the utility of lexicon-based predictors, inspecting words from a wide spectrum of likelihoods.

We calculated for each stimulus its neighborhood density and also its log bigram frequency. Bigram frequency is another measure of word likelihood, which could potentially improve the performance of the lexicon-based prediction. In French, the calculations were based on the phonetically transcribed ***Lexique*** electronic dictionary (New et al. 2001, 65,632 items). We calculated these for each singular and plural, making two neighborhood density-based predictors and and two bigram frequency-based predictors for each stimulus. In Brazilian Portuguese, we used the LABEL-LEX electronic dictionary (96,136 items), which we converted to phonetic transcription. In this case, each stimulus comes in three forms (singular, faithful plural, alternating plural), leading to three neighborhood density-based predictors and and three bigram frequency-based predictors.

Longer words have fewer neighbors on average, and indeed our monosyllabic items have significantly more neighbors than the polysyllabic items. This holds true both in French (t-test, $\boldsymbol{t(58.1)} = 8.3$, $\boldsymbol{p} < .0001$) and in Portuguese (t-test, $\boldsymbol{t(65.8)} = 7.03$, $\boldsymbol{p} < .0001$). This is shown in Figure 11, where neighborhood density is negatively correlated with alternations in both

languages. Similarly, monosyllables have higher bigram frequencies, simply by virtue of each bigram frequency being lower than one, and shorter items having fewer bigrams. This holds in French (t-test, $t(84.4) = 11.0$, $p < .0001$) and in Portuguese (t-test, $t(75.4) = 8.5$, $p < .0001$).

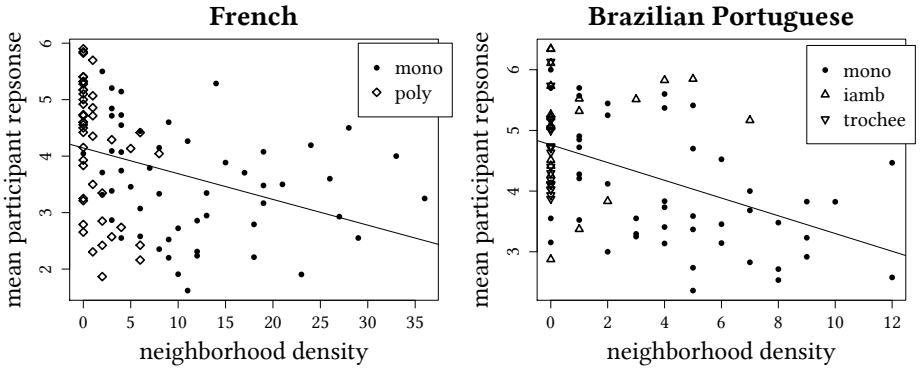

Figure 11: Neighborhood density as a predictor of participant responses

Thus, it is quite obviously the case that these lexicon-based measures distinguish short words from long words; the question is only how well these measures can predict the participants' responses to the stimuli.

We assess the relative contribution of these theories using comparison of nested models (as in §6.1), this time comparing our monosyllabicity criterion with neighborhood density and bigram frequency. Starting with Brazilian Portuguese, we constructed three models: a model for monosyllabicity (identical to the one in Table 4), a model for neighborhood density, and a superset model that contains the predictors of both. The result, reported in Table 6, is that monosyllabicity is the best predictor of the participants' responses, leaving neighborhood density with no contribution to make beyond it.

		χ^2	df	p-value
Portuguese	monosyllabicity	21.1	4	$<.0005$
	neighborhood density	2.6	4	$>.1$

Table 6: Monosyllabicity is the best predictor of participants' responses to nonce words in Brazilian Portuguese

The result stays unchanged with bigram frequency added to the neighborhood density model. Even combined, neighborhood density and bigram frequency are entirely subsumed by monosyllabicity in Portuguese. Again, this is not a null result we present in Table 6: neighborhood density and bigram frequency are each excellent and highly significant predictors of the participants' responses when considered without monosyllabicity.

