

# Generalization of French and Portuguese Plural Alternations and Initial Syllable Protection\*

Michael Becker  
michael.becker@stonybrook.edu

Lauren Eby Clemens  
lclemens@albany.edu

Andrew Nevins  
a.nevins@ucl.ac.uk

Version of February 17, 2016.

## *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 right-edge alternations more than polysyllabic items are, an asymmetry we attribute to privileged protection of initial syllables. We implement the analyses of the two languages using constraint-based grammars that take trends learned across the lexicon and predict the treatment of nonce words. Five large-scale nonce word tasks confirm the productivity of the trend in both languages.

Keywords: morphophonological alternations, irregular plurals, initial syllable faithfulness, neighborhood density, durational asymmetries.

## 1 Initial Syllable Faithfulness

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, specifically implemented in terms of *initial syllable faithfulness* as a means of protecting the former but not the latter from alternations. While 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), the degree of granularity of the mechanisms that regulate grammatical generalization to novel forms remains an open question.

One family of hypotheses is that speakers use discrete formal primitives such as syllable count or feature-markedness constraints, as in frameworks such as Optimality Theory (Prince & Smolensky 1993/2004), and that constraint rankings may 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, 2008b; Coetzee 2008; Becker 2009; Becker, Ketrez & Nevins 2011) and the USELISTED approach (Zuraw 2000; Hayes & Londe 2006; Becker, Nevins & Levine 2012). While these authors differ in their commitment to the universality of the constraint set and its naturalness, it is often implicitly assumed that universality and naturalness are guiding principles, if not ironclad rules. In the present paper, we demonstrate that in both French and Brazilian

---

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

Portuguese, monosyllables are protected from plural alternations, whereas polysyllables are impacted more strongly. We attribute this asymmetric 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. 2011). We contrast our discrete analysis with hypotheses that use fine-grained, gradient approaches that refer to phonetic duration, neighborhood density, and similar measures.

The case studies at hand are two processes affecting plural formation in what historically were lateral-final nouns in French and Portuguese. Both French and Portuguese have undergone a series of diachronic changes affecting the realization of final laterals in both singular and plural forms, and the telescoping of some of these changes has led to an irregular set of alternations. The extension of these alternations to novel words is limited in French but as we demonstrate, can be elicited in experimental settings, whereas in Portuguese it is strongly productive outside the laboratory as well. By hypothesis, the generalization of the morphophonological process of plural formation to novel words has fallen along grammatically-defined contours, in which number of syllables and vowel quality play a role within the context of discretely-defined constraints on the process.

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 discuss alternative explanations to the initial syllable effect, and §7 concludes.

## 2 A study of alternations in the French lexicon

The point of departure for the first set of studies is based on French plural morphology as measured in the lexicon, where we demonstrate that in the real words of French, alternations impact polysyllables more than monosyllables. This size-based asymmetry is a novel observation, to our knowledge. The trend is statistically significant, and we offer an analysis that captures the trend grammatically by appealing to initial syllable faithfulness.

### 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 shown in (1), the same [al] → [o] alternation is also fairly common for [al]-final adjectives, e.g. [lwajal ~ lwajo] 'loyal'. The adjectives show the same kind of lexically-specific variation that nouns do, e.g. [nazal ~ nazo] 'nasal' vs. [naval ~ naval] 'naval'.

(1) Unfaithful, alternating French plurals<sup>1</sup>

	alternation	singular	plural	
a.	[al] → [o]	mal bokal зувнал	mo boko зувно	‘evil’ ‘jar’ ‘newspaper’
b.	[aj] → [o]	baj травaj supivaj	bo травo supivo	‘lease’ ‘work’ ‘basement window’

These alternations contrast with near-minimal pairs of nouns and adjectives that keep the same form in the singular and in the plural, as in (2). In addition, some variable nouns and adjectives can take either kind of plural, e.g. [val] ‘valley’, [bɛɾkaj] ‘home’. The alternations do not always impact homophonous nouns and adjectives equally in our data, e.g. the adjective [final ~ fino] ‘final’ always alternates, but the corresponding noun is variable [final ~ final / fino] ‘end’.

## (2) Faithful, non-alternating French plurals

	singular	plural	
a.	bal ʃakal karnaval	bal ʃakal karnaval	‘ball’ ‘jackal’ ‘carnival’
b.	maj detaj evõtaj	maj detaj evõtaj	‘hammer’ ‘detail’ ‘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 ~ maʎs] ‘evil, sg./pl.’ of Gallo-Roman (7<sup>th</sup> century) turned to [mal ~ maʎ] and then monophthongized to [mal ~ moʃ] by the end of Early Old French (12<sup>th</sup> century). Soon after, the plural suffix was lost in all but liaison contexts, producing the modern [mal ~ mo]. The [j]-final nouns in (1b-2b) had a palatal lateral that velarized preconsonantly, and thus followed a similar path, from e.g. [baʎ ~ baʎs] ‘lease’ to [baʎ ~ baʎs] to [baʎ ~ baʎs] to [baʎ ~ bos] and with the loss of the plural suffix, [baʎ ~ bo]. The simplification of the palatal lateral in the 18<sup>th</sup> century produced the modern [baj ~ bo] (Pope 1952; see also Bennett 1997).

Up until Early Old French (12<sup>th</sup> 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 alternation 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 (16<sup>th</sup> century)

<sup>1</sup>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⟩ ‘newspaper’, ⟨soudpirail ~ soudpiraux⟩ ‘basement window’.

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.

In concluding this overview 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. Finally, we mention that [l/j] alternations affect five adjectives that acquire a final [l] or [j] in the feminine and before vowel-initial singular nouns (i.e., liaison contexts; 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 adjectives remain unchanged in the plural.

## 2.2 Trends in the French irregular plural

To assess the distribution of the alternations among the real words of French, we extracted all 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] ‘sky’, we focused on the 672 [al/aj]-final items. A native speaker of French reviewed all of the monosyllabic [al/aj] items and a random sample of the polysyllabic ones, and classified all of the items for which they knew the plural as alternating, non-alternating, or variable. Items without a known plural were discarded, leaving 16 monosyllables and 102 polysyllables, for a total of 118 masculine [al/aj]-final nouns and adjectives (listed in Appendix A). Since our focus is on the difference between monosyllables and polysyllables, there was little benefit in increasing the number of polysyllables in our sample. The results are shown in Figure 1, where it is evident that the alternations impact polysyllabic [al/aj]-final items more than monosyllabic ones. The dataset is available at <http://becker.phonologist.org/projects/FrenchPortuguese/>.

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 2013). The dependent variable was a binary distinction between alternating and non-alternating plurals. Variable words, which can take either kind of plural (10 of the 118 items), were counted as non-alternators; we also tried counting them as alternators, which led to nearly identical results (cf. our MaxEnt analysis below, which allows the variability to be modeled). The predictors used were final *consonant*, a binary factor that contrasted [l] and [j]; *monosyllabic*, a binary factor that contrasted monosyllables with polysyllables; *token frequency* (taken from *Lexique* and log-transformed); and *neighborhood density* (calculated as explained in §6 below). A model with *consonant* and *token frequency* was significantly improved by addition of *monosyllabic* (ANOVA likelihood test,  $\chi^2(1) = 10.6$ ,  $p < .005$ ). None of the interactions made a

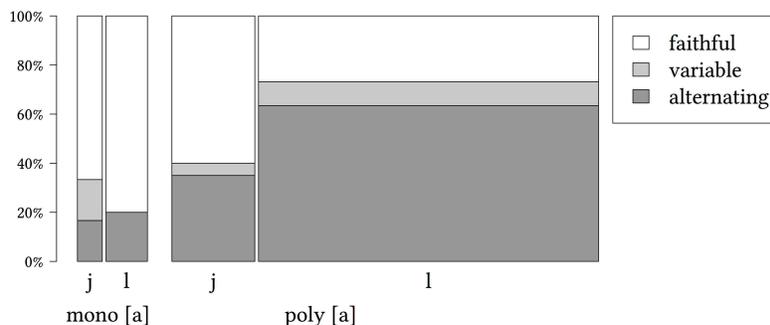


Figure 1: Mosaic plot of French [al/aj]-final singular nouns and adjectives ( $n=118$ ), by monosyllabicity and final consonant. Monosyllables are less likely to take the alternating [o] plural.

	$\beta$	$SE(\beta)$	$z$	$p$ -value
(Intercept)	.07	.20	.32	
<i>consonant=[l]</i>	.46	.21	2.19	<.05
<i>token frequency</i>	.50	.22	2.31	<.05
<i>monosyllabic</i>	-.73	.25	-2.89	<.005

Table 1: Regression model for the French lexicon. Positive  $\beta$  reflects more alternation.

significant improvement, nor did *neighborhood density*.<sup>2</sup> The final model, in Table 1, enjoys low collinearity ( $\kappa=1.6$ ).

The model in Table 1 shows that final [l] is conducive to significantly more alternations than final [j], frequent items are conducive to significantly more alternations than infrequent items, and polysyllables are conducive to significantly more alternations than monosyllables. In other words, monosyllables are protected from alternations, and this effect is predicted to apply to novel items. Neighborhood density, however, is not a significant predictor of alternations in the lexicon. This statistical model, based on the real words of French, predicts that a novel word of French will be more likely to alternate if it ends in [al] rather than [aj]. Even more strongly, it predicts that a nonce word is more likely to alternate if it is polysyllabic. These predictions are tested with three experiments in §3 below, where we will see that the weaker effect of the final consonant is not extended to nonce words, but the predicted difference between monosyllables and polysyllables is borne out strongly and clearly. Token frequency, while relevant to the existing words of the language, is irrelevant to novel words, which are all equally new to speakers.

### 2.3 Initial syllable faithfulness protects monosyllables

We have seen that the [al/aj  $\rightarrow$  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

<sup>2</sup>To reduce the collinearity in the model, all predictors were  $z$ -transformed, giving rise to the following ranges: *monosyllabic*:  $-.39 + 2.51$ , *consonant*:  $-1.87 + .53$ , and *token frequency*:  $-1.98 + 2.64$ . For further detail on collinearity, and how it must be kept low to give rise to trustworthy regression models, see e.g. Baayen (2008; §6.2.2). Applying a  $z$ -transformation (and centering more generally) is one of several ways to reduce collinearity; for further details and practical tutorials, see Jaeger (2013). Concerns about collinearity also informed the decision to exclude *part of speech* from the model, since all but one [j]-final items are nouns, and thus *part of speech* is highly correlated with *consonant*; however, we return to this issue in Section 3.4.

grammar, as evidenced by the fact that they extend it to novel items. We propose that *initial syllable faithfulness* is responsible for the effect (Trubetzkoy 1939; Steriade 1994; Beckman 1997, 1998; Casali 1998; Barnes 2006; Jesney 2009; Becker 2009; Becker et al. 2011, 2012). 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 the disappearance of the plural suffix (represented only in the orthography) removed the environment that conditions the change. 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 does apply 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’). Thus, all the continuant sonorants that are allowed word-finally in nouns participate in the alternation (unless one treats the fricative [ʁ] as a liquid due to its distribution in clusters).

We present our synchronic analysis of the French plural in three parts. First, §2.3.1 uses the autosegmental theory of mutation (Wolf 2007) to cause the [al/aj→o] change. Then in §2.3.2, the USELISTED approach allows established lexical items to maintain a stable plural, while simultaneously allowing lexical statistics to project probabilistically to novel items. Finally, in §2.3.3, *initial syllable faithfulness* captures the resistance of monosyllabic items to the [al/aj→o] change.

### 2.3.1 Alternation caused by a floating [+round]

We propose that the plural affix is a floating [+round] feature that docks on the stem’s final [a], following Wolf’s (2007) autosegmental theory of mutation. The feature [−low] is not needed in the underlying representation of this affix, as markedness constraints ban a [+low, +round, +back] vowel in French. The floating affix is required to dock by MAX(float), which 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(round) 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- $\sigma_1$ (round)

/bal + [+round]/	IDENT- $\sigma_1$ (round)	MAX(float)	IDENT(round)
a.  bal		*	
b. bo	*!		*

- (4) [bo.kal ~ bo.ko] ‘jar’ allows the floating /[+round]/ to dock

/bokal + [+round]/	IDENT- $\sigma_1$ (round)	MAX(float)	IDENT(round)
a. bo.kal		*!	
b.  bo.ko			*

The winner [bo.ko] in (4) needs to be more optimal than other, more faithful candidates, such as [bo.kol]. Following Wolf’s (2007:§4.1) analysis of DhoLuo, where alignment constraints force deletion, we use ALIGN-R(affix, stem), a constraint that requires right-alignment of the affixal [+round] with the stem, outranking MAX(approx), a constraint that bans deletion of sonorant continuants, as see in (5).

- (5) Root-final consonant deleted in [bo.kal ~ bo.ko] ‘jar’ due to ALIGN-R(affix, stem)  $\gg$  MAX

/bokal + [+round]/	IDENT- $\sigma_1$ (round)	MAX(float)	IDENT(round)	ALIGN	MAX(approx)
a. bo.kal		*!			
b. bo.kol			*	*!	
c.  bo.ko			*		*

The complete analysis must also prevent the non-low vowels of the language from alternating, e.g. \*/gil + [+round]/  $\rightarrow$  \*[gy]. We exclude the front unrounded vowels [i, e,  $\epsilon$ ,  $\tilde{\epsilon}$ ] with a faithfulness constraint that is specific to front vowels, IDENT(round)/front (independently needed in this language, which contrasts roundness on front vowels only). We exclude the round vowels [u, o,  $\omega$ ,  $\tilde{\omega}$ , y,  $\omega$ ,  $\omega$ ] with \*VACDOC (Wolf 2007), a constraint that penalizes the vacuous docking of a [+round] feature on a round vowel. The remaining non-front, non-round vowel, [ã], is exceedingly rare before word-final sonorants, i.e. the constraint \*[ã][+son]# is generally respected in the language, and removes potential targets for the floating [+round]. Alternatively, one could enlist \*MAP constraints (Zuraw 2013) to rule out the unattested mappings.<sup>3</sup>

To prevent the deletion of final consonants other than [l] and [j], e.g. \*/gak + [+round]/  $\rightarrow$  \*[go], we employ MAX(obstruent) and MAX(nasal) to block the alternations from words that do not end in a sonorant continuant. It is thus safe to allow the floating round feature to attach to any stem in the language; the analysis limits the possibility of docking to [al/aj]-final items.

