The Surfeit of the Stimulus: Analytic biases filter lexical statistics in Turkish devoicing neutralization[*]

Michael Becker	Nihan Ketrez	Andrew Nevins
Reed College	Yale University	Harvard University

November 22, 2008

Abstract

Some sublexical statistical regularities of Turkish phonotactics are productively extended in nonce words, while others are not. In particular, while stop-voicing alternation rates in the lexicon can be predicted by the place of articulation of the stem-final stop, by word-length, and by the preceding vowel quality, this stop-voicing alternation is only productively conditioned by place of articulation and word-length. Speakers' responses in forced-choice and production tasks demonstrate that although they are attuned to the place of articulation and size effects, they ignore preceding vowels, even though the lexicon contains this information in abundance. We interpret this finding as evidence that speakers distinguish between phonologically-motivated generalizations and accidental generalizations. We propose that Universal Grammar, a set of analytic biases, acts as a filter on the generalizations that humans can make: UG contains information about possible and impossible interactions between phonological elements. Omnivorous statistical models that do not have information about possible interactions incorrectly reproduce accidental generalizations, thus failing to model speakers' behavior.

---

[*] For their valuable comments and suggestions on several versions of this work, we are greatly indebted to Adam Albright, Nicoleta Bateman, Kathryn Flack, Lyn Frazier, Bruce Hayes, William Idsardi, Karen Jesney, Shigeto Kawahara, Michael Key, John Kingston, Joe Pater, Kathryn Pruitt, John McCarthy, Harker Rhodes, Engin Sezer, Jesse Snedeker, Donca Steriade, Anne-Michelle Tessier, Bert Vaux, Andy Wedel, Matt Wolf, and Kie Zuraw. We would also like to thank the audience at the 2004 HUMDRUM, especially Luigi Burzio, Jane Grimshaw, Alan Prince and Ellen Woolford, the audience at the Yale 2006 Turkish Linguistics Workshop, especially İlhan Çağrı, Atakan İnce, and Jaklin Kornfilt, the audience at the 81st LSA annual meeting in Anaheim, CA, especially Sharon Inkelas and Eduardo Gil, the Harvard phonotactics working group, especially Iris Berent, Peter Graff, and Tim O'Donnell, and the audience at the University of Toronto linguistics department, especially Peter Avery, Elan Dresher, Yoonjung Kang, Alexei Kochetov, and Naomi Nagy. Thanks to the many speakers of Turkish who helped with various aspects of this project, including Mine Doğucu, Anıl and Alper Duman, Beste Kamali, Süleyman Ulutaş, Leyla Zidani-Eroğlu, several students at Boğaziçi university, and the many kind people who filled our questionnaires. Thanks also to the children and parents at the Yeni Çilek and Ayışığı kindergartens in Levent, İstanbul. Any remaining errors would be ours.

1. Introduction

Learners and language users can and often do use statistical properties of linguistic input to discover hidden structure and make predictive generalizations about newly-encountered items (e.g. Coleman & Pierrehumbert (1997); Bailey & Hahn (2001); see Saffran (2003), Hay & Baayen (2005), Chater & Manning (2006) for recent overviews). While these abilities to track statistical regularities in the input appear to be very powerful, at the same time they also appear to be constrained: some patterns are more readily detected and used than others. For example, Bonatti et al. (2005) found that adult learners exposed to artificial grammars were much better at extracting transitional probability regularites over consonants than equally matched transitional probabilites over vowels, suggesting that learners preferentially pay more attention to statistics within consonantal frames. In a study of infant learning of phonotactic patterns, Saffran & Thiessen (2003) showed that infants learned statistical patterns that grouped together /p/, /t/, /k/ (i.e. voiceless stops) as a class of items comprising the first sound in artificial word tokens much better than patterns that grouped /p/, /d/, /k/ as this class, again suggesting that statistical learning may be less efficient when the regularities are inconsistent with natural language structure.