For French, following the same procedure gives a similar result: the contribution of monosyllabicity is highly significant ($\chi^2(3) = 21.0$, $p < .0005$), while the contribution of neighborhood density is only marginal ($\chi^2(3) = 7.6$, $p > .05$). The result stays unchanged with the addition

of bigram frequency. However, when neighborhood density is log-transformed, it outperforms monosyllabicity, although it does not subsume it. Log-transformed neighborhood density makes a bigger contribution to the superset model (χ^2***(3)*** $= 16.6$, $\boldsymbol{p} < .001$), while the contribution of monosyllabicity is small but still significant (χ^2***(3)*** $= 9.1$, $\boldsymbol{p} < .05$).

This result leaves monosyllabicity as a necessary part of the picture for French, but it also adds log-transformed neighborhood density as a strong predictor of participants' responses beyond monosyllabicity.

There is a way, however, in which neighborhood density is unlike monosyllabicity and phonetic duration. By its very definition, neighborhood density is computed based on a lexicon, and it is sensitive to the size of the lexicon: the more lexical items a speaker has, the more neighbors some words will have, and neighborhood density will become informationally richer. We sought to measure this dependence of neighborhood density on lexicon size, and see what kind of lexicon is needed to make neighborhood density a better predictor than monosyllabicity.

We randomly sampled from Lexique 1,000 times, taking 50 samples from each of 20 different lexicon sizes, ranging from ∼6,500 (10% of the whole list) to ∼59,000 (90% of the list). For each of these one thousand samples, we fitted three models: one based on monosyllabicity (like the one in Table 2), one based on log transformed neighborhood density, and a superset model that includes both. We measured the contribution of the two nested models, getting two χ^2 scores. Figure 12 shows the difference between the two scores for each of the 1,000 samples. Both χ^2 scores are at 3 degrees of freedom, and thus can be compared on this abstract, unitless scale. The raw results are available at [link redacted].

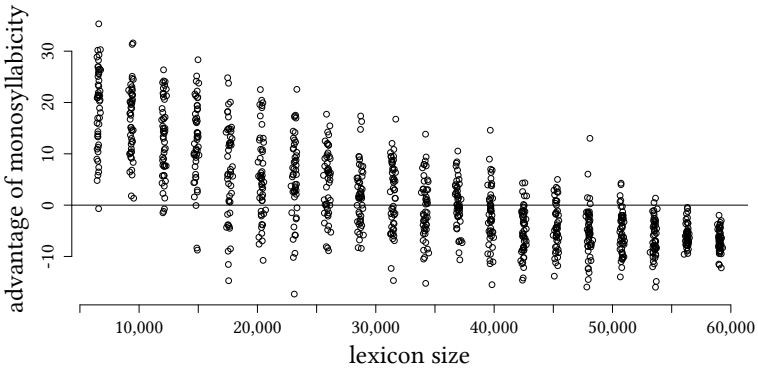

Figure 12: Monosyllabicity vs. neighborhood density in 1,000 random samples from the French lexicon. Monosyllabicity is the better predictor for lexicons of up to 30,000 items; neighborhood density is the better predictor for lexicons of 39,000 items and above.

With lexicons that have 39,000 items or more, log transformed neighborhood density is a significantly better predictor of participants' responses than monosyllabicity (paired t-test at ∼39,000, ***t(49)*** $= -2.20$, $\boldsymbol{p} < .05$), with the advantage of neighborhood density being stronger for bigger lexicons. With lexicons that have 30,000 items or less, however, monosyllabicity is

significantly better than neighborhood density, even when it is log transformed (paired t-test at ∼29,000, $t(49) = 2.43$, $p < .05$), with the advantage of monosyllabicity being stronger for smaller lexicons. Recall that without the log transform, monosyllabicity is the better predictor for lexicons of any size.

A reduction in lexicon size is sure to reduce the informational power of neighborhood density; simply computing neighborhood density on a smaller lexicon is a trivial way to reduce its effect. Our simulation is important, however, in showing that neighborhood density only amasses enough power to overcome monosyllabicity from a lexicon that is only marginally plausible in size. It is an entirely reasonable assumption that most of our participants did not have vocabulary size above the threshold at which neighborhood density kicks in as more useful; a point that can be empirically tested in future work. According to Seidenberg & McClelland (1990), the average size of an adult lexicon is 30,000 words.