In the lexicon, [l] is more deletable than [j]; we can incorporate this fact into our analysis by further relativizing MAX to the particular type of deleted segment, i.e. introducing MAX(glide), a

<sup>3</sup>Lexique lists only two [ãl]-final nouns, and no [ãj]-final words of any kind. Both nouns are listed with a faithful plural. Given the extreme rarity of final [ãl], the absence of alternations could also be an accidental gap.

constraint that penalizes the deletion of [j] but not [l]. As we will see in §3, however, speakers failed to extend the greater deletability of [l] to nonce words. We propose that speakers are biased against learning a discontinuous treatment of the sonority scale: since French prevents deletion of low sonority segments (obstruents and nasals), learners expect the deletion pattern to continue to follow the sonority scale, making liquids either less deletable or as deletable as glides, but not more deletable. The French lexicon offers a disfavored pattern, and speakers leave it unlearned. Such “surfeit of the stimulus” (Becker et al. 2011, 2012) cases, or anti-Universal patterns, would be available to purely statistical learners that are not equipped with the biases that humans bring to the table.

One could imagine that speakers would learn to prefer [al→o] over [aj→o] by deploying a markedness constraint against final [al] or final [l]; however, given that final [(a)l] is extremely common in French across the entire lexicon, it is hard to see how it could be assigned a sufficiently large weight to tip the scale.<sup>4</sup> We return to the question of the final consonant in §2.3.4 and §3.4.

### 2.3.2 USELISTED protects established lexical items

Most existing lexical items in our data have an established plural. For example, the plural of [bokal] ‘jar’ is normally the unfaithful [boko], while the faithful plural is practically never attested for this word. Novel items and infrequent items are often subject to more variation, e.g. both plurals are attested for [val] ‘valley’ (see §2.1). The USELISTED approach (Zuraw 2000; Hayes & Londe 2006; Becker et al. 2012) is specifically designed to account for the difference between established and novel items, as seen in a variety of languages. A probabilistic grammar is created based on the statistical trends in the existing lexical items, but these items are protected from the same trends by the USELISTED constraint. The original proposal in Zuraw (2000) used a stochastic version of Optimality Theory, but because stochastic OT was since discovered to be intractable (Pater 2008a; a.o.), we use the mathematically solid Maximum Entropy approach (Goldwater & Johnson 2003; Hayes & Wilson 2008; White 2014), implemented using Hayes & Wilson’s (2008) MaxEnt Grammar Tool.<sup>5</sup>

A MaxEnt grammar is a multinomial logistic regression model, and thus is quite similar to the binary logistic regression model we presented in Table 1; both models use the probability of the alternating plural as the dependent variable.<sup>6</sup> Using the MaxEnt Grammar Tool with its default

---

<sup>4</sup>A related explanation for the final-consonant mismatch between the lexicon and the nonce words would capitalize on the fact that the [o] product of deletion is identical whether its source is [al] or [aj]. A preference for product-oriented generalizations could steer speakers away from the effect of the source (Bybee & Slobin 1982; Bybee & Moder 1983; Bybee 2001; Kapatsinski 2009, 2011); however, Becker & Gouskova (to appear) show that source-oriented generalizations can be productive given sufficient evidence.

<sup>5</sup>Deriving probabilistic predictions from the lexicon can also be achieved with other constraint-based approaches, such as constraint cloning (Pater 2006, 2008b; Coetzee 2008; Becker 2009; Becker, Ketrez & Nevins 2011). Much like USELISTED, constraint cloning produces consistent plurals for established items and gradient predictions for novel items; however, deriving exact numerical predictions using constraint cloning is controversial, and we thus prefer the well-understood MaxEnt approach.

<sup>6</sup>Binominal and multinomial logistic regression models differ in a few respects: (1) The multinomial model allows items with probabilities that are intermediate between zero and one, and thus takes into account the variation in the lexicon. (2) The multinomial model as implemented in the MaxEnt grammar tool included priors that prevent implausibly large estimates in case of separation (=the near perfect alignment of [ε] with lack of alternation), and thus allows us to use both [a] and [ε] items in the model. (3) The multinomial model is expected to double as a generative linguistic model, and thus its predictors are expected to be plausible constraints. As a corollary, generative models will normally keep constraint weights/coefficients uniformly positive or uniformly negative, to prevent a constraint such as \*CODA from switching between penalizing codas and rewarding codas. Regression models more generally allow estimates to be either

settings ( $\mu = 0$ ,  $\sigma = 100,000$ ), we trained the grammar on the probabilities of the alternating plurals in the lexicon, which included the 118 [al/aj]-final words from §2.2, and a sample of 100 [el/ɛj]-final words from Lexique. Five constraints were used in the analysis: four of these are seen in the tableau in (6), with the weights assigned to them by the MaxEnt grammar tool, while the fifth, IDENT(round), was assigned a weight of zero, i.e. it made no numerical contribution to the analysis. The harmony ( $\mathcal{H}$ ) of each plural is the sum of multiplying each violation mark by the weight of the violated constraint; this harmony is then exponentiated, and dividing these exponents by their sum in each tableau yields the expected probability. The tableau in (6) shows two plurals for the nonce monosyllabic singular [gnal]: the faithful plural [gnal] and the alternating [gno]. The expected probability of the alternating monosyllabic plural [gno] is 35%, less than that of the faithful plural.

(6) MaxEnt prediction for an [al]-final monosyllable

/gnal + [+rd]/	IDENT (round)/front $w = 4.83$	IDENT- $\sigma_1$ (round) $w = 1.29$	MAX(glide) $w = 1.01$	MAX(float) $w = .67$	$\mathcal{H}$	$p$
a. gnal				*	-.67	.65
b. gno		*			-1.29	.35

To summarize, the MaxEnt analysis is trained on the plurals of real words of French and uses their distribution to calculate constraint weights. While the plurals of existing words are protected by USELISTED, the plurals of nonce words are assigned a probability based on their weighted violations of the constraints in (6). The constraint USELISTED was not used during training, and thus was not assigned a weight by the MaxEnt tool. Its weight can be assumed to be arbitrarily high.

### 2.3.3 Protection of monosyllables

We have seen in (6) that the [a→o] change in the nonce word [gnal ~ gno] violates IDENT- $\sigma_1$ (round), since the stem change impacts the word’s first (and only) syllable. In contrast, the same [a→o] change in a nonce polysyllable such as [guval ~ guvo] does not violate IDENT- $\sigma_1$ (round), since the initial syllable [gu] is left unchanged (7). As a result, the predicted probability of the alternation in an [al]-final polysyllable with the given constraint weights is 0.66, and more likely than the faithful form, in contrast to tableau (6).

(7) MaxEnt prediction for an [al]-final polysyllable

/guval + [+rd]/	IDENT (round)/front $w = 4.83$	IDENT- $\sigma_1$ (round) $w = 1.29$	MAX(glide) $w = 1.01$	MAX(float) $w = .67$	$\mathcal{H}$	$p$
a. guval				*	-.67	.34
b. guvo					0	.66

It is worth discussing the differences between an account based on protection of initial syllables, as proposed here, and one based on specific protection of *monosyllables*. While the two kinds

positive or negative. (4) The multinomial model as implemented in the MaxEnt grammar tool does not estimate error, and thus does not generate  $p$ -values, though see Hayes et al. (2012) for the use of likelihood tests with MaxEnt models.

of protection overlap completely in the cases we examine in this paper, the predictions diverge in other cases. For example, the differential protection of codas in the initial and non-initial syllables of polysyllabic words in Tamil (Beckman 1997, 1998) does not single out monosyllables as special. Conversely, a hypothetical case of prefixation that impacts the left edge of polysyllables but not monosyllables, if observed, would require a protection of monosyllables, and would not be amenable to a treatment in terms of initial syllable protection. While cases of monosyllable-specific faithfulness are uncommon in the literature, some relevant evidence may come from Kirk & Demuth (2006), an acquisition study in which children showed higher accuracy for coda consonants in monosyllabic words vs. iambic or trochaic disyllabic words. We adopt initial syllable faithfulness here for its broader coverage of the cases that are known to us and the majority of cases that are described in the literature, while fully acknowledging that the specific set of results here could in fact be handled entirely with monosyllabic-specific constraints.

### 2.3.4 Summary

The proposed analysis uses a floating [+round] plural suffix, which allows the alternation to occur as a synchronic morphophonological process. The floating feature can be safely attached to any given noun or adjective, since the grammar only allows the feature to dock on [al/aj]-final words; other words are protected from change with a family of faithfulness constraints.

A weighted constraint MaxEnt grammar is trained on the established words of the language to assign constraint weights, and uses these weights to predict the probability of alternation of novel words. When the alternation impacts a monosyllable, it incurs a violation of *initial syllable faithfulness*, which decreases the predicted acceptability of the alternation. In §3 below, we show that speakers extend this differential treatment of monosyllables from the words in their lexicon to nonce words.

Our analysis strongly penalizes the alternation of [ɛl/ɛj]-final nouns using IDENT(round)/front, a constraint that is independently needed in this language, which contrasts rounding on front vowels. We will see that participants strongly reject the alternation with [ɛ].

In the lexicon, [l] is more deletable than [j], which the analysis captures with a MAX(glide) constraint that penalizes the deletion of [j]. In our experiments, however, participants do not seem to disfavor the deletion of [j]. If indeed one concludes that [l] and [j] are equally deletable, the model can be revised to diminish the effect of MAX(glide), keeping its weight closer to zero using a Bayesian prior (cf. similar use of priors in MaxEnt models in Wilson 2006, White 2014, a.o). In the next section, we assess the extent to which speakers have generalized the trends found throughout the lexicon when it comes to governing the treatment of alternation in nonce words.

## 3 Nonce testing the French plural

The irregular plurals of French have received little attention in the literature, perhaps due to a prevalent intuition about their lack of productivity. Indeed, the alternation is synchronically unnatural, and does not seem to extend to loanwords. The number of [al/aj]-final monosyllables is rather small ( $n = 16$ ), and thus might not intuitively suggest overwhelming evidence for their

protection. We present a series of three experiments that establish that French speakers do extend the alternation to nonce words, and in particular, strongly protect the monosyllabic ones.

In Experiment 1 and Experiment 2, we asked participants to choose between a faithful plural and an unfaithful plural. The choice is presented on a scale of 1–7 in Experiment 1, and binarily in Experiment 2. In both experiments, the nonce words are presented auditorily using naturally produced speech. Additionally, in Experiment 2, participants were asked to rate the singular before they judge the plural. In both experiments, monosyllables were protected from alternation.

In the naturally produced materials of Experiment 1 and 2, the phonetic duration of the final [al/aj] sequences is inversely proportional to the duration of the whole word. Thus, the [al] of the short word [gnal] is phonetically longer in duration than the [al] of longer word [guval]. Increased segmental duration has been suggested as a source of protection from alternations, thus making the [al] of monosyllables less likely to change to [o] (see §6). In experiment 3, we switch from naturally recorded stimuli to synthesized stimuli that keep segmental duration constant, i.e. the [al] of [gnal] and the [al] of [guval] were identical in duration. Participants nonetheless protected monosyllables; the implications are discussed further in §6.

### 3.1 Experiment 1: scalar judgments

#### 3.1.1 Participants

The participants were recruited online using word of mouth, and volunteered their time and effort without compensation. Data was gathered from 185 people who completed the survey and self-identified as being born in France and being at least 18 years old; the rest was discarded. 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). According to self-reports, 116 females and 65 males completed the experiment; 4 did not answer this question. 83 participants listed no other language spoken, or indicated that they are monolingual. Of those who listed other languages, 76 indicated some knowledge of English, 18 Spanish, 15 German, and small numbers of other languages.

#### 3.1.2 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); recall that alternations in the lexicon are largely limited to [a] as compared to [ɛ], and we included the latter as a baseline for comparison. All nonword 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 B. 12 loanwords from English were used as fillers, e.g. [stek] 'steak', [dolar] '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 the 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. The recordings were not manipulated in any other way, other than normalizing the intensity with Praat (Boersma & Weenink 2015), and sounded natural.

### 3.1.3 Procedure

The experiment was run in Experigen (Becker & Levine 2012), and 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. 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 choice of more items containing [a] than [ɛ] reflects our expectation for a greater rate of alternations in the former. The choice of more [al] than [aj] approximates the greater prevalence of [al] in the lexicon. An additional 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 do not 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. Frame sentences were created for each of the 12 fillers.

Before the experiment began, each 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⟩. In the following screen, the item [dal] appeared in its frame sentence, as schematized in Figure 2. A sound button played the plural [dal] when pressed, and upon completion, another sound button appeared. When pressed, the plural [do] was played, and the question "Which plural do you prefer?" appeared, with seven numbered buttons. The edges of the scale were labeled "the one on the left" and "the one on the right". Once one of the seven buttons was pressed, a new screen with the next item appeared.

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 demographic questions (country of origin, current location, year of birth, sex, other languages spoken, suggestions).

### 3.1.4 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); this comparison simply reflects the distribution based on vowel quality found in the lexicon. Both of these effects can be seen in Figure 3, where error bars show 95% confidence

J'aimerais bien avoir un petit *dal*.  
C'est pas ta mère qui passe là avec deux \_\_\_\_\_ ?

> >

Quel pluriel préférez-vous ?

celui à gauche  1  2  3  4  5  6  7 celui à droite

Figure 2: Example trial showing a singular base and two plural forms of it. The order of the two plural forms was randomized, and participants were asked to indicate whether they preferred the plural on the left or the plural on the right.

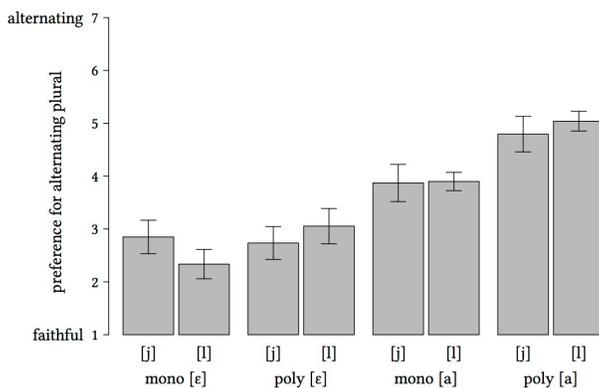


Figure 3: Experiment 1: French nonce monosyllables protected from alternations ( $n=185$ )

intervals. Figure 3 also shows an interaction: the monosyllabicity effect is much stronger for [a] than it is for [ε]. The difference between [j]-final and [l]-final items was overall rather small (3.6 vs. 4.0). The raw results are available at <http://becker.phonologist.org/projects/FrenchPortuguese/>.