In this paper, we examine a number of predictive statistical phonotactic regularities found within the Turkish lexicon, some natural and some unnatural from the point of view of phonological typology, and examine whether they are all kept track of and used to an equal extent in on-line judgement tasks involving novel words. By examining whether adult speakers of a language with robust statistical regularities will detect and extend the use of unnatural patterns in generalization tasks, we can provide potential evidence for the role of analytic biases as active filters on extraction of sublexical statistics.

Voicing alternations in Turkish are observed at the right edges of nouns, as in (1). Nouns that end in voiceless aspirated stop in their bare form, such as the pre-palatal stop [t͡ʃʰ], can either retain that [t͡ʃʰ] in the possessive (1a-b), or the [t͡ʃʰ] of the bare stem may alternate with the voiced [d͡ʒ] in the possessive (1c-d).[1]

---

[1] Turkish orthography doesn't represent aspiration, as it is predictable from a combination of voicing and morphological structure. For a discussion of laryngeal features in Turkish, see §4.1.

(1)

	bare stem	possessive	
a.	at͡ʃʰ	at͡ʃʰ-ɨ	‘hunger’
b.	anat͡ʃʰ	anat͡ʃʰ-ɨ	‘female cub’
c.	tat͡ʃʰ	tad͡ʒ-ɨ	‘crown’
d.	amat͡ʃʰ	amad͡ʒ-ɨ	‘target’

Turkish exhibits a contrast between the voiced stops [b, d, d͡ʒ, ɡ] and the voiceless aspirated stops [pʰ, tʰ, t͡ʃʰ, kʰ] in onset position, e.g. *tʰer* ‘sweat’ vs. *der* ‘give-aorist’. In coda position, however, the contrast is lost, with stops appearing voiceless and aspirated through complete phonetic neutralization (Kopkallı 1993; Wilson 2003). This restriction on the distribution of voiced stops applies productively to loanwords, e.g. *ropʰ* ‘dress’ < French *robe*. Voiced coda stops are allowed in the initial syllable of the word, e.g. *ad* ‘name’ or *abla* ‘older sister’, and in a limited number of exceptional words.

When nouns that end in a voiceless stop are follwed by a vowel-initial suffix, the final stop may surface with its voiced counterpart, e.g. *d͡ʒopʰ* ‘club’ vs. the possessed form *d͡ʒo**b**-u* ‘club.3SG’; however, when followed by a consonant-initial suffix, the final stop remains in coda position and thus stays voiceless: *d͡ʒopʰ-lar* ‘club.plural’. This alternation occurs in 54% of the nouns of the language (Inkelas et al. 2000), and applies productively to loanwords, e.g. *gurupʰ* vs. *gurub-u* ‘group.3SG’. For the remaining 46% of stop-final nouns, the stop is voiceless in all suffixed forms of the word, e.g. *sopʰ* ∼ *sopʰ-u* ‘clan.3SG’ ∼ *sopʰ-lar* ‘clan.plural’.

The velar stops [kʰ,ɡ] contrast in onset position, e.g. *so.**k**ʰakʰ* ‘street’ vs. *ga.**g**a* ‘beak’. In word-final position, they neutralize to the voiceless stop [kʰ]. While post-consonantal dorals, as in *renkʰ* ∼ *ren**g**-i* ‘color’, display the general process of voicing alternation, intervocalic velar stops undergo deletion rather than voicing, i.e. when nouns ending in *postvocalic* velar stop are followed a vowel-initial suffix, the velar stop deletes e.g. *ete**k**ʰ* ∼ *ete-i* ‘skirt’ (Zimmer & Abbott (1978), Sezer (1981)). Since voicing alternation and deletion are in complementary distribution, depending on the segment that precedes the final dorsal, we treat the two processes as one. Additionally, as will be shown below, whether a noun stem shows the k/∅ alternation or not is correlated with the same type of sublexical statistics as other stop consonant alternations, thereby justifying a unified treatment for the purpose of the current experimental inquiry.