Due to its reliance on the lexicon, neighborhood density only emerges as useful when speakers have a particularly large vocabulary size, unlike monosyllabicity and duration-based measures, which do not require a comparison to a larger set of items. This difference is particularly important for children who are acquiring the language: monosyllabicity and duration-based measures are available immediately, and monosyllabicity is useful immediately, while neighborhood density only becomes useful much later, when the lexicon is large and rather adult-like. In the crucial early stages, the advantage of monosyllabicity is strongest. If indeed speakers switch to rely on neighborhood density at some point, they will also need a mechanism that points them in this direction. The exact details of the acquisition mechanism aside, the prediction is made that speakers with smaller lexicons (children or adults) will rely on monosyllabicity, while speakers with larger lexicons may additionally rely on neighborhood density.

To summarize, the situation is simple in Portuguese: monosyllabicity subsumes the predictive power of neighborhood density, based on 96,000 items, even when log transformed and when bigram frequency is added. In French, the same obtaines with raw neighborhood density, but log-transformed neighborhood density makes a strongly significant contribution to the model, leaving monosyllabicity with a smaller (yet siginificant) contribution. This result depends on a French lexicon of at least 39,000 lexical items; monosyllabicity is the better predictor for lexicons of at most 30,000 lexical items. The reliance of neighborhood density on a substantial lexicon suggests directions for work with children and adults with small lexicons, and require a mechanism for noticing the increased usefulness of neighborhood density as the speaker's lexicon develops.

7 General discussion and conclusions

In this paper, we offered two studies of plural alternations, in French and Portuguese. We found that in both languages, monosyllables are protected from alternations, and we analyzed the effect in terms of initial syllable faithfulness (Trubetzkoy 1939; Steriade 1994; Beckman 1997, 1998; Casali

1998; Barnes 2006; Jesney 2009; Becker 2009; Becker et al. 2011a). Fundamental to the proposal was the claim that syllable count is a salient property of words, and readily exploited by learners.

Our approach is based on the traditional generative approach, which seeks to account for phonological patterns using discrete formal mechanisms. We make a modest departure from tradition, augmenting the discrete approach with the ability to track lexical statistics, using constraint cloning (Pater 2006, 2009; Coetzee 2008; Becker 2009; Becker et al. 2011a). The gradient behavior of individual lexical items is derived from a discrete mechanism, just like a continuous binomial distribution is derived from the discrete heads-or-tails nature of a coin.

The underlying mechanism of initial syllable faithfulness is potentially universal, possibly even innate, even though the evidence from our two test-cases does not speak to that directly. It is possible that the positional faithfulness mechanism can be induced simply by giving the speaker the notion of monosyllabicity, or maybe even just the notion of syllable; see an effort in this general direction in Daland et al. (2011). Until such an induction mechanism is developed, however, we assume that speakers bring the protection of initial syllables to the table, regardless of language-specific input.

We compare our approach with two approaches that rely on gradient and language-specific measurements, one relying on phonetic duration (Barnes 2006; Giavazzi 2010), and one relying on neighborhood density (Ussishkin & Wedel 2009). These approaches point in the right direction, with word-size being correlated with segmental duration and neighborhood density. These informationally-rich, continuous approaches, however, do not predict human responses as accurately and as strongly as our discrete approach. Without appealing to the notion of monosyllabicity, the gradience and richness of these approaches is to their detriment, as they offer smaller empirical coverage. Richness is not guaranteed to improve accuracy.

While the dominance of monosyllabicity is absolute in Portuguese, the picture is somewhat more complicated in French. Here, log-transformed neighborhood density has stronger predictive power when given a lexicon of over ∼39,000 items. This reliance on a large lexicon makes the prediction that speakers will respond differently to nonce words depending on the size of their lexicon, with younger speakers relying on immediately available measures such as monosyllabicity, and with older speakers increasingly relying on the lexicon as it grows.

In addition to its reliance on the lexicon, neighborhood density also differs from phonetic duration and positional faithfulness in that its predictive power here is purely due to the size-based nature of the alternation. When alternations target longer items, as they do in French and Portuguese, they correlate with neighborhood density, because neighborhood density is universally correlated with word-length. In other positionally-governed alternations, such as those sensitive to stress, neighborhood density is not expected to predict alternations at all. Thus, the success of neighborhood density is limited to a specific kind of positional effect, and has no promise as a more general approach to positional phenomena. In this sense, the phonetically based approach is more promising as a theory of positional faithfulness than the lexicon-based approach,

even if it does not achieve the same predictive power as our discrete approach.