To assess the statistical strength of these effects, a mixed-effects linear regression model was fitted using the *lmer* function from the *lme4* package (Bates & Maechler 2013) in R, with rating of the alternation as the dependent variable.<sup>7</sup> The following predictors were used: *monosyllabic*, a binary predictor that contrasted monosyllables and polysyllables, *vowel*, a binary predictor that contrasted [a] with [ε], and *consonant*, a binary predictor that contrasted [l] with [j]. A fully crossed model is reported in Table 2.<sup>8</sup> The model confirms that alternations are significantly dispreferred with monosyllables (negative  $\beta$ ) and significantly preferred with [a] relative to [ε], just as expected based on the lexicon (§2). A significant interaction shows that the effect of monosyllabicity is stronger with [a]. The difference between [l] and [j] came out as a weak trend, suggesting that the

<sup>7</sup>On the use of a linear regression to model a bounded 1–7 scale, see the introduction to §3.2. A logistic regression with an arbitrary cut-off at 4 yielded an essentially identical result.

<sup>8</sup>The model reported in Table 2 is a fully crossed model, which includes *monosyllabic*, *vowel*, and *consonant* and all of their interactions as fixed effects and as random effects by *participant*. We also included all within-item predictors as random effects by *item*, i.e. we included *vowel*, *consonant*, and their interaction, but not *monosyllabic*, since it was a between-item predictor; items were coded without their final vowel and consonant. Following an anonymous reviewer’s recommendation, no model simplification was undertaken. The predictors were coded with a mean of zero and a range of one, viz. *monosyllabic*:  $-.54 +.46$ , *vowel*:  $-.69 +.31$ , and *consonant*:  $-.69 +.31$ . This fully crossed model has low collinearity measures ( $VIF \leq 1.14$ ,  $\kappa = 2.09$ ), calculated using *mer-utils*, by Austin Frank, available at <https://github.com/aufrank/R-hacks>.

	$\beta$	SE( $\beta$ )	$t$	$p$ -value <sup>9</sup>
(Intercept)	3.90	.10	40.84	
<i>monosyllabic</i>	-.83	.13	-6.41	<.001
<i>vowel=[a]</i>	1.67	.14	12.23	<.001
<i>consonant=[l]</i>	.06	.13	.47	>.1
<i>monosyllabic:vowel</i>	-.60	.24	-2.53	<.05
<i>monosyllabic:consonant</i>	-.35	.22	-1.56	>.1
<i>vowel:consonant</i>	.25	.25	1.02	>.1
<i>monosyllabic:vowel:consonant</i>	.58	.48	1.21	>.1

Table 2: Experiment 1: Preference for alternating plurals on a 1–7 scale in French nonce words. Positive  $\beta$  reflects more alternation.

two consonants may be equally deletable.

### 3.1.5 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. We will show in §3.4 that the lexical predictions are strongly correlated with the experimental results. 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, and thus from the linguist’s perspective, there is no clear benefit to encoding the size-based faithfulness effect in the synchronic grammar. The discrete and grammatical treatment of this pattern is apparently irresistible to the learner, who generalizes over the lexicon using salient grammatical principles, even if those generalizations no longer apply to existing words over a natural alternation.

## 3.2 Experiment 2: binary judgements

Experiment 1 established that French speakers prefer alternating plurals most in polysyllables that have [a] in their singular form. Two questions remain, however. Firstly, it is not clear what role the wellformedness of the singular might have on the choice of plural. Secondly, the use of a linear regression in the analysis of responses on a 1–7 scale is potentially problematic, because it assumes that participants are using the scale linearly (i.e., the distance between 1 and 2 reflects the same difference in acceptability as the distance between 2 and 3, etc). To alleviate these concerns, the next experiment asked participants to rate the singular base before they judge the plural forms. Plural forms are presented for judgment in a binary yes-or-no task, to be analyzed with a logistic model. The materials and methods were kept exactly the same as in Experiment 1, with the sole difference being in the format of the individual trials.

<sup>9</sup>Exact p-values could not be obtained with *lme4* or *pvals.fnc*, due to current disagreement in the community about the best way to calculate them. We used the normal approximation to estimate p-values, i.e. we assumed infinite degrees of freedom and used p-values from the z distribution.

J'aimerais bien avoir un petit .

Évaluez ce mot; pourrait-il entrer dans la langue française ?

non      oui

C'est pas ta mère qui passe là avec deux  ?

Acceptez-vous cette forme comme le pluriel du mot donné ?

C'est pas ta mère qui passe là avec deux  ?

Acceptez-vous cette forme comme le pluriel du mot donné ?

Figure 4: Example trial showing a singular base, rated on a 1–5 scale, and two plural forms of it, each presented for binary judgment. The order of the two plural forms was randomized.

### 3.2.1 Participants

The participants were recruited online using word of mouth, and volunteered their time and effort without compensation. Data was gathered from 52 people who completed the survey and self-identified as being born in France and being at least 18 years old; the rest was discarded. The server logs indicate that these 52 participants took on average 7 minutes to complete the experiment (range 3–27 minutes, median 5). The participants reported an average age of 32 (range 22–66, median 30). According to self-reports, 31 females and 21 males completed the experiment. 9 participants listed no other language or indicated that they are monolingual. Of those who listed other languages, 41 indicated some knowledge of English, 13 Spanish, 10 German, and smaller numbers of other languages.

### 3.2.2 Materials

The materials are identical to those of Experiment 1.

### 3.2.3 Procedure

Each item was first presented in the singular, in a frame sentence (as in Experiment 1), but with the singular written on a button. Pressing the button played the singular form and then a question was displayed, asking the participant to rate the item as a word of French on a scale of 1–5, as shown in Figure 4. Once one of the 1–5 buttons was pressed, all buttons were disabled (turned gray and unpressable), and one of the plurals (either faithful or alternating, randomly chosen) was presented in a frame sentence containing a sound button. When pressed, the plural was played, and a question was displayed, asking the participant whether the plural was acceptable (yes or no). Once one of the two buttons was pressed, both were disabled, and the other plural was presented in the same way, again asking for a yes or no judgment of its acceptability. Pressing one of the final two buttons then moved the participant to the next item.

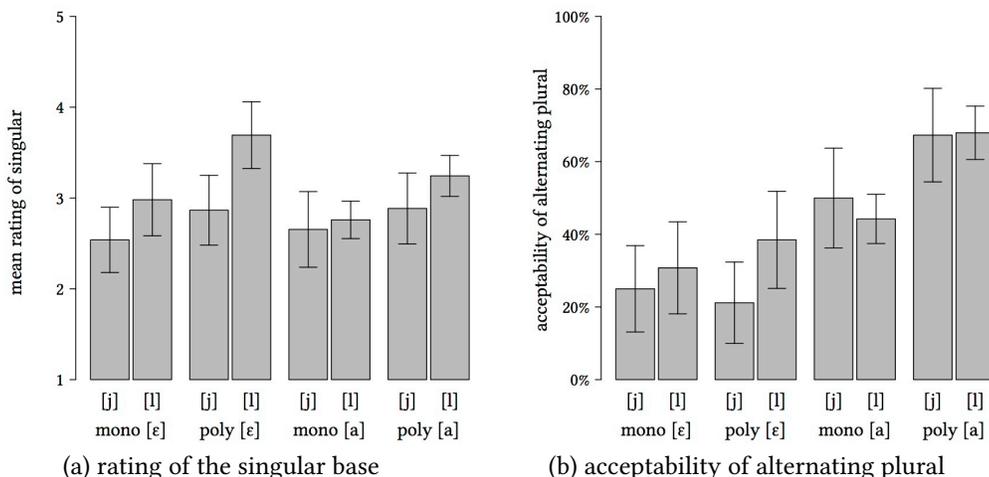


Figure 5: Experiment 2: Acceptability of alternating plurals in French nonce words ( $n=52$ )

### 3.2.4 Results

As in Experiment 1, speakers accepted the alternating [o/ø] plurals with monosyllabic items significantly less than with polysyllabic items (40% vs. 55%). In addition, they accepted alternating plurals with [a] significantly more than plurals with [ε] (55% vs. 29%). Both of these effects can be seen in the bar plots in Figure 5b. Figure 5b also shows an interaction: the monosyllabicity effect is much stronger for [a] than it is for [ε]. The difference between [j]-final and [l]-final items was overall rather small and insignificant (41% vs. 50%), and it was even smaller for the [a] items (59% vs. 54%).

In this experiment, speakers also rated the singulars on a scale of 1–5 before they judged the plural forms, seen in Figure 5a. Overall, participants found the singulars to be of intermediate acceptability (3.0 on the 1–5 scale). Polysyllables were rated higher than monosyllables (3.2 vs. 2.7), and items with final [l] were rated higher than items with final [j] (3.0 vs. 2.7). Items with [ε] were rated slightly higher than items with [a] (3.0 vs. 2.9). This small difference between [ε] and [a] confirms that speakers successfully separated the rating of the singulars from the acceptability of the plural forms.

The acceptability of the faithful plurals mirrored the acceptability of the alternating plurals, so we do not report these here. The complete raw results are available at <http://becker.phonologist.org/projects/FrenchPortuguese/>.

The acceptability of the alternating plurals was modeled with a mixed-effects logistic regression, using the following predictors: *base*, or the acceptability of the singular base on the 1–5 scale, and the full crossing of the three predictors *monosyllabic*, *vowel*, and *consonant*, defined as in Experiment 1, as reported in Table 3.<sup>10</sup>

<sup>10</sup>The model was constructed as follows: We started with a fully crossed model that had *base*, and the full crossing of *monosyllabic*, *vowel*, and *consonant* as fixed effects and as random slopes for *participant* and *item*, excluding a random slope for *monosyllabic* given *item*, as explained above. Since this model did not converge, random slopes were removed gradually, following the procedure in Barr et al. (2013). The first model that converged lost all random slopes for *item*, while *participant* only lost the three-way interaction and the interaction of *vowel* and *consonant*. No simplification of the fixed effects was attempted. The predictors were coded with a mean of zero and a range of one, with *base* coded on a range from  $-0.49$  to  $+0.51$ , and the other predictors coded as in Experiment 1. This final model enjoyed low collinearity measures ( $VIF \leq 1.3$ ).

	$\beta$	SE( $\beta$ )	$z$	$p(> z )$
(Intercept)	-.10	.17	-.60	
<i>base</i>	.73	.30	2.48	<.05
<i>monosyllabic</i>	-.84	.21	-3.99	<.001
<i>vowel=[a]</i>	1.58	.25	6.34	<.001
<i>consonant=[l]</i>	.08	.22	.37	>.1
<i>monosyllabic:vowel</i>	-1.05	.50	-2.10	<.05
<i>monosyllabic:consonant</i>	-.48	.42	-1.14	>.1
<i>vowel:consonant</i>	-.82	.44	-1.87	>.05
<i>monosyllabic:vowel:consonant</i>	.70	.88	.79	>.1

Table 3: Experiment 2: Acceptability of alternating plurals in French nonce words, showing protection of monosyllables. Positive  $\beta$  reflects more alternation.

The model in Table 3 shows that alternating plurals are highly significantly preferred in polysyllables and with [a], including a significant interaction. The difference between [l] and [j] did not reach significance at the .05 level. The correlations of *base* with the other predictors were very low ( $r < .14$ ), suggesting that the observed effects of monosyllabicity and stem vowel go above and beyond the rating of the singular base. To confirm this point, we factored out the potential correlation of *base* with *monosyllabic* and *vowel* using residualization, i.e. taking the remaining predictive power of *base* once it did its best to predict *monosyllabic* and *vowel*; the residualized model is essentially identical to the one reported here.

### 3.2.5 Discussion

Overall, this experiment confirms and replicates the results of Experiment 1. Experiment 2 sharpens the interpretation of the results of Experiment 1 in two ways: Firstly, the singular base form was rated on its own, in addition to the plurals. The regression results suggest that the protection of monosyllables is statistically independent from the phonotactic wellformedness of the base. Secondly, the plurals were judged binarily, and thus could be modeled statistically using a logistic regression. This alleviates the potential concern about modeling a scalar rating using a linear regression. Importantly, Experiment 2 confirms the three strong effects found in Experiment 1: the difference between monosyllables and polysyllables, the difference between [a] and [ε], and their interaction. Neither experiment showed a significant effect of the final consonant. In short, the grammatically-mediated difference between monosyllables and polysyllables in nonce plural formation was robustly upheld with a wholly different rating methodology.

## 3.3 Experiment 3: no duration cues

Cross-linguistically, monosyllables are usually produced with segments that are phonetically longer than segments in polysyllables; this trend holds true in French generally, and it holds of the materials used in Experiments 1 and 2 specifically. Thus, a monosyllable such as [gnal] has a phonetically longer [al], while a disyllabic [guval] has a phonetically shorter [al]. As will be discussed in §6 below, longer segmental duration has been proposed as a source of protection from alternations, and according to this view, our materials in Experiments 1 and 2 provided participants

$\kappa=2.03$ ).

with a phonetic cue that could be used in the judgments of the plural forms. To control for this effect, the current experiment uses artificially created auditory stimuli that maintain constant segmental duration; the experiment is otherwise identical to Experiment 2.

### 3.3.1 Participants

The participants were recruited online using word of mouth, and volunteered their time and effort without compensation. Data was gathered from 71 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 71 participants took on average 6 minutes to complete the experiment (range 3–18 minutes, median 5). The participants reported an average age of 35 (range 19–79, median 31). According to self-reports, 46 females and 22 males completed the experiment; 3 did not say. 5 participants listed no other language or indicated that they are monolingual. Of those who listed other languages, 62 indicated some knowledge of English, 25 Spanish, 14 German, and smaller numbers of other languages.

### 3.3.2 Materials

The same words were used as in Experiment 1 and 2. The auditory materials were not recorded by a native speaker, but rather generated by the MBROLA text-to-speech system (Dutoit et al. 2006). Vowels were generated specifying a level pitch of 130Hz, and a duration of 280ms in final (stressed) position and 140ms in non-final (unstressed) position. Consonants were specified with durations ranging from 80ms to 160ms, depending on their identity, but not on their position. The pitch and duration values were chosen to match the averages obtained from the naturally produced materials. The materials, then, had identical final syllables in, e.g., [zal] and [vøzal]. While the resulting materials sounded decidedly robotic, their segments were easily recognizable.

### 3.3.3 Procedure

The procedure of Experiment 2 was identical to Experiment 1.

### 3.3.4 Results

The results are overall similar to those in Experiment 2, with alternating plurals being accepted more with [a] than with [ɛ] (55% vs. 17%), with [l] more than with [j] (48% vs. 32%), and with polysyllables more than monosyllables (48% vs. 39%), as seen in Figure 6b. Again, the ratings of the singulars show a preference for polysyllables over monosyllables (3.5 vs. 3.1 on the 1–5 scale), for [l] over [j] (3.4 vs. 3.1) and for [ɛ] slightly more than [a] (3.4 vs. 3.2), as seen in Figure 6a.