The distinction between alternating and non-alternating stops is traditionally captured within generative phonology as the difference between an underlying voiced stem-final stop in the case of *d͡ʒopʰ* ∼ *d͡ʒob-u* and an underlying voiceless stem-final stop in the case of *sopʰ* ∼ *sopʰ-u*, with the underlying contrast being neutralized in word-final coda position (Lees 1961). While the difference between alternating and non-alternating nouns may be captured in a variety of alternate theoretical frameworks which do not incorporate the possibility of underlying representations (e.g. via reference to identity-relations vs. lack thereof among surface forms alone (Burzio 2002; Albright 2008), it is clear that under any way of representing morphophonemic alternation, Turkish nouns fall into two distinct classes of words, one of which alternates and one of which doesn't.

Whether the final stop of a given noun will or will not alternate is unpredictable. However, the noun's size strongly correlates with its status: Most monosyllabic nouns do not alternate, while most poly-syllabic nouns do. Section §2 discusses several other factors that correlate with voicing alternations, and shows that Turkish speakers use only a subset of the available factors: They use the noun's size and the place of articulation of the final stop, but they do not use the quality of the vowel that precedes the word-final stop. A back vowel before a word-final [t͡ʃʰ], for instance, correlates with more alternations, but Turkish speakers seem to ignore this correlation. This language-specific pattern can be understood given a cross-linguistic perspective: Typological observations commonly correlate the distribution of voice with a word's size and a consonant's place of articulation, but rarely or never with the quality of a neighboring vowel. Indeed, speakers are reluctant to learn patterns that correlate vowel height with the voicing of a neighboring consonant (Moreton 2008).

From a cross-linguistic perspective, it is unsurprising that mono-syllabic nouns would behave differently from poly-syllabic nouns with respect to the voicing alternation. Initial syllables are often protected from markedness pressures, showing a wider range of contrasts and an immunity to alternations (Beckman 1998). Specifically in Turkish, the privileged status of the laryngeal features [voice] and [s.g.] in initial syllables is not only seen in voicing alternations. Generally in the language, a coda stop followed by an onset stop will surface with the laryngeal features of the onset stop (e.g. *is.tʰi**b**.dat* 'despotism', **is.tʰi**pʰ**.dat*), but a coda stop in the initial syllable may surface with its independent laryngeal specification (e.g. *ma**kʰ**.bul* 'accepted', *e**b**.kʰem* 'mute').

The backness of a neighboring vowel, however, is never seen to interact with a consonant's voicing. While such a connection is mildly phonetically plausible (vowel backness correlates with tongue-root position, which in turn correlates with voicing), there is no known report of any language where consonant voicing changes depending on the backness of a neighboring vowel, or vice versa. Given this gap in the universal inventory of possible phonological interactions, it is no longer surprising that in Turkish, speakers show no sign of using vowel backness as a predictor of voicing alternations.

In Optimality Theory (Prince & Smolensky 1993/2004), typological observations are encoded in the structure of the universal inventory of constraints (CON). The constraints and their interactions produce all and only the observed sound patterns of the world's languages. The preferred status of initial syllables is encoded with a set of faithfulness constraints specific to initial syllables. The lack of interaction between vowel backness and voicing is encoded by the exclusion of constraints from CON that refer to some value of [±back] next to some value of [±voice], e.g. *[+back][+voice]. In the absence of such constraints, there is never a reason to change one of these features in the presence of the other, and the lack of interaction is predicted. The account of the Turkish facts offered here capitalizes on these aspects of CON, while remaining agnostic about the mechanism that excludes these constraints, be it by assuming an innate set of constraints (as has been assumed since Prince & Smolensky 1993/2004, and in the context of learning by Tesar & Smolensky 1998, 2000; Tesar 1998; Prince 2002; Hayes 2004; Jarosz 2006; Tesar & Prince 2006 among others), or by a mechanism of constraint induction (as in Hayes & Wilson 2008, Flack 2007).