The model presented in this paper, couched in terms of Optimality Theory, makes further predictions beyond the mere deployment of a monosyllabicity criterion. The analysis derives plural forms via comparison between candidates, which naturally identifies the locus of alternation by canceling shared violations (see Prince 2002). The lexical statistics are only tracked using constraints that distinguish between winners and losers, leaving other elements of the stem unable to interfere. For example, the difference between simple and complex onsets, which is discrete and binary, is not expected to be a good predictor of alternations, even though it correlates with differences in duration and word-likelihood. Indeed, onset complexity makes no improvement in prediction in either language, even among the monosyllables only (French: $\chi^2(1) = 3.19$, $p > .05$, Brazilian Portuguese: $\chi^2(1) = .06$, $p > .1$)

In conclusion, our use of initial syllable faithfulness allows us to understand the trends in the plurals of French and Portuguese in terms of a larger cross-linguistic pattern of regulation of strong positions. In the study of phonotactics, initial syllables show a wider range of contrasts in a variety of languages (Steriade 1994, Beckman 1997, 1998, a.o.), and thus pattern with other strong positions, such as stressed syllables, roots, and nouns (Smith 2002, 2010). Strong positions are also subject to augmentation effects; this should be observable in initial syllables, and the prediction can be tested in future work. Thus, our approach is not only numerically more accurate in the two cases we explored here, it also accords with a wider range of known facts, and makes predictions about the range of possible phenomena. The micro-patterns of initial syllables' treatment in exceptional plurals in these two languages reflect typological macro-patterns, as well as salient cognitive universals related to word-edges in phonotactic learning (Endress & Mehler 2010). Despite the rich swaths of multidimensional data that computers can store and access, human learners of alternations seem to adhere to the well-worn adage "Keep it Simple".

Appendix A French nonce words

Items with the average response they were given on the scale of 1–7, where 1 represents a faithful plural (identical to the singular), and 7 represents an alternating plural.
Alternating plurals end in [ø] for [ɛj]-final items and [o] otherwise.

[al]		[aj]	
dal	3.60	daj	n/a
fal	4.00	faj	2.93
fʁal	4.26	fʁaj	5.29
gnal	5.21	gnaj	5.50
kʁal	3.71	kʁaj	3.50
pnal	4.55	pnaj	4.71
psal	4.84	psaj	4.07
pʁal	3.79	pʁaj	3.33
sal	3.25	saj	4.50
skʁal	3.74	skʁaj	4.09
snal	3.46	snaj	4.73
spʁal	3.71	spʁaj	3.38
stʁal	4.44	stʁaj	5.14
sval	3.07	svaj	2.87
tʁal	3.89	tʁaj	3.17
vʁal	4.15	vʁaj	4.60
zal	3.35	zaj	2.86
ʒal	4.19	ʒaj	4.08
deʁnal	4.92	deʁnaj	4.86
guval	5.32	guvaj	5.00
gzovʁal	4.60	gzovʁaj	5.90
istʁal	4.29	istʁaj	5.85
moʁnal	5.70	moʁnaj	4.71
oskʁal	4.71	oskʁaj	4.60
oʁsodal	5.41	oʁsodaj	4.15
peʁetal	5.28	peʁetaj	4.56
pʁykal	5.17	pʁykaj	5.10
segal	4.14	segaj	4.35
skʁənal	5.32	skʁənaj	4.50
smyʁkal	4.74	smyʁkaj	3.83
smønal	5.17	smønaj	3.93
vøzal	5.40	vøzaj	5.82
ʒistal	5.07	ʒistaj	4.42

[ɛl]		[ɛj]	
dɛl	2.21	dɛj	2.72
dʁɛl	2.24	dʁɛj	2.35
knɛl	2.55	knɛj	4.05
kʁɛl	2.79	kʁɛj	2.31
nɛl	1.90	nɛj	2.95
psɛl	2.58	psɛj	3.32
pʁɛl	2.52	pʁɛj	1.91
ʁɛl	2.55	ʁɛj	3.48
zɛl	1.62	zɛj	2.20
fanɛl	4.42	fanɛj	2.30
faʁasɛl	2.79	faʁasɛj	3.25
gʁasɛl	2.42	gʁasɛj	2.57
paʁtɛl	4.04	paʁtɛj	1.87
ʁafɛl	2.74	ʁafɛj	3.50
ʁomɛl	2.16	ʁomɛj	2.85
ʃotɛl	2.42	ʃotɛj	3.21
ʒyvɛl	3.35	ʒyvɛj	2.65