The statistical analysis was performed as for Experiment 2, and reached similar results.<sup>11</sup> The final model in Table 4 shows that alternating plurals are significantly more acceptable when the rating of the base is higher and with the vowel [a]. More importantly, alternating plurals are

<sup>11</sup> We started with a fully crossed model that had *base*, and the full crossing of *monosyllabic*, *vowel*, and *consonant* as fixed effects and a random slopes for *participant* and *stem*. Since this model did not converge, random slopes were removed gradually. The first model that converged lost all random slopes for *stem*, while *participant* was only left with random slopes for *monosyllabic*, *vowel*, *consonant*, and the interaction of *vowel* and *monosyllabic*. This final model enjoyed low collinearity measures ( $VIF \leq 1.6$ ,  $\kappa = 2.11$ ).

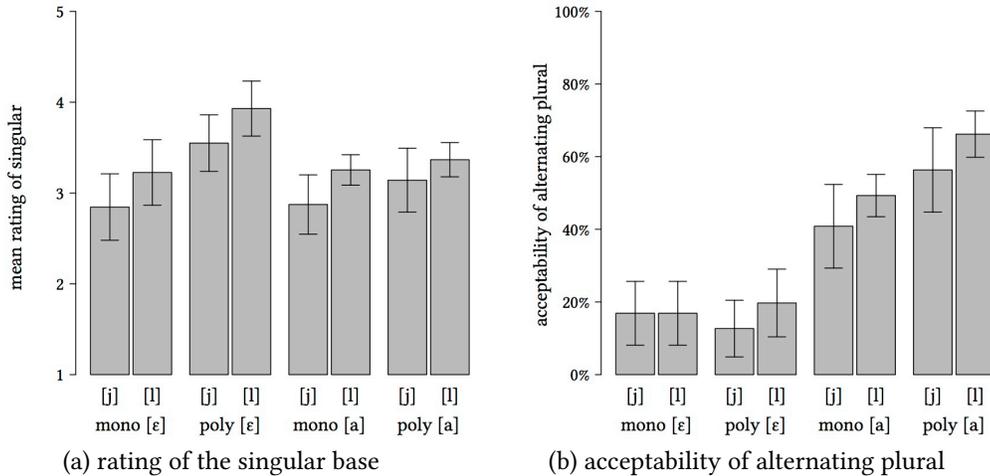


Figure 6: Experiment 3: Acceptability of alternating plurals in French nonce words ( $n=71$ )

	$\beta$	SE( $\beta$ )	$z$	$p(> z )$
(Intercept)	-.78	.22	-3.51	
<i>base</i>	1.13	.31	3.62	<.001
<i>monosyllabic</i>	-1.03	.27	-3.84	<.001
<i>vowel=[a]</i>	3.43	.39	8.84	<.001
<i>consonant=[l]</i>	.39	.22	1.81	>.05
<i>monosyllabic:vowel</i>	.41	.71	.57	>.1
<i>monosyllabic:consonant</i>	-.33	.41	-.79	>.1
<i>vowel:consonant</i>	.35	.50	.69	>.1
<i>monosyllabic:vowel:consonant</i>	.79	.98	.80	>.1

Table 4: Experiment 3: Acceptability of alternating plurals in French nonce words. Positive  $\beta$  reflects more alternation.

accepted significantly more often in polysyllables than in monosyllables. There was no significant interaction of vowel and monosyllabicity.

As reported for Experiment 2 above (§3.2.4), we found that residualizing *monosyllabic* and *vowel* against *base* makes very little change to the model for Experiment 3, confirming that the preference for alternating plurals in polysyllables goes above and beyond the preference for polysyllabic singulars. We also fitted a regression model for the combined results of Experiment 2 and Experiment 3, with an added *experiment* factor; this combined model is only minimally different from the ones we report for each experiment individually, and no significant difference between the experiments was found.

### 3.4 Comparison of lexicon with experimental results

The three experiments confirm that speakers prefer alternating plurals when they are polysyllabic, and faithful plurals when they are monosyllabic. Experiment 2 and Experiment 3 also show that this preference goes above and beyond the higher ratings given to polysyllabic singular bases. Experiment 3 confirms that speakers prefer alternations in polysyllables even in the absence of segmental durational cues for monosyllabicity in the auditory materials they heard. This

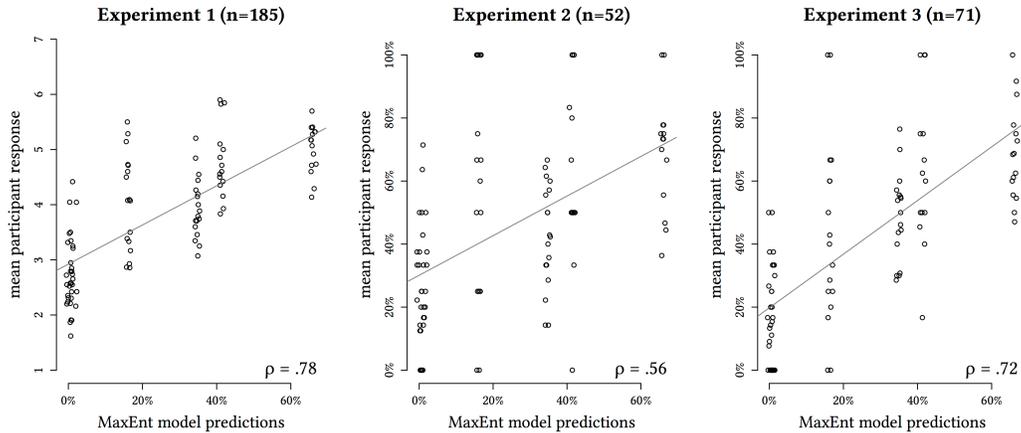


Figure 7: Participants’ judgment of alternating plurals of nonce items, plotted by the plurals’ probability as predicted by the MaxEnt analysis, with horizontal jitter.

result shows that speakers have their own internal expectations about the distribution of plural alternations and that these expectations do not require support from the phonetic properties of the particular tokens presented to them.

All three experiments also showed a preference for alternations with [a] relative to [ɛ], as expected. There was a significant interaction of vowel and monosyllabicity in Experiments 1 and 2. In Experiment 3, a much larger estimate for the vowel effect left the interaction insignificant.

Comparing the predictions of our MaxEnt analysis in §2.3 to the participants’ responses per item, we see a strong and highly significant correlation, shown in the three panels of Figure 7; the Spearman rank correlation coefficients are .78, .56, and .72 for the three experiments, all significant at  $p < .001$ . The lowest correlation, which is still rather strong, is found in the experiment with the smallest number of participants. Differently stated, there is a close match between the predictions of our lexicon-trained grammar and the observed treatment of nonce words. Our model by no means captures the entirety of the variance in the experimental results; in particular, there is a great deal of variance within each of the categories imposed by the MaxEnt model. It is thus possible that the prediction could be further improved perhaps using more fine-grained details regarding onset profiles.

The three experiments did not show a significant effect of final consonant. This result was somewhat unexpected given the predictions of our lexicon regression model and our MaxEnt analysis, and the general expectation that lexical trends usually extend to nonce words (Hayes & White 2013, contra Becker et al. 2011). We suggested that this “surfeit of the stimulus” case is due to the sonority hierarchy: if deletion of low-sonority segments is categorically blocked, speakers are biased to extend this trend upwards in the scale, not allowing liquids to be more deletable than glides. This effect can be incorporated into the MaxEnt analysis with a defeasible Bayesian prior that requires the weight of MAX(glide) to be closer to zero. However, an additional explanation suggests itself: recall footnote 2, in which we mentioned the collinearity of final consonant with part-of-speech. Should learners opt for part-of-speech as the relevant generalization and completely ignore final consonant as a result, no effect of final consonant would be expected in our experiments. To test such a hypothesis, one would need to design an experiment explicitly

contrasting *l*-final nouns and adjectives.

We conclude this section by focusing on the importance of the monosyllabic vs. polysyllabic effect in determining whether a noun will undergo the irregular plural alternation in French. In the next section, we turn to a parallel alternation in Portuguese, where 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.

## 4 A study of alternations in the Brazilian Portuguese lexicon

We now turn to another language where monosyllables are protected from alternations, 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 one finds a [w → j] alternation, e.g. [ʒofi'naw ~ ʒofi'najs] 'newspaper'. While the languages have a common Latin root, the changes occurred after the languages separated, and the details of historical development diverge. Nonetheless, the Portuguese pattern also falls along similar lines, whereby initial syllables are protected from the resulting irregular plural formation process for this specific group of nouns.

### 4.1 The Brazilian Portuguese regular plural

For most nouns in Portuguese, the plural is completely regular and predictable. In this language, nouns overwhelmingly end in a vowel, a glide, or one of the consonants [f, s].<sup>12</sup> There are two plural suffixes available, in complementary distribution: [-s] after vowels and glides, otherwise [-is] after consonants. More specifically, the [-s] allomorph is observed after an oral vowel, [j], or any nasal vowel other than [ẽw̃], e.g. [ˈbafiku ~ ˈbafikus] 'ship', [e'ɾɔj ~ e'ɾɔjs] 'hero', [ifẽ ~ ifẽs] 'hyphen'. The [-is] allomorph is regularly added to nouns that end in [f, s], e.g. [ˈflofi ~ ˈfloris] 'flower', [ˈvɔjs ~ ˈvɔzis] 'voice'. No suffix is added to nouns that already end in an unstressed vowel and [s] in the singular, e.g. [ˈlapis] 'pencil', [ˈõnibus] 'bus'.

This leaves the nouns that end in [w] or in [ẽw̃]. The rest of this paper focuses on [w]; for the similar yet distinct patterning of [ẽw̃], 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'nejs] 'ring', as in (8). We analyze this alternation pattern as taking the plural suffix /-is/, with concomitant fusion of the stem-final [w] and the suffix-initial [i].

When the stem-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 [ˈfɔsiw ~ ˈfɔsejs] 'fossil' (see also Huback 2007).

<sup>12</sup> The word-final pronunciation of these two consonants varies greatly between dialects. The former can be heard as [x], [χ], [h], [r] or [ɹ], the latter as [ʃ]. These sounds usually assimilate in voicing to a following consonant. They consistently alternate with [r] and [z], respectively, when suffixed in the plural.

## (8) Brazilian Portuguese alternating, unfaithful plurals

shape	singular	plural	spelling	
mono	'saw	'sajs	⟨sal⟩	'salt'
	'mɛw	'mɛjs	⟨mel⟩	'honey'
	'pɾɔw	'pɾɔjs	⟨prol⟩	'advantage'
iamb	ʒofi.'naw	ʒofi.'najs	⟨jornal⟩	'newspaper'
	a.'nɛw	a.'nɛjs	⟨anel⟩	'ring'
	ba.'hiw	ba.'his	⟨barril⟩	'barrel'
trochee	'ni.vɛw	'ni.vɛjs	⟨nível⟩	'level'
	'fɔ.siw	'fɔ.sɛjs	⟨fóssil⟩	'fossil'

In contrast to nouns that take the unfaithful plural, other [w]-final nouns form their plural with simple suffixation of [s] and no further change, e.g. [gow ~ 'gows] 'goal', as in (9). We analyze these as taking the plural suffix /-is/ with concomitant deletion of the suffixal [i].

## (9) Brazilian Portuguese non-alternating, faithful plurals

shape	singular	plural	spelling	
mono	'paw	'paws	⟨pau⟩	'stick'
	'hɛw	'hɛws	⟨réu⟩	'culprit'
	'gow	'gows	⟨gol⟩	'goal'
iamb	ka.'kaw	ka.'kaws	⟨cacau⟩	'cocoa'
	tro.'fɛw	tro.'fɛws	⟨troféu⟩	'trophy'
	mu.'zɛw	mu.'zɛws	⟨museu⟩	'museum'
trochee	'aw.kow	'aw.kows	⟨alcool⟩	'alcohol'

Historically, unfaithful plurals originate from nouns with a final lateral (de Kolovrat 1923). Deletion of intervocalic laterals, which happened across the board in Galician-Portuguese (9<sup>th</sup> century), affected the plural, taking [sal ~ 'salis] 'salt' to [sal ~ 'sais]. Much later (18<sup>th</sup> century), coda [l] vocalized to [w], yielding [saw ~ 'sais], with hiatus resolution creating the glide in [saw ~ 'sajs]. The faithful plurals are traced back to nouns that originally ended in [u] or [w], which simply took the plural [s], e.g. [mu'zɛw ~ mu'zɛws] 'museum'. Some nouns with faithful plurals had an intervocalic lateral that deleted in both singular and plural, e.g. [palo ~ 'palos] 'stick' leading to [pao ~ 'paos] and from there via raising of unstressed mid vowels and hiatus resolution to the modern [paw ~ 'paws].

With the vocalization of final laterals, the historical distinction between final [l] and final [w] can only be heard in the plural, as a distinction between alternation and non-alternation. This naturally leads to some fluctuation: some nouns that have normative faithful plurals have developed an innovative unfaithful plural, e.g. [de'graw] 'step' is often heard pluralized colloquially as [de'grajs]. Similar variation is seen in a number of nouns, e.g. [ʃa'pɛw] 'hat', [tro'fɛw] 'trophy', as documented in Abaurre (1983), 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, [saw] 'salt', which has the normative unfaithful plural

[sajs], has the innovative faithful plural [saws]. By hypothesis, initial syllable protection is the factor that causes the innovation of the faithful plural in this monosyllabic noun and others like it; again, we acknowledge that our account is equally consistent with protection of monosyllables rather than initial syllables.

These same trends also extend to loanwords, with monosyllables such as [ˈgow], from English ‘goal’, receiving faithful plurals, and polysyllables such as [kokeˈtɛw], from English ‘cocktail’, receiving unfaithful plurals. A very small number of nouns, most of them monosyllabic, take the plural suffix [-is] with the stem-final [w] surfacing as [l], e.g. [ˈmaw ~ ˈmalis] ‘evil’. In terms of our analysis, these plurals violate IDENT(lateral). The pattern’s rarity suggests very limited productivity, and we thus left it out from our nonce word experiments. Our discussion and investigation is limited to those dialects that have fully merged the historic coda lateral with the labiovelar glide, making the plural only partially predictable. Dialects that maintain the coda lateral, as in some southern Brazilian dialects and most European Portuguese dialects, deserve separate attention in future work.

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. While the productivity is mostly limited to experimental tasks in French, the Portuguese alternation is widely observed to be productive in everyday language use.

### 4.3 Trends in the irregular plurals

We collected the [w]-final nouns and adjectives and their plurals from two word lists.<sup>13</sup> 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 <http://becker.phonologist.org/projects/FrenchPortuguese/>. 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 8 shows all 387 nouns in our list, plotted by prosodic shape. Only about 29% of the monosyllables take unfaithful plurals, while 88% of polysyllables do, i.e. monosyllables are protected from alternation. The rate of unfaithful plurals is lower for iambs (87%) than it is for trochees (96%).