Appendix B Brazilian Portuguese nonce words

Items with the mean response per item on the 1–7 scale, where 1 is a back plural (e.g. [ˈbɾaw ∼ ˈbɾaws]) and 7 is a front plural (e.g. [ˈbɾaw ∼ ˈbɾajs]).

mono	
ˈbɾaw	3.59
ˈdaw	4.47
ˈfaw	2.58
ˈfɾaw	2.74
ˈpɾaw	2.83
ˈtɾaw	3.45
ˈbɛw	3.73
ˈbɾɛw	4.14
ˈkɛw	5.44
ˈpɾɛw	3.55
ˈpɛw	5.41
ˈtɾɛw	3.52
ˈbɔw	3.25
ˈbɾɔw	6.00
ˈdɾɔw	4.21
ˈkɔw	5.60
ˈpɔw	3.83
ˈtɾɔw	4.28
ˈvɔw	4.90
ˈgɾɔw	4.85
ˈʒɔw	5.57
ˈfew	2.53
ˈfɾew	2.36
ˈkew	2.71

mono (cont.)	
ˈnew	2.92
ˈprew	3.15
ˈgɾew	3.29
ˈʒew	3.37
ˈbɾow	3.41
ˈfow	3.48
ˈkɾow	3.00
ˈmow	4.00
ˈnow	3.23
ˈtow	3.82
ˈdɾiw	5.70
ˈkiw	3.83
ˈniw	4.52
ˈpɾiw	3.55
ˈtɾiw	3.14
ˈziw	4.70
ˈbɾuw	5.70
ˈdɾuw	4.72
ˈfuw	3.14
ˈmuw	3.68
ˈpɾuw	5.25
ˈtuw	5.37
ˈvuw	4.12

trochee	
ˈtahtew	5.74
ˈʃastow	4.40
ˈʒaɲew	5.04
ˈfɛskow	4.10
ˈpɛspow	4.74
ˈvɛhpew	6.13
ˈkɔzew	5.20
ˈtɔmew	5.19
ˈzɔfow	3.87
ˈhefew	5.17
ˈkeskew	4.73
ˈtehtow	4.18
ˈhosew	5.00
ˈkospow	4.02
ˈsohkew	4.10
ˈsikow	4.12
ˈʒiɲew	4.03
ˈʧibow	3.93
ˈduhnow	4.64
ˈhuntew	4.71
ˈtumow	4.30

iamb	
biˈɲaw	5.85
kuˈtaw	5.51
maˈhaw	5.17
buˈtɛw	5.73
paˈmɛw	6.34
pɾiˈzɛw	5.26
kuˈpɔw	5.08
niˈsɔw	6.34
gahˈtɔw	5.08
haˈsew	3.83
suˈpew	3.37
ʤiˈmew	2.88
vaˈɾow	5.83
ziˈbow	5.21
zuˈgow	5.05
duˈʒiw	4.26
ʃaˈmiw	5.32
ʧiˈviw	5.52
tuˈzuw	6.11
ʒaˈnuw	4.42
ʧiˈɾuw	4.51

References

Albright, Adam & Bruce Hayes (2003). Rules vs. Analogy in English past tenses: a computational/experimental study. ***Cognition* 90**. 119–161.

Baayen, R. H. (2011). ***languageR: Data sets and functions with "Analyzing Linguistic Data: A practical introduction to statistics"***.. R package.

Barnes, Jonathan (2006). ***Strength and Weakness at the Interface: Positional Neutralization in Phonetics and Phonology***. Berlin/New York: Mouton de Gruyter.

Bates, Douglas, Martin Maechler & Ben Bolker (2011). ***lme4: Linear mixed-effects models using S4 classes***.

Becker, Michael (2009). ***Phonological Trends in the Lexicon: The Role of Constraints***. Ph.D. dissertation, University of Massachusetts Amherst.