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 8. Among the monosyllables, the lax vowels [a, ɛ, ɔ] are most conducive to alternating plurals, while the tense vowels [e, o, i, u] are most likely to have faithful plurals. While most of the evidence for the effect of vowel tenseness comes from the iambs,

<sup>13</sup>One is LABEL-LEX, available at <http://label.ist.utl.pt/pt/apresentacao.php>. The other is available at [http://artsci.wustl.edu/~tcpol10/palavras\\_explic.html](http://artsci.wustl.edu/~tcpol10/palavras_explic.html).

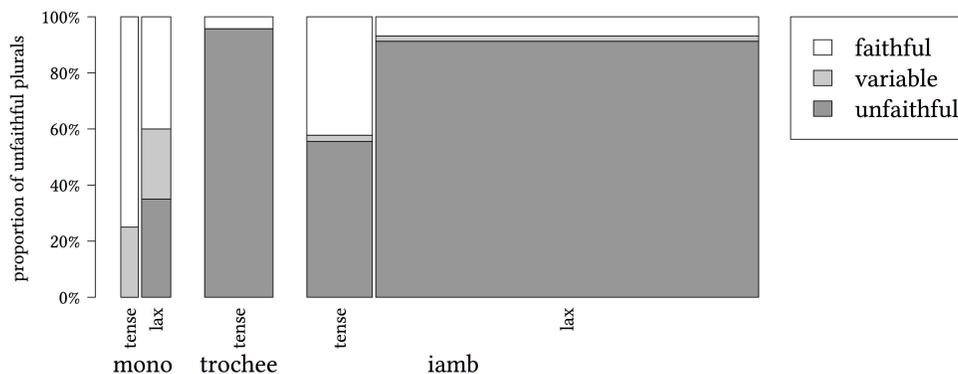


Figure 8: Mosaic plot of Brazilian Portuguese w-final words ( $n=387$ ), by prosodic shape and vowel tenseness; tense = [e, o, i, u], lax = [a, ε, ɔ].

	$\beta$	$SE(\beta)$	$z$	$p(> z )$
(Intercept)	1.22	.29	4.25	
<i>monosyllabic</i>	-4.43	.61	-7.31	<.001
<i>iamb vs. trochee</i>	2.87	.78	3.67	<.001
<i>lax</i>	2.41	.37	6.44	<.001

Table 5: Regression model for the Portuguese lexicon. Positive  $\beta$  reflects higher likelihood of the unfaithful plural.

the effect is visible in the monosyllables as well.

A regression analysis with R's *glm* function was used 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 and alternating plurals. Since 20 of the 387 items had an intermediate faithfulness level, a cutoff level of 50% was chosen; other cutoff levels hardly made any difference.

The model includes a predictor for *prosodic shape*, coded as one factor that contrasts monosyllables and polysyllables, and one that contrasts iambs and trochees. We also included *lax*, a predictor that coded the laxness of final vowel of the root, contrasting the lax [a, ε, ɔ] with the tense [e, o, i, u].<sup>14</sup> This model has reasonably low collinearity ( $\kappa=4.65$ ). The interaction of *lax* and *monosyllabicity* caused separation and was left out,<sup>15</sup> but a likelihood ratio test shows that the interaction does not make a significant improvement to the model ( $\chi^2(1) = 1.7, p > .1$ ).

The model in Table 5 confirms that polysyllables are significantly more likely to take alternating plurals than monosyllables, and trochees are significantly more likely to take alternating plurals

<sup>14</sup>Prosodic shape was coded as two binary factors: the monosyllabicity predictor was coded as +.5 for monosyllables and -.5 for polysyllables, and the iamb vs. trochee predictor was coded as 0 for monosyllables, -.5 for iambs and +.5 for trochees. The laxness predictor was coded as -.73 for tense and +.27 for lax, these values ensuring a mean tenseness of zero. For an overview of contrast coding techniques, see [http://www.ats.ucla.edu/stat/r/library/contrast\\_coding.htm](http://www.ats.ucla.edu/stat/r/library/contrast_coding.htm). These techniques allow the integration of predictors that have more than two levels without introducing unwanted correlations in the model. Other vowel features were explored as predictors, but did not show consistent effects, including height-based groupings that would single out the mid-vowels.

<sup>15</sup>Since the tense monosyllables are (almost) all faithful, the predictor and the outcome align perfectly, and the estimate for the interaction approaches infinity; this is a problem of *separation*, which is a type of *non-identifiability*. See §5.8 of Gelman & Hill (2007) for further detail. A model using the *bayesglm* function from the *arm* package yielded a non-significant interaction.

than iambs. The lax vowels [a, ε, ɔ] are significantly more likely to take alternating plurals than the tense vowels [e, o, i, u].

Token frequency, as estimated by the SUBTLEX-BR corpus of Brazilian Portuguese, added no predictive value to the model, as reported in Tang et al. (2013). Model comparison revealed that while monosyllabicity made a significant improvement ( $\chi^2(1) = 55.7, p < .001$ ) token frequency makes no significant improvement ( $\chi^2(1) = .2, p > .1$ ). While token frequency is significant in a model that does not encode monosyllabicity, the role of token frequency is entirely subsumed by monosyllabicity.

#### 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 initial syllable protection. As we discussed and implemented in §2.3 for French, a MaxEnt grammar combined with the USELISTED approach allows established items to maintain a fixed plural, while novel items are free to vary according to grammatically-filtered trends.

Assuming that the plural suffix is underlyingly /-is/ for consonant-final stems, [w]-final nouns will form their plurals either by deleting the [i], as in /gow + is/ → [gows] ‘goals’ (violating MAX), or by fusing the [w] and [i] to [j], as in /a'new + is/ → [a'nejɐs] ‘rings’, violating IDENT(back) and the anti-fusion constraint UNIFORMITY.<sup>16</sup>

The difference between monosyllables and polysyllables can be attributed to an initial syllable-specific version of IDENT(back).<sup>17</sup> The analysis is outlined in the tableaux in (10-11) and largely parallels our analysis of French in §2.3. Here, IDENT-σ<sub>1</sub>(back) protects a monosyllable such as [gow] ‘goal’ from changing its stem [w], but MAX forces the alternation in a polysyllable such as [a'new] ‘ring’.

(10) Monosyllabic [gow] protected from alternation

/gow <sub>1</sub> + i <sub>2</sub> s/	IDENT-σ <sub>1</sub> (back)	MAX	IDENT(back)
a. <sup>☞</sup> gow <sub>1</sub> s		*	
b. goj <sub>1,2</sub> s	*!		*

(11) Polysyllabic [a'new] fuses with the suffixal [i]

/a'new <sub>1</sub> + i <sub>2</sub> s/	IDENT-σ <sub>1</sub> (back)	MAX	IDENT(back)
a. a'new <sub>1</sub> s		*!	
b. <sup>☞</sup> a'nej <sub>1,2</sub> s			*

The fully faithful concatenation of the base and suffix is ungrammatical in Portuguese, e.g. \*[gowis], \*[a'newis]. We attribute this to a constraint against [w] in onset position, which holds without exception in the derived environment of the plural, and also fairly generally before high vowels throughout the language.

<sup>16</sup>For further detail regarding the evaluation of faithfulness constraints in cases of fusion, see Pater (1999).

<sup>17</sup>This could also be implemented with an initial-syllable version of UNIFORMITY, but we stick to IDENT(back) in the text.

To capture the vowel laxness effect, the analysis requires a constraint that penalizes the alternating plural glides in the presence of tense vowels, i.e. \* $[+tense]j$  (named \*SHALLOWDIPH in (12-13) below). As discussed more fully in Nevins (2012), this constraint militates against shallow, poorly vertically dispersed diphthongs – those that combine a tense vowel with a high glide. In contrast, the steeper diphthongs  $[ɛj]$ ,  $[ɔj]$ ,  $[aj]$  are created more freely. This mirrors the typological preference for steep diphthongs that maximize height differences, cf. the ubiquity of  $[aj]$  and the rarity of  $[uj]$ . Here, we abstract away from the fact that this contrast between steep and shallow diphthongs is limited to the derived environment of plurals; Portuguese allows all possible vowel-glide combinations in monomorphemic words, e.g.  $[boj]$  ‘bull’,  $[hej]$  ‘king’.

As we did for French in §2.3, we fit a USELISTED/MaxEnt grammar that was trained on the probability of the unfaithful plural in the lexicon, and predicts the likelihood or acceptability of the unfaithful plural in nonce words. As before, the MaxEnt Grammar Tool was used with its default settings ( $\mu = 0$ ,  $\sigma = 100,000$ ). We used a total of five constraints in the analysis, four of which are shown with their assigned weights in (12-13). The fifth constraint, IDENT(back), was assigned a weight of zero, and thus makes no numerical contribution to the analysis.

Faithful plurals incur a violation of MAX, while alternating plurals incur violations within the two constraints of the IDENT(back) family. In a polysyllable with a lax vowel like  $[pri'zɛw]$  (12), the alternation impacts the stressed syllable, but not the initial syllable. The created diphthong  $[ɛw]$  is steep and thus satisfies \*SHALLOWDIPH. In a monosyllable with a tense vowel (13), the alternation impacts the stressed syllable, which is also the initial syllable, and it creates the shallow diphthong  $[ej]$ . The equally shallow  $[ew]$  is not derived, and is thus grandfathered in (cf. McCarthy 2003). These three violations add up to lower the expected probability of the alternation.

(12) MaxEnt prediction for an iamb with a stressed lax vowel

$/pri'zɛw + is/$	Max $w = 5.30$	IDENT- $\sigma$ (back) $w = 2.83$	IDENT- $\sigma_1$ (back) $w = 2.81$	*SHALLOW DIPH $w = 2.19$	$\mathcal{H}$	$p$
a. $pri'zɛws$	*				-5.30	.08
b. $pri'zɛjs$		*			-2.83	.92

(13) MaxEnt prediction for a monosyllable with a stressed tense vowel

$/grew + is/$	Max $w = 5.30$	IDENT- $\sigma$ (back) $w = 2.83$	IDENT- $\sigma_1$ (back) $w = 2.81$	*SHALLOW DIPH $w = 2.19$	$\mathcal{H}$	$p$
a. $'grews$	*				-5.30	.93
b. $'grejs$		*	*	*	-7.83	.07

We note that our analysis assumes that the underlying representation of the singular is identical to its surface form, i.e. we do not encode the alternating or non-alternating plural patterning of  $[w]$ -final nouns in their underlying representation. As a result, the grammar can make gradient predictions based on prosodic and segmental factors that follow the lexical distribution of alternations, providing the information that is needed for the derivation of novel words. This marks a departure from the traditional structuralist analysis of Mattoso Câmara (1953), who assumes

underlying laterals for alternating nouns. Our analysis (and purely-auditory experiments) do not rely on the orthographic cues to alternations (where alternations are often signalled with the letter ⟨l⟩). In fact, the innovative alternating plurals in nouns that are spelled without ⟨l⟩, such as ⟨chapéu⟩ with the plural [ʃaˈpejs] ‘hats’, suggest that orthography does not determine whether alternations occur in natural usage. In the two experiments we discuss below, items were presented purely auditorily, thus removing any orthographic cues. To summarize, our analysis predicts that speakers monitor the distribution of faithful and unfaithful plurals in terms of monosyllabicity, vowel quality, and stress. 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 Nonce testing the Brazilian Portuguese plural

The Portuguese lexicon provides strong evidence for the effect of monosyllabicity (which we take to be ultimately an initial syllable protection effect). The [w~j] alternation is very much active, extending to nonce words and overtaking plurals of established lexical items, but never in monosyllables. Nevertheless, we sought to establish the productivity of the alternation experimentally, and in particular its sensitivity to monosyllabicity. The effect of vowel laxness is strong as well, and we will see that it is indeed extended to nonce words. The effect of stress is weaker in the lexicon than the effects of monosyllabicity and laxness, and we will see that it did not extend to nonce words; we discuss this matter in §5.3.

The two experiments we present here closely follow our methodology in §3, with judgments solicited on a 1–7 scale in Experiment 4, and binarily in Experiment 5.

### 5.1 Experiment 4: scalar judgments

#### 5.1.1 Participants

The participants were recruited online using word of mouth, and volunteered their time and effort without compensation. We gathered data from 181 people who completed the experiment 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). According to self-reports, 95 females and 51 males completed the experiment; 35 did not specify.

When asked about the variety 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. 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.

### 5.1.2 Materials

We created a total of 89 [w]-final nonce word items: 47 monosyllabic, 21 trochaic, and 21 iambic. The full list of items, with by-item results, is in Appendix C. In addition, we constructed 21 [s]-final items as fillers, 7 in each category of monosyllabic, trochaic, and iambic. As we did in French, the items' onsets spanned a wide range of phonotactic wellformedness, from extremely common onsets like [d] and [f] to more uncommon ones like [br] and [dr] (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 unaware of the purpose of the task. The list included each noun in three forms: the singular, a faithful plural, and an unfaithful plural. The recording and processing was done as in §3.1.2.

### 5.1.3 Procedure

The experiment was run in *Experigen* (Becker & Levine 2012), as in §3.1.3. 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 [[a'pɛw] 'hat', which is known to vary in the plural between the standard [[a'pɛws] and the colloquial [[a'pɛjs].

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, asking participants which plural was preferred by use of the scale. Pressing one of the seven buttons moved the participant to the next item. The order of plurals was randomized. The presentation and layout closely followed the procedure for Experiment 1 (see Figure 2), except that the singulars were not presented orthographically at all.

After the participant responded to all 24 target items and fillers, they were asked to answer a few demographic questions, such as their country of origin, age, and so forth.

### 5.1.4 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 9. The raw results are available at <http://becker.phonologist.org/projects/FrenchPortuguese/>.

As for the effect of the vowel that preceded the word-final [w], faithful plurals were chosen most often with the tense vowels [e, o, i, u] (4.3), and unfaithful plurals were chosen most often with the lax mid vowels [a, ɛ, ɔ] (4.8).

To assess the statistical strength of these effects, we used a linear mixed-effects model, with the rating of the alternation as the dependent variable. The predictors are those that were used in the lexicon model in §4.3. The fully crossed model, reported in Table 6,<sup>18</sup> enjoys low collinearity

<sup>18</sup> This is a fully crossed model, with *prosodic shape* and *lax* and their interaction, both as fixed effects and as random slopes given *item* and *participant*. The monosyllabicity and stress predictors were coded exactly as in §4.3. The laxness predictor was coded as  $-.28$  for tense and  $+.72$  for lax, these values ensuring a mean tenseness of zero.

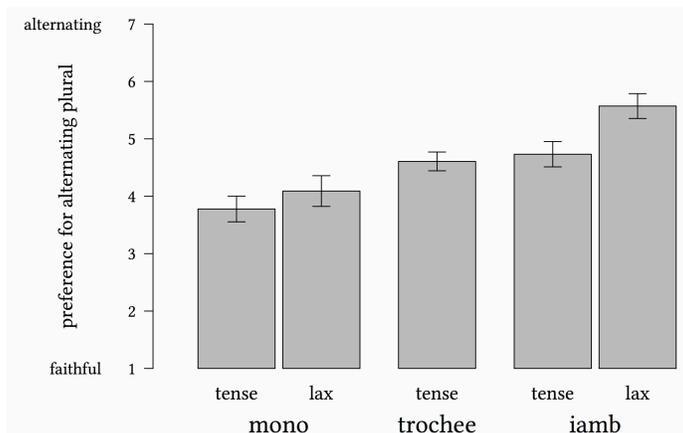


Figure 9: Experiment 4: Monosyllables protected from alternations in Portuguese nonce words ( $n=181$ ).

	$\beta$	SE( $\beta$ )	$t$	$p$ -value <sup>19</sup>
(Intercept)	4.40	.11	39.22	
<i>monosyllabic</i>	-1.03	.18	-5.61	<.001
<i>iamb vs. trochee</i>	-.09	.28	-.32	>.1
<i>lax</i>	.63	.21	2.91	<.005
<i>monosyllabic:lax</i>	-.48	.43	-1.12	>.1

Table 6: Experiment 4: Preference for alternating plurals on a 1–7 scale in Portuguese nonce words. Positive  $\beta$  reflects higher acceptability of the unfaithful plural.

measures ( $\kappa=2.16$ ,  $\text{VIF}\leq 1.39$ ).

The model in Table 6 shows that unfaithful, alternating plurals are highly significantly preferred in polysyllables and dispreferred in monosyllables. Among the polysyllables, there was no difference between iambs and trochees. Alternating plurals were also highly significantly preferred following the lax vowels [a, ε, o], mirroring the distribution in the lexicon. The interaction of *lax* and *monosyllabic* was insignificant; the interaction of *lax* with *iamb vs. trochee* cannot be included in the model, because all trochees have a final tense vowel.

### 5.1.5 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. While Huback (2007) makes a strong case for the effect of token frequency in plural formation for real words of Portuguese, Tang et al. (2013) present model comparisons based on a much larger corpus and demonstrate no advantage of token frequency as a predictor

<sup>19</sup>As in §3.1.4,  $p$ -values were estimated using the normal approximation.

beyond syllable count. Naturally when it comes to experimental results, token frequency makes no prediction for nonce items, which are all equally infrequent and equally unfamiliar to the listener.

## 5.2 Experiment 5: binary judgments

While the results of Experiment 4 were clear, we sought to further confirm them with a more refined methodology, as we did for French. We switch from a 1–7 scale to a binary choice, and concomitantly a logistic model for the analysis. Additionally, we ask speakers to rate the singular before the plurals are judged, to tease apart any effect of the singulars' wellformedness.

### 5.2.1 Participants

The participants were recruited online using word of mouth, and volunteered their time and effort without compensation. We gathered data from 72 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 72 participants took on average 10 minutes to complete the experiment (range 5–24 minutes, median 9). The participants reported an average age of 30 (range 19–62, median 28). According to self-reports, 35 females and 21 males completed the experiment; 16 did not specify.

When asked about the variety of Portuguese they speak, 18 participants listed São Paulo, 17 listed Rio de Janeiro, 15 Minas Gerais, 6 Rio Grande do Sul, and a smaller number of other places in Brazil. 9 participants listed no other language or indicated that they are monolingual. Of those who listed other languages, 39 indicated some knowledge of English, 14 French, 12 English, and a small number of other languages.

### 5.2.2 Materials and procedure

The experiment used the same structure and materials used in Experiment 4, with the only change being the presentation of each item. The presentation followed the same structure as Experiment 2, with the singular base rated first on a scale of 1–5, and then each plural judged as either acceptable or unacceptable with a binary choice. The materials were only presented auditorily, as in Experiment 4, and unlike Experiment 2.

### 5.2.3 Results

The results largely reflect the results of Experiment 4: alternating plurals were accepted least often with monosyllables (53%), and significantly more often with trochees (67%) and iambs (76%). Alternating plurals were accepted most often with the lax vowels [a, ε, ɔ] (70%) and least often with the tense vowels [e, o, i, u] (61%). Looking only at polysyllables with a tense vowel in their final syllable, there is no significant difference between iambs (68%) and trochees (67%).

The statistical analysis was performed using a mixed-effects logistic regression model, with the acceptability of the unfaithful plural as the dependent variable. The predictors were the same ones used in §5.1.4, with the addition of *base*, the acceptability of the singular base on the 1–5 scale. The fully crossed model, reported in Table 7, enjoys low collinearity measures ( $\kappa=2.28$ ,  $\text{VIF}\leq 1.33$ ).<sup>20</sup>

---

<sup>20</sup> The fully crossed model used *prosodic shape* and *lax* and their interaction, as well as *base*, both as fixed effects and as

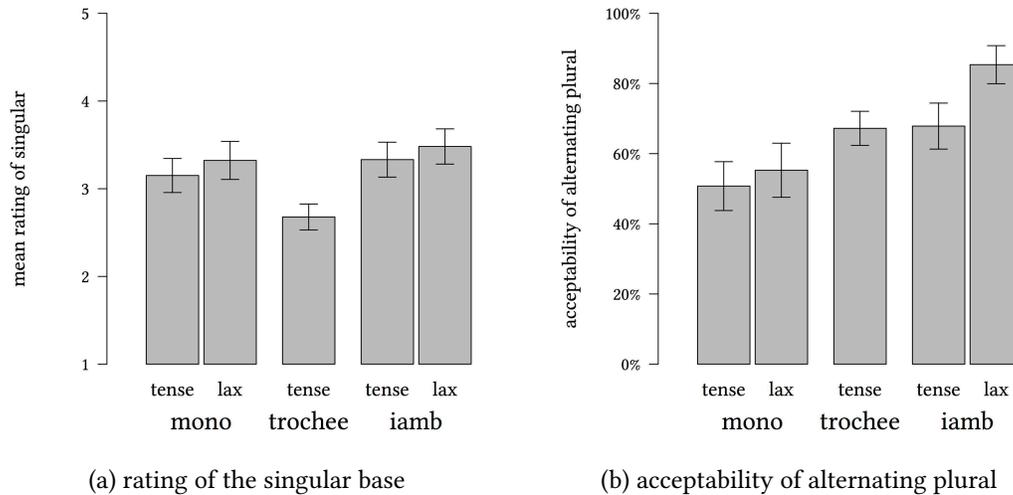


Figure 10: Experiment 5: Monosyllables protected from alternations in Brazilian Portuguese nonce words ( $n=72$ ).

	$\beta$	SE( $\beta$ )	$z$	$p(> z )$
(Intercept)	.78	.14	5.58	
<i>base</i>	.61	.28	2.20	<.05
<i>monosyllabic</i>	-1.32	.22	-6.07	<.001
<i>iamb vs. trochee</i>	.16	.29	.54	>.1
<i>lax</i>	1.03	.27	3.88	<.001
<i>monosyllabic:lax</i>	-1.46	.53	-2.73	<.01

Table 7: Experiment 5: Acceptability of alternating plurals in Portuguese nonce words. Positive  $\beta$  reflects higher acceptability of the unfaithful plural.

Alternations were accepted significantly more often with polysyllables and when the stem’s last vowel was lax. The interaction was significant (the effect of laxness was stronger on polysyllables), unlike in Experiment 4. Alternating plurals were significantly more acceptable when the rating of the base was higher, as seen in French.

### 5.3 Comparison of lexicon with experimental results

The results of Experiment 4 and Experiment 5 both confirm that Brazilian Portuguese speakers protect monosyllabic stems from alternation, and that alternations are more acceptable when the stem’s final vowel is lax. The effect of the lax vowel is stronger in iambs than it is in monosyllables; this interaction came out significant in Experiment 5, and as a trend in the same direction in Experiment 4. In the lexicon, we were unable to estimate the interaction of laxness and monosyllabicity; this is because there is ample evidence for the effect of vowel laxness among the iambs, observable even in the innovative alternating plural ⟨chapéis⟩ [ʃaˈpɛjs] ‘hats’ vs. the lack of such an innovation in ⟨museus⟩ [muˈzɛws] ‘museums’. On the other hand, the small number of

random slopes given *item* and *participant*. The monosyllabicity and stress predictors were coded exactly as in §4.3 and in §5.1.4. The laxness predictor was coded as  $-.30$  for tense and  $+.70$  for lax, these values ensuring a mean tenseness of zero. The rating of the base was scaled using R’s *scale* function dividing by 4, resulting in a range from  $-.53$  to  $+.47$ .

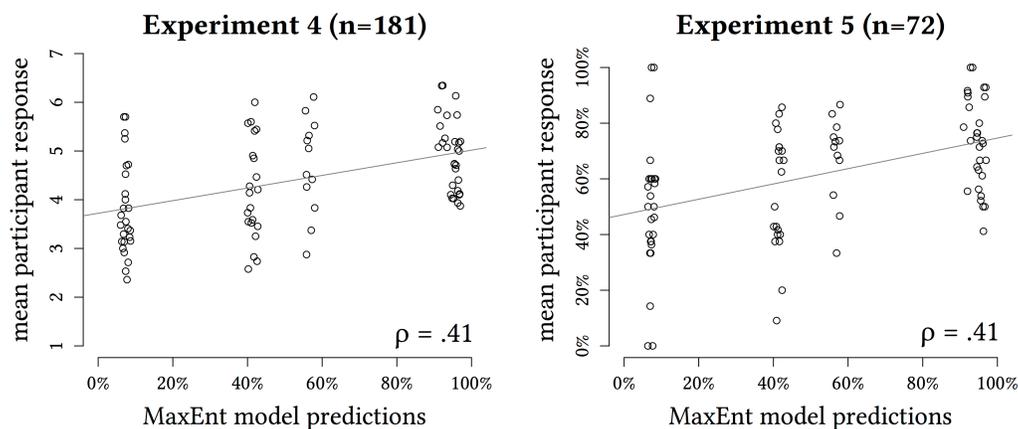


Figure 11: Participants' judgment of alternating plurals of nonce items, plotted by the plurals' probability as predicted by the MaxEnt analysis, with horizontal jitter.

monosyllables provides rather meager evidence for any laxness effect, leading to non-identifiability when attempting to estimate the interaction. It would seem that speakers are somewhat reluctant to fully extend the laxness effect from the iambs to the monosyllables, yielding the observed (somewhat weak) interaction.

Our analysis of the lexicon also predicted that among the polysyllables, alternating plurals would be more acceptable with trochees than with iambs, yet this trend was not extended to nonce words. In our MaxEnt model, the difference between trochees and iambs is modeled using IDENT- $\sigma$ (back), a constraint that protects stressed syllables, and which accrues its large weight by training on established plurals in the lexicon. More broadly in the language, however, we see mixed evidence for protection of stressed syllables: while they display a wider range of vocalic contrasts, stressed syllables are also subject to more nasalization and diphthongization than unstressed syllables. Looking beyond the [w]-final nouns, a study of the plurals of [ẽw̃]-final nouns again suggests that plural alternations impact stressed syllables more than unstressed ones, both in the lexicon (Abaurre 1983) and in preliminary results we have conducted from a nonce word study. Thus, a complete picture of the plural morphology and beyond suggests a limited role for faithfulness to the stressed syllable in Brazilian Portuguese. If indeed faithfulness to backness in the stressed syllable is somewhat limited in the language, this can be incorporated into the MaxEnt analysis by imposing a Bayesian prior on IDENT- $\sigma$ (back).

Figure 11 shows the correlation between the predictions of the MaxEnt grammar in §4 and the participants' choices across both experiments. The correlation is positive and highly significant (Spearman's rank correlation,  $\rho = .41$ ,  $p < .001$  for both experiments). As in French, a sizable proportion of the variation remains unexplained, perhaps suggesting directions for future research investigating factors beyond the strong effects established here for prosodic shape and vowel quality. We turn to a few directions for further investigation in the next section.

## 6 Potential alternative approaches to word-size effects

We have established that monosyllables are repeatedly protected from alternations in French and Brazilian Portuguese, and offered an analysis in terms of initial syllable faithfulness. While the details of implementation differ somewhat 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 both in experimental terms and in terms of synchronic explanation. For example, the items displayed a wide variety of consonant clusters that differ substantially in length, both in the monosyllables and in the polysyllables. In addition, we were careful to represent a wide range of phonotactics in our materials, from extremely common sound combinations up to the least frequently attested sequences in each language.

It may turn out that some of the effects we have observed may have their root in factors that that lie beyond the categorical structural description of phonological words, and in this section we turn to a brief discussion of the potential for phonetically-grounded explanations that rely on raw segmental duration (Barnes 2006; Giavazzi 2010), and lexicon-based explanations that rely on neighborhood density (Ussishkin & Wedel 2009; Stausland Johnsen 2011).

### 6.1 Duration-based approaches

The idea that phonological contrast is based on perceptibility is well-established in the literature (Hayes 1999; Steriade 2001/2008, and many others since). In Optimality Theory, faithfulness constraints work towards the preservation of contrasts, especially in positions of prominence (Smith 2002, 2010). Prominence can be established in several different ways, e.g. intensity of stop burst, pitch peaks, and especially duration. In particular, Barnes (2006) argues that duration is the driving force behind alternations in Turkish and a variety of other languages. Beginning with perception-based accounts of positional prominence, Barnes (2006) maintains that positional asymmetries such as initial syllable faithfulness are not part of Universal Grammar *per se*. Instead, he argues that their explanation is phonetically-grounded, and in particular, grounded in the acoustic cues associated with stress: “I will argue that all patterns of phonological licensing asymmetries attested in initial position are entirely predictable on the basis of the documented phonetic characteristics of those syllables” (Barnes 2006; 165). Similar notions are pursued by Giavazzi (2010), although she focuses on the role of stress, not on the initial syllable. In her discussion of positional faithfulness, Giavazzi (2010; 23) claims that “the grammar of prosodically-conditioned processes is not determined by the special status of privileged positions”, and additionally refers specifically to the central role of duration; asserting “The blocking of neutralization, I claim, does not arise from the grammatical pressure of positional faithfulness constraints, rather it results from the action of metrical constraints on prominence ( $\acute{D}_{dur} \geq X$  and  $\acute{V}_{Energy} \geq X$ ) and from their side-effects” (p. 56). The common theme in both of these works is that protection from alternation is due to phonetic prominence, and in particular increased duration. In a somewhat similar vein, Lehiste (1972) and subsequent work have claimed that in a short word, each individual segment is on average longer than the corresponding segment in a long word, and

thus it stands to reason that short words will resist alternations by virtue of having longer and thus more prominent segments.