Becker, Michael, Nihan Ketrez & Andrew Nevins (2011a). The surfeit of the stimulus: Analytic biases filter lexical statistics in Turkish laryngeal alternations. ***Language* 87**. 84–125.

Becker, Michael & Jonathan Levine (2010). Experigen - an online experiment platform. Available at https://github.com/tlozoot/experigen.

Becker, Michael & Andrew Nevins (2009). Initial-syllable faithfulness as the best model of word-size effects in alternations. Talk given at NELS 40.

Becker, Michael, Andrew Nevins & Jonathan Levine (2011b). Asymmetries in generalizing alternations to and from initial syllables. Ms. available at http://becker.phonologist.org/projects/english/.

Beckman, Jill (1997). Positional faithfulness, positional neutralisation and Shona vowel harmony. ***Phonology* 14**. 1–46.

Beckman, Jill (1998). ***Positional Faithfulness***. Ph.D. dissertation, University of Massachusetts Amherst, Amherst, MA.

Benua, Laura (1997). ***Transderivational Identity: Phonological Relations Between Words***. Ph.D. dissertation, University of Massachusetts Amherst.

Boersma, Paul & David Weenink (2011). Praat: Doing phonetics by computer. [Computer program].

Bybee, Joan (1995). Regular morphology and the lexicon. ***Language and Cognitive Processes* 10**. 425–455.

Bybee, Joan (2001). ***Phonology and language use***. Cambridge University Press.

Casali, Roderic (1998). ***Resolving Hiatus***. Garland, New York.

Coetzee, Andries W. (2008). Grammaticality and ungrammaticality in phonology. ***Language* 84**. 218–257.

Côté, Marie-Hélène (2011). French liaison. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.) ***The Blackwell Companion to Phonology***, Blackwell.

Daland, Robert, Bruce Hayes, Marc Garellek, James White, Andrea Davis & Ingrid Normann (2011). Explaining sonority projection effects. ***Phonology* 28**. 197–234.

Endress, Ansgar D. & Jacques Mehler (2010). Perceptual constraints in phonotactic learning. ***Journal of Experimental Psychology: Human Perception and Performance* 36**. 235–250.

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

Escudero, Paola, Paul Boersma, Andreia Rauber & Ricardo Bion (2009). A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. ***Journal of the Acoustical Society of America* 126**. 1379–1393.

Gelman, Andrew & Jennifer Hill (2007). ***Data Analysis Using Regression and Multilevel/Hierarchical Models***. Cambridge University Press.

Giavazzi, Maria (2010). ***The phonetics of metrical prominence and its consequences for segmental phonology***. Ph.D. dissertation, MIT.

Gomes, Christina Abreu & Carolina Gonçalves Manoel (2010). Flexão de número na gramática de criança e na gramática do adulto. ***Veredas*** . 122–134.

Hayes, Bruce (1999). Phonetically-driven phonology: The role of optimality theory and inductive grounding. In M. Darnell, E. Moravscik, M. Noonan, F. Newmeyer & K. Wheatly (eds.) ***Functionalism and Formalism in Linguistics, Volume I: General Papers***, John Benjamins. 243–285.

Hayes, Bruce & Zsuzsa Londe (2006). Stochastic phonological knowledge: The case of Hungarian vowel harmony. ***Phonology* 23**. 59–104.

Hooper [Bybee], Joan B. (1976). Word frequency in lexical diffusion and the source of morphophonological change. In W. Christie (ed.) ***Current Progress in Historical Linguistics***, NorthHolland. 96–105.

Huback, Ana Paula da Silva (2007). ***Efeitos de freqüência nas representações mentais***. Ph.D. dissertation, Universidade Federal de Minas Gerais, Belo Horizonte.

Inkelas, Sharon & Cemil Orhan Orgun (1995). Level ordering and economy in the lexical phonology of Turkish. ***Language* 71**. 763–793.

Inkelas, Sharon, Cemil Orhan Orgun & Cheryl Zoll (1997). The implications of lexical exceptions for the nature of the grammar. In Iggy Roca (ed.) ***Derivations and Constraints in Phonology***, Oxford: Clarendon. 393–418.

Jesney, Karen (2009). Positional faithfulness, non-locality, and the harmonic serialism solution. Proceedings of the 39th Meeting of the North East Linguistic Society.

Luce, Paul A. & D.B. Pisoni (1998). Recognizing spoken words: The Neighborhood Activation Model. ***Ear and Hearing* 19**. 1–36.

New, Boris, Christophe Pallier, Ludovic Ferrand & Rafael Matos (2001). Une base de données lexicales du français contemporain sur internet: Lexique. ***L'Année Psychologique* 101**. 447–462.

Pater, Joe (1999). Austronesian nasal substitution and other *NÇ effects. In René Kager, Harry van der Hulst & Wim Zonneveld (eds.) ***The Prosody–Morphology Interface***, Cambridge: Cambridge University Press. ROA 160-1196.

Pater, Joe (2006). The locus of exceptionality: Morpheme-specific phonology as constraint indexation. In Leah Bateman & Adam Werle (eds.) ***UMOP: Papers in Optimality Theory III***, Amherst, MA: GLSA. 1–36.

Pater, Joe (2009). Morpheme-specific phonology: Constraint indexation and inconsistency resolution. In Steve Parker (ed.) ***Phonological Argumentation: Essays on Evidence and Motivation***, Equinox. 1–33.

Petrova, Olga, Rosemary Plapp, Catherine Ringen & Szilárd Szentgyörgyi (2006). Voice and aspiration: Evidence from Russian, Hungarian, German, Swedish, and Turkish. ***The Linguistic Review* 23**. 1–35.

Pierrehumbert, Janet (2001). Exemplar dynamics: Word frequency, lenition and contrast. In Joan Bybee & Paul Hopper (eds.) ***Frequency and the Emergence of Linguistic Structure***, John Benjamins. 135–157.

Pope, Mildred Katharine (1952). ***From Latin to Modern French***. London: Butler & Tanner.

Prince, Alan (2002). Arguing optimality. ROA 562-1102.

Prince, Alan & Paul Smolensky (1993/2004). ***Optimality Theory: Constraint Interaction in Generative Grammar***. Oxford: Blackwell. [ROA-537].

Pycha, Anne, Sharon Inkelas & Ronald Sprouse (2007). Morphophonemics and the Lexicon: A Case Study from Turkish. In M. J. Solé, P. Beddor & M. Ohala (eds.) ***Experimental Approaches to Phonology***, Oxford University Press. 369–385.

R Development Core Team (2011). R: A language and environment for statistical computing. Vienna, Austria. (http://www.R-project.org).

Seidenberg, Mark S. & James L. McClelland (1990). More words but still no lexicon: Reply to besner et al. (1990). ***Psychological Review* 97**. 447–452.

Smith, Jennifer L. (2002). ***Phonological Augmentation in Prominent Positions***. Ph.D. dissertation, University of Massachusetts Amherst.

Smith, Jennifer L. (2010). Category-specific effects. In Marc van Oostendorp, Colin J. Ewen,

Elizabeth Hume & Keren Rice (eds.) ***The Blackwell Companion to Phonology***, Blackwell.
Stausland Johnsen, Sverre (2011). ***The origin of variation in Norwegian retroflexion***. Ph.D. dissertation, Harvard University.
Steriade, Donca (1994). Positional neutralization. Ms., UCLA.
Steriade, Donca (2001/2008). The phonology of perceptibility effects: The p-map and its consequences for constraint organization. In Kristin Hanson & Sharon Inkelas (eds.) ***The Nature of the Word: Studies in Honor of Paul Kiparsky***, Cambridge, MA: MIT Press. 151–180.
Trubetzkoy, Nikolai (1939). ***Grundzüge der Phonologie***. Göttingen: Vandenhoeck & Ruprecht.
Ussishkin, Adam & Andrew Wedel (2009). Lexical access, effective contrast and patterns in the lexicon. In Paul Boersma & Silke Hamann (eds.) ***Perception in Phonology***, Mouton de Gruyter. 267–292.
Wolf, Matthew (2007). For an autosegmental theory of mutation. In Leah Bateman, Michael O'Keefe, Ehren Reilly & Adam Werle (eds.) ***UMOP 32: Papers in Optimality Theory III***, Amherst, MA: GLSA. 315–404.
Zuraw, Kie (2000). ***Patterned Exceptions in Phonology***. Ph.D. dissertation, UCLA.