To preliminarily test the merits of this line of explanation and assess the role of duration for our materials, we started with measuring the duration of the alternating part and non-alternating part in each item. 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 a predetermined threshold.

In both languages, shorter words have significantly longer alternating portions and longer words have significantly shorter alternating portions: In French (al/aj/ɛl/ɛj) are significantly longer in monosyllables than they are in polysyllables (427ms vs. 369ms, t-test,  $t(73.0) = 4.48$ ,  $p < .001$ ). The same is true in Portuguese, where the stem's final vowel and following [w] are significantly longer in monosyllables (462ms vs. 384ms, t-test,  $t(75.6) = 6.2$ ,  $p < .001$ ). According to Barnes' and Giavazzi's proposals, the increased duration of the alternating portions would be the source of protection from alternation. Crucially, however, the relevant concept of duration does not refer directly to individual tokens, but rather to a more abstract notion of duration, as we demonstrated with Experiment 3. In this experiment, all segmental durations were held constant, yet participants still protected the shorter words from alternation. If the relevant notion of duration is indirect and coarse-grained enough to generalize beyond individual exemplars that may not comply to the pattern, then its grammatical formalization as a difference between monosyllables and polysyllables could arise as one possible outcome of this step of abstraction.

Our binary, monosyllabic-based explanation points in the same direction as these duration-based explanations, and determining whether one explanation is superior to another requires statistical techniques of model comparison. A variety of such techniques are only recently making their way into the language sciences, with a great deal of divergence in the community about what is the most informative such measure – be it likelihood tests, AIC, BIC, DIC, or Random Forests. As consensus on these matters does not yet exist, we have made our materials, results, and phonetic durations available as supplementary, downloadable material at <http://becker.phonologist.org/projects/FrenchPortuguese/>, in hopes that future research can conduct more detailed comparisons.

## 6.2 Lexical similarity-based approaches

We turn to the second source of explanation for the patterns and variation in our experimental results, based on the notion of alternation as regulated by lexical neighborhoods. 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 [tɪp] include [tɪp], [tɪæp], [stɪp], 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* alternations for real words, but no effect for nonce words. Thus, the connection between neighborhood density and alternation rates remains somewhat ill-understood. The lexical status of the stimulus may influence the potential effect of neighborhood density and phonotactic probability on phonological processing. In particular, as Storkel & Morrisette (2002; p.29) remark "Lexical processing is predicted to dominate language tasks involving real words because real words have a lexical representation. In contrast, phonological processing is predicted to dominate language tasks involving nonwords because nonwords have no lexical representation." Moreover, as we have found in a series of simulations, the predictive strength of neighborhood density increases very sharply with the overall lexicon size. This would suggest that if neighborhood density is to be employed by language speakers, it would only become useful after a substantial increase in the overall lexicon, perhaps well after phonological acquisition has solidified. By contrast, a factor such as monosyllabicity is immediately available to language learners regardless of lexicon size.

The specific causal role of neighborhood density on alternation rate, as proposed by Ussishkin & Wedel (2009), has been empirically assessed by Pycha et al. (2007), and later by Becker & Nevins (2009) and Becker et al. (2011), all of whom found that neighborhood density is a rather weak predictor of morphophonological laryngeal alternations in Turkish. In the present section, we assess the relevance of this claim for the alternations at hand 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.

For each of our experimental stimuli, we calculated its neighborhood density and also its log phone 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). Neighbors were defined as existing lexical items that were one segment away from the source word via deletion, addition, or substitution. We calculated these for each singular and plural, making two neighborhood density-based predictors 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 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,  $t(58.1) = 8.3$ ,  $p < .001$ ) and in Portuguese (t-test,  $t(65.8) = 7.03$ ,  $p < .001$ ). Neighborhood density is indeed negatively correlated with alternations in both languages and across all five experiments. 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,  $t(84.4) = 11.0$ ,  $p < .001$ ) and in Portuguese (t-test,  $t(75.4) = 8.5$ ,  $p < .001$ ). Thus, it is quite obviously the case that these lexicon-based measures distinguish short words from long

	all items		monosyllables only		polysyllables only	
Experiment 1	$r = -.49$	$p < .001$	$r = -.28$	$p < .05$	$r = -.61$	$p < .001$
Experiment 2	$r = -.51$	$p < .001$	$r = -.52$	$p < .001$	$r = -.50$	$p < .005$
Experiment 3	$r = -.42$	$p < .001$	$r = -.33$	$p < .05$	$r = -.56$	$p < .001$
Experiment 4	$r = -.45$	$p < .001$	$r = -.46$	$p < .005$	$r = +.19$	$p > .1$
Experiment 5	$r = -.53$	$p < .001$	$r = -.46$	$p < .001$	$r = -.09$	$p > .1$

Table 8: Correlation of neighborhood density with the participants’ choices, by item; overall, separately for monosyllables, and separately for polysyllables.

words; the question is only how well these measures can predict the participants’ responses to the stimuli. Our own comparisons of these predictors versus the initial syllable faithfulness criteria proved inconclusive, as we obtained divergent results for one or the other depending on whether we used ANOVA likelihood tests, AIC, BIC, DIC, or Random Forests, and which random effect structure we used when these models did not converge; given the lack of current consensus, we did not believe it would be helpful to report five sets of inconclusive results at this point. Following a reviewer’s suggestion, we demonstrate the inconclusiveness of the results with a set of Pearson correlation tests, shown in Table 8 for the by-item correlation of neighborhood density with the participants’ choice of alternating plural. We see that when all of the items are considered, i.e., when neighborhood density and monosyllabicity cover the same ground, all five experiments show a strong and significant negative correlation: larger neighborhoods correlate with fewer plural alternations. When considering monosyllables and polysyllables separately, cracks appear: the correlation coefficients for neighborhood density among the monosyllables drop dramatically for two of the three French experiments, while the correlation among the polysyllables loses all predictive power in both Portuguese experiments. Similar inconsistencies appeared with many of the statistical techniques we tried. The overall picture that emerges is that monosyllabicity is a consistently strong element of the analysis, while the contribution of neighborhood density fluctuates. As discussed above, we have made all measurements available on an accompanying website, and we very much hope that future research can specifically examine whether *initial syllable protection* can be subsumed and/or bolstered by lexicon-based measurements such as neighborhood density.

### 6.3 Summary

We have briefly investigated the merits of two alternative approaches to the one which motivated our research question: 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 in both French and Portuguese. We therefore leave open the possibility that correlated “proxy” categories may emerge from the kinds of positive feedback mechanisms explored in Wedel (2007), in which a gradient and variable distribution of patterns may over time become converted into a more categorical pattern.

If indeed gradient factors such as neighborhood density or phonetic duration can underlie and

ultimately trigger the development of a categorical pattern (as in Wedel 2007), interesting questions still remain about the mechanism that triggers the change towards a preference for a binary factor such as monosyllabicity over other, continuous factors, and about the kind of predicted acquisition path given such a theory. It is our hope that the results reported here can contribute towards resolving such fundamental questions in acquisition and grammaticalization, in conjunction with interdisciplinary modeling efforts.

## 7 General discussion and conclusions

In this paper, we offered two case studies of plural alternations, in French and Portuguese, based on five experiments. 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. 2011). Fundamental to the proposal was the claim that monosyllabicity is a salient property of words, and readily exploited by learners. Our results here from French and Portuguese join similar results from Turkish, where again monosyllables resist an stem-final alternation that applies more liberally to polysyllables (Becker et al. 2011). These three languages can be equally construed as protecting either initial syllables or monosyllables.

Becker et al. (2012) report that the opposite pattern is found in the English lexicon, where the plural stem-final alternation (e.g. *leaf* ~ *leaves*) impacts monosyllables more than polysyllables among existing words. In a nonce word task, English speakers impacted monosyllables and polysyllables equally, showing that the monosyllabicity effect is asymmetric: the grammar cannot productively subject monosyllables to alternations more strongly than polysyllables. Indeed, when given an unbiased artificial language, English speakers act exactly like French/Portuguese/Turkish speakers, preferring to protect monosyllables. To show that initial syllable faithfulness offers protection beyond the realm of monosyllables, Becker et al. (2012) also tested English speakers on an artificial language where all words were disyllabic, and alternations impact the word's first or second vowel. The speakers protected the initial syllable in these cases, suggesting that the protection of monosyllables should in fact be understood as a special case of initial syllable protection.

Our experimental results in this paper show that speakers of both languages attended to the monosyllabicity of the stem and the quality of its last vowel, and replicated these lexical trends, which were also the most reliable factors in their lexicons. Two weaker lexical trends were not replicated: In French, speakers did not replicate the greater deletability of [l], and in Portuguese, speakers did not prefer alternations in unstressed syllables. We suggested that speakers were biased against assigning a sufficiently large weight to two relevant faithfulness constraints, one against the deletion of glides (in French) and one protecting stressed syllables (in Portuguese). Similar mismatches have been discussed by Becker et al. (2011, 2012); Hayes et al. (2009); Hayes & White (2013), and may ultimately inform a broader theory of the biases that speakers bring to the task of learning their language.

Our analysis in terms of initial syllable faithfulness is based on a generative approach that seeks to account for phonological patterns using discrete formal mechanisms. Historically, such

frameworks focused on entirely regular phenomena, as in Prince & Smolensky (1993/2004); we deployed a theory augmented with the ability to track lexical statistics over irregular morphophonology using MaxEnt grammars (Goldwater & Johnson 2003; Hayes & Wilson 2008; White 2014). The gradient patterning of novel items is derived from a discrete mechanism that assigns a probability distribution over outputs based on a binary monosyllabicity criterion.

The underlying effect of initial syllable faithfulness is potentially universal (perhaps even innate, although the evidence from our two test-cases does not speak to that directly). It is possible that the positional faithfulness mechanism can be induced by furnishing the speaker with 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 asymmetric protection of initial syllables to the table, regardless of language-specific input.

In conclusion, the formalization in terms 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. The micro-patterns of initial syllables' treatment in the irregular plurals of French and Portuguese are understood here to reflect typological macro-patterns, as well as salient cognitive universals related to word-edges in phonotactic learning (Endress & Mehler 2010). While humans may have the ability to track rich swaths of multidimensional continuous data, we see here the explanatory advantage afforded by a maximally simple discrete understanding of morphophonological lexical pattern in terms of syllable count. Given the consistent results across our five experiments, we contend that the overall picture is one in which learners formalize generalizations that can be governed at the syllable level even when they regulate segmental alternations.

## Appendix A French lexicon

The French lexicon used in §2, with Part of Speech (N=noun, A=adjective) and status of plural (0=fairful, .5=variable, 1=alternating). The items are listed by monosyllabicity and final consonant.

item	PoS	plural	item	PoS	plural	item	PoS	plural
aj	N	.5	anal	A	0	mōdjal	A	1
baj	N	1	āsestkal	A	1	moval	N	1
bɔkaj	N	0	ostkal	A	.5	myval	A	1
ɔkaj	N	0	otonal	A	1	naɔval	N	0
taj	A	0	aval	N	0	nazal	A	1
taj	N	0	banal	A	.5	nazal	N	1
bal	N	0	bākal	A	0	natal	A	0
kal	N	0	bestjal	A	1	naval	A	0
fal	N	0	bokal	N	1	oval	A	1
gval	N	0	boɔeal	A	.5	oval	N	0
mal	A	1	kaɔnaval	N	0	papal	A	1
mal	N	1	sātikal	A	1	paɔbal	N	0
mal	N	0	jakal	N	0	paɔsjal	A	0
pwal	N	0	ʃanal	N	1	paskal	A	.5
ɔal	N	0	ʃəval	N	1	paskal	N	.5
val	N	0	koval	A	0	postal	A	1
atɔkaj	N	.5	koval	N	0	paɔsjal	A	1
bɔkaj	N	0	dɔksal	A	1	ɔesital	N	0
bōzaj	N	0	dɔal	A	0	ɔektal	A	1
kamaj	N	0	egal	A	1	ɔegal	N	0
ʃādaj	N	0	fatal	A	0	ɔezjonal	A	1
kobaj	N	0	feodal	A	1	ɔezjonal	N	1
koɔkaj	N	1	feodal	N	1	ɔenal	A	1
detaj	N	0	festival	N	0	ɔival	N	1
emaj	N	1	filjal	A	1	ɔwajal	A	1
eruvātaj	N	0	final	A	1	ɔɔval	A	1
evātaj	N	0	final	N	.5	ɔɔval	N	1
fwaɔkaj	N	0	fraktal	A	0	ɔinal	N	1
guɔɔɔkaj	N	1	frōtal	A	1	ɔosjal	N	1
pɔɔtaj	N	0	frōtal	N	1	ɔpasjal	A	1
samɔɔkaj	N	0	glasjal	A	0	ɔspesjal	A	1
supɔɔkaj	N	1	zɔɔnal	N	1	ɔspektkal	A	.5
tɔɔkaj	N	1	zovjal	A	1	ɔstɔɔnal	A	1
vātaj	N	1	legal	A	1	ɔtɔɔmal	A	1
vitɔkaj	N	1	lokal	A	1	ɔtōbal	A	0
			lokal	N	1	ɔtɔɔbal	A	1
			lwajal	A	1	ɔtɔɔvjal	A	1
			maɔsjal	A	1	ɔvagal	A	1
			mātal	A	1	ɔvasal	N	1
			mātal	N	1	ɔvenal	A	.5
			metal	N	1			
			mistkal	N	0			

## Appendix B French nonce words

Items with the average response they were given: in Experiment 1 ( $n=185$ ) on the scale of 1–7 (higher score = stronger preference for alternating plurals), and in Experiment 2 ( $n=52$ ) & 3 ( $n=71$ ), the proportion of choice of alternating plurals.

Alternating plurals end in [ø] for [ej]-final items and [o] otherwise.

[l]	Exp. 1	Exp. 2	Exp. 3	[j]	Exp. 1	Exp. 2	Exp. 3
dal	3.60	0.42	0.44	daj	n/a	n/a	n/a
fal	4.00	0.14	0.50	faj	2.93	0.25	0.50
fʁal	4.26	0.64	0.57	fʁaj	5.29	0.67	1.00
ɡnal	5.21	0.62	0.76	ɡnaj	5.50	0.67	0.25
kʁal	3.71	0.40	0.55	kʁaj	3.50	0.00	0.60
pnal	4.55	0.60	0.70	pnaj	4.71	0.60	0.17
pʁal	3.79	0.50	0.56	pʁaj	3.33	1.00	0.00
psal	4.84	0.67	0.55	psaj	4.07	1.00	0.25
sal	3.25	0.14	0.30	saj	4.50	0.00	0.20
skʁal	3.74	0.33	0.56	skʁaj	4.09	0.25	0.67
snal	3.46	0.57	0.40	snaj	4.73	0.50	0.40
spʁal	3.71	0.43	0.29	spʁaj	3.38	n/a	0.67
stʁal	4.44	0.56	0.60	stʁaj	5.14	1.00	1.00
sval	3.07	0.36	0.54	svaj	2.87	0.25	0.60
tʁal	3.89	0.33	0.31	tʁaj	3.17	1.00	0.00
vʁal	4.15	0.50	0.44	vʁaj	4.60	1.00	0.43
zal	3.35	0.22	0.46	zaj	2.86	0.50	0.29
zʁal	4.19	0.29	0.30	zʁaj	4.08	0.75	0.33
deʁnal	4.92	0.47	0.88	deʁnaj	4.86	1.00	0.50
ɡuval	5.32	0.73	0.68	ɡuvaj	5.00	1.00	0.50
ɡzovʁal	4.60	0.67	0.55	ɡzovʁaj	5.90	0.50	1.00
istʁal	4.29	0.70	0.56	istʁaj	5.85	0.50	0.75
mɔʁnal	5.70	0.78	0.92	mɔʁnaj	4.71	0.50	0.67
oʁsodal	5.41	0.73	0.78	oʁsodaj	4.15	0.50	1.00
oskʁal	4.71	0.44	0.60	oskʁaj	4.60	0.50	0.62
peʁetal	5.28	1.00	0.69	peʁetaj	4.56	0.50	0.45
pʁykal	5.17	0.78	0.61	pʁykej	5.10	0.80	0.50
segal	4.14	0.75	0.50	segaj	4.35	0.83	0.40
skʁənal	5.32	0.75	0.62	skʁənaj	4.50	0.50	0.75
smɔnal	5.17	0.56	0.75	smɔnaj	3.93	0.33	0.60
smʁyʁkal	4.74	0.36	0.47	smʁyʁkaj	3.83	1.00	1.00
vøzal	5.40	0.73	0.73	vøzaj	5.82	0.00	n/a
zistal	5.07	1.00	1.00	zistaj	4.42	0.67	0.17
dɛl	2.21	0.00	0.00	dɛj	2.72	0.20	0.38
dʁɛl	2.24	0.33	0.14	dʁɛj	2.35	0.25	0.50
kɛl	2.55	0.71	0.27	kɛj	4.05	0.38	0.13
kʁɛl	2.79	0.50	0.00	kʁɛj	2.31	0.33	0.00
nɛl	1.90	0.14	0.00	nɛj	2.95	0.00	0.00
pʁɛl	2.52	0.20	0.00	pʁɛj	1.91	0.50	0.17
psɛl	2.58	0.33	0.50	psɛj	3.32	0.22	0.17
ʁɛl	2.55	0.25	0.25	ʁɛj	3.48	0.00	0.11
zɛl	1.62	0.14	0.09	zɛj	2.20	0.12	0.25
fanel	4.42	0.00	0.00	fanej	2.30	0.00	0.00
fakasel	2.79	0.33	0.33	fakasej	3.25	0.38	0.00
ɡʁasel	2.42	0.20	0.30	ɡʁasej	2.57	0.12	0.38
paʁtel	4.04	0.38	0.15	paʁtej	1.87	0.17	0.00
ʁafɛl	2.74	0.64	0.00	ʁafɛj	3.50	0.25	0.20
ʁomɛl	2.16	0.25	0.00	ʁomɛj	2.85	0.17	0.33
ʃotel	2.42	0.20	0.20	ʃotej	3.21	0.00	0.08
zyvɛl	3.35	0.50	0.33	zyvɛj	2.65	0.43	0.00

## Appendix C Brazilian Portuguese nonce words

Items with the mean response per item. In Experiment 4 ( $n=181$ ), the response is on the 1–7 scale, where 1 is a faithful plural (e.g. [ˈbraw ~ ˈbraws]) and 7 is an alternating plural (e.g. [ˈbraw ~ ˈbrajs]). In Experiment 5 ( $n=72$ ), 0 is a faithful plural and 1 is an alternating plural.

	Exp. 4	Exp. 5		Exp. 4	Exp. 5
ˈbraw	3.59	0.38	ˈtahtew	5.74	0.76
ˈdaw	4.47	0.40	ˈfastow	4.40	0.50
ˈfaw	2.58	0.42	ˈzapew	5.04	0.71
ˈfraw	2.74	0.43	ˈfeskow	4.10	0.56
ˈpraw	2.83	0.09	ˈpespow	4.74	0.52
ˈtraw	3.45	0.67	ˈvehpew	6.13	0.93
ˈbɛw	3.73	0.86	ˈkozew	5.20	0.89
ˈbrɛw	4.14	0.71	ˈtɔmew	5.19	0.75
ˈkɛw	5.44	0.62	ˈzɔfow	3.87	0.64
ˈprɛw	3.55	0.70	ˈhefew	5.17	0.73
ˈpɛw	5.41	0.50	ˈkeskew	4.73	0.76
ˈtrɛw	3.52	0.83	ˈtehtow	4.18	0.67
ˈbɔw	3.25	0.40	ˈhɔsew	5.00	0.93
ˈbrɔw	6.00	0.71	ˈkɔspow	4.02	0.50
ˈdrɔw	4.21	0.78	ˈsohkew	4.10	0.80
ˈkɔw	5.60	0.43	ˈsikow	4.12	0.54
ˈpɔw	3.83	0.20	ˈzɪpew	4.03	0.41
ˈtrɔw	4.28	0.38	ˈʃibow	3.93	0.61
ˈvɔw	4.90	0.67	ˈduhnow	4.64	0.63
ˈgɔw	4.85	0.80	ˈhuntew	4.71	0.67
ˈzɔw	5.57	0.70	ˈtumow	4.30	0.74
ˈfɛw	2.53	0.58	biˈpaw	5.85	0.74
ˈfrɛw	2.36	0.54	kuˈtaw	5.51	0.89
ˈkɛw	2.71	0.50	maˈhaw	5.17	0.56
ˈnɛw	2.92	0.38	buˈtɛw	5.73	0.91
ˈprɛw	3.15	0.33	paˈmɛw	6.34	1.00
ˈgrɛw	3.29	0.14	prɪˈzɛw	5.26	0.86
ˈzɛw	3.37	0.60	kuˈpɔw	5.08	0.79
ˈbrɔw	3.41	0.00	niˈsɔw	6.34	0.92
ˈfɔw	3.48	0.50	gahˈtɔw	5.08	1.00
ˈkrɔw	3.00	0.38	haˈsɛw	3.83	0.54
ˈmɔw	4.00	0.40	suˈpɛw	3.37	0.33
ˈnɔw	3.23	0.00	dʒiˈmɛw	2.88	0.47
ˈtɔw	3.82	0.46	vaˈrɔw	5.83	0.75
ˈdriw	5.70	0.67	ziˈbɔw	5.21	0.67
ˈkiw	3.83	0.36	zuˈgɔw	5.05	0.71
ˈniw	4.52	0.60	duˈziw	4.26	0.74
ˈpriw	3.55	0.45	ʃaˈmiw	5.32	0.79
ˈtriw	3.14	0.33	ʃiˈviw	5.52	0.83
ˈziw	4.70	1.00	tuˈzuw	6.11	0.87
ˈbrɔw	5.70	1.00	zaˈnuw	4.42	0.68
ˈdruw	4.72	0.89	ʃiˈruw	4.51	0.73
ˈfuw	3.14	0.60			
ˈmuw	3.68	0.57			
ˈpruw	5.25	0.60			
ˈtuw	5.37	0.40			
ˈvuw	4.12	0.60			

## References

- Abaurre, Bernadete (1983). Alguns casos de formação de plural em português: uma abordagem natural. *Cadernos de Estudos Lingüísticos* 5. 127–156.
- Albright, Adam & Bruce Hayes (2003). Rules vs. Analogy in English past tenses: a computational/experimental study. *Cognition* 90. 119–161.
- Baayen, R. Harald (2008). *Analyzing Linguistic Data: A practical introduction to statistics*. Cambridge University Press.
- Barnes, Jonathan (2006). *Strength and Weakness at the Interface: Positional Neutralization in Phonetics and Phonology*. Berlin/New York: Mouton de Gruyter.
- Barr, Dale J., Roger Levy, Christoph Scheepers & Harry J. Tily (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68. 255–278.
- Bates, Douglas & Martin Maechler (2013). *lme4: Linear mixed-effects models using Eigen and Eigen*. R package version 0.999999-4.
- Becker, Michael (2009). *Phonological Trends in the Lexicon: The Role of Constraints*. Ph.D. dissertation, University of Massachusetts Amherst.
- Becker, Michael & Maria Gouskova (to appear). Source-oriented generalizations as grammar inference in Russian vowel deletion. *Linguistic Inquiry* Available as lingbuzz/001622.
- Becker, Michael, Nihan Ketrez & Andrew Nevins (2011). The surfeit of the stimulus: Analytic biases filter lexical statistics in Turkish laryngeal alternations. *Language* 87. 84–125.
- Becker, Michael & Jonathan Levine (2012). 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 (2012). Asymmetries in generalizing alternations to and from initial syllables. *Language* 88. 231–268.
- 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.
- Bennett, William (1997). Problematic plural forms in French. In Raymond Hickey & Stanislaw Puppel (eds.) *Language History and Linguistic Modelling, Vol II*, Mouton de Gruyter. 1363–1370.
- Boersma, Paul & David Weenink (2015). 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.
- Bybee, Joan & Carol Lynn Moder (1983). Morphological classes as natural categories. *Language* 2. 251–270.
- Bybee, Joan & Dan Slobin (1982). Rules and schemas in the development and use of the english past tense. *Language* 58. 265–289.
- 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. 2685–2710.
- Daland, Robert, Bruce Hayes, Marc Garellek, James White, Andrea Davis & Ingrid Normann (2011). Explaining sonority projection effects. *Phonology* 28. 197–234.
- de Kolovrat, Georges (1923). *Etude sur la vocalisation de la consonne l dans les langues romanes*. Paris: Jouve et cie.
- Dutoit, T., V. Pagel, N. Pierret, F. Bataille & O. van der Vrecken (2006). The Mbrola Project: Towards A Set Of High Quality Speech Synthesizers Free Of Use For Non Commercial Purposes. In

- Proceedings of the fourth international conference on spoken language processing*. Philadelphia, vol. 3, 1393–1396.
- 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.
- 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.
- Goldwater, Sharon & Mark Johnson (2003). Learning OT constraint rankings using a maximum entropy model. In Jennifer Spenader, Anders Eriksson & Osten Dahl (eds.) *Proceedings of the Stockholm Workshop on Variation within Optimality Theory*. 111–120.
- 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. Moravcsik, 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.
- Hayes, Bruce & James White (2013). Phonological naturalness and phonotactic learning. *Linguistic Inquiry* 44. 45–75.
- Hayes, Bruce & Colin Wilson (2008). A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39. 379–440.
- Hayes, Bruce, Colin Wilson & Anne Shisko (2012). Maxent grammars for the metrics of Shakespeare and Milton. *Language* 88. 691–731.
- Hayes, Bruce, Kie Zuraw, Péter Siptár & Zsuzsa Londe (2009). Natural and unnatural constraints in Hungarian vowel harmony. *Language* 85. 822–863.
- Huback, Ana Paula da Silva (2007). *Efeitos de frequência nas representações mentais*. Ph.D. dissertation, Universidade Federal de Minas Gerais, Belo Horizonte.
- Jaeger, Florian (2013). LSA summer institute class on mixed effect regression. Access at [wiki.bcs.rochester.edu/HlpLab/LSA2013Regression](http://wiki.bcs.rochester.edu/HlpLab/LSA2013Regression).
- Jesney, Karen (2009). Positional faithfulness, non-locality, and the harmonic serialism solution. *Proceedings of the 39th Meeting of the North East Linguistic Society*.
- Kapatsinski, Vsevolod (2009). *The architecture of grammar in artificial grammar learning: Formal biases in the acquisition of morphophonology and the nature of the learning task*. Ph.D. dissertation, Indiana University.
- Kapatsinski, Vsevolod (2011). What statistics do learners track? rules, constraints or schemas in (artificial) grammar learning. In Stefan Th. Gries & Dagmar Divjak (eds.) *Frequency effects in language: Learning and processing*, Mouton de Gruyter. 53–82.
- Kirk, Cecilia & Katherine Demuth (2006). Accounting for variability in 2-year-olds' production of coda consonants. *Language Learning and Development* 2. 97–118.
- Lehiste, Ilse (1972). The timing of utterances and linguistic boundaries. *Journal of the Acoustical Society of America*. 2018–2024.
- Luce, Paul A. & D.B. Pisoni (1998). Recognizing spoken words: The Neighborhood Activation Model. *Ear and Hearing* 19. 1–36.
- Mattoso Câmara, Joaquim (1953). *Para o estudo da Fonêmica Portuguesa*. Rio de Janeiro: Simões.
- McCarthy, John J. (2003). Comparative markedness. *Theoretical Linguistics* 29. 1–51.
- Nevins, Andrew (2012). Vowel lenition and fortition in Brazilian Portuguese. *Letras de Hoje* 47. 228–233.
- 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, Michael O’Keefe, Ehren Reilly & Adam Werle (eds.) *Papers in Optimality Theory III*, Amherst, MA: GLSA. 259–296.
- Pater, Joe (2008a). Gradual learning and convergence. *Linguistic Inquiry* 39. 334–345.
- Pater, Joe (2008b). Morpheme-specific phonology: Constraint indexation and inconsistency resolution. In Steve Parker (ed.) *Phonological Argumentation: Essays on Evidence and Motivation*, London: Equinox. 123–154.
- Pope, Mildred Katharine (1952). *From Latin to Modern French*. London: Butler & Tanner.
- 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 (2013). R: A language and environment for statistical computing. Vienna, Austria. (<http://www.r-project.org>).
- 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. 2439–2463.
- 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.
- Storkel, Holly L. & Michele L. Morrisette (2002). The lexicon and phonology: Interactions in language acquisition. *Language, Speech, and Hearing Services in Schools* 33. 24–37.
- Tang, Kevin, Andrew Nevins & Michael Becker (2013). Prosody drives alternations: evidence from a 51 Million Word Corpus of Brazilian Portuguese. Poster presented at the Linguistic Society of America Meeting, Boston MA.
- 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.
- Wedel, Andrew B. (2007). Feedback and regularity in the lexicon. *Phonology* 24. 147–185.
- White, James (2014). Evidence for a learning bias against saltatory phonological alternations. *Cognition* 130. 96–115.
- Wilson, Colin (2006). Learning phonology with substantive bias: an experimental and computational study of velar palatalization. *Cognitive Science* 30. 945–982.
- 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.
- Zuraw, Kie (2013). \*MAP constraints. Ms. UCLA.