We propose a version of Optimality Theory where the pattern of individual lexical items is recorded in terms of lexically-specific constraint rankings (cf. Pater 2006, 2008; Anttila 2002; Inkelas et al. 1997; Itô & Mester 1995; Coetzee 2008). A noun with a non-alternating final stop, like *anat͡ʃʰ* ∼ *anat͡ʃʰ-ɨ*, is associated with the ranking IDENT(lar) $\gg$ *Vt͡ʃV, meaning that faithfulness to laryngeal features outweighs the markedness pressure against voiceless intervocalic palatal stops. A noun with a final alternating stop, like *amat͡ʃʰ* ∼ *amad͡ʒ-ɨ*, is associated with the opposite ranking, i.e. *Vt͡ʃV $\gg$ IDENT(lar). This assumes that the final stop in *amat͡ʃʰ* is underlyingly voiceless and unaspirated, and that it surfaces unfaithfully in *amad͡ʒ-ɨ*, contrary to the traditional generative analysis of Turkish (Lees 1961; Inkelas & Orgun 1995; Inkelas et al. 1997). This aspect of the analysis is discussed and motivated in §4.

Given this approach, the pattern of mono-syllabic nouns, like *at͡ʃʰ* ∼ *ad͡ʒ-ɨ*, can be recorded separately from the pattern of poly-syllabic nouns, by using a faithfulness constraint that protects the laryngeal features of stops in the base's initial syllable, IDENT(lar)$_{\sigma1}$. The existence of constraints in CON that are specific to initial syllables allows Turkish speakers to learn separate lexical trends for monosyllabic and polysyllabic nouns. On the other hand, in the absence of universal constraints that relate laryngeal features and vowel backness, the backness of the stem-final vowel cannot be used in recording the pattern of any lexical items, and this aspect of the lexicon goes ignored by speakers.

To encode lexically-specific constraint rankings, the version of Optimality Theory used here is one augmented by a mechanism of constraint cloning (Pater 2006, 2008). In this theory, language learners detect that their language requires opposite rankings of a pair of constraints, and then clone one of those constraints. In the Turkish case, speakers realize that some lexical items require IDENT(lar) $\gg$ *Vt͡ʃV and some lexical items require the opposite ranking. They clone one of the constraints, say IDENT(lar), and then non-alternating nouns are associated with the clone of IDENT(lar) that ranks over *Vt͡ʃV, and alternating nouns are associated with the clone that ranks under *Vt͡ʃV.

The resulting grammar contains two lists of nouns, as every *t͡ʃ*-final noun of Turkish is listed under one of the clones of IDENT(lar). Since most *t͡ʃ*-final nouns do alternate, most nouns will be listed with the clone that ranks below *Vt͡ʃV. Now suppose a speaker encounters a novel noun in its bare form, and they are required to produce the possessive form. The grammar allows the final stop to either alternate or not alternate, but the alternating pattern is more likely, since more nouns are listed with the clone of IDENT(lar) that ranks below *Vt͡ʃV. Cloned constraints allow speakers to reach a grammar that records the pattern of known items, and then project that pattern probabilistically onto novel items.

The full analysis of Turkish will involve the general faithfulness constraint IDENT(lar) and the more specific IDENT(lar)$_{\sigma 1}$, to protect final stops from becoming voiced, and additionally MAX and MAX$_{\sigma 1}$, to protect final dorsals from deleting (see §5.6). These faithfulness constraints conflict with a series of markedness constraints against voiceless stops, either between two vowels (*VpV, *VtV, *Vt͡ʃV, *VkV) or between a sonorant consonant and a vowel (*RpV, *RtV, *Rt͡ʃV, *RkV). Each stop-final noun of Turkish is listed under a pair of conflicting constraints, or equivalently, each pair of conflicting constraints accumulates a list of lexical items, and this listing allows the speaker to project the lexical statistics onto novel nouns.

Speakers' ability to project trends from their lexicon onto novel items is a well-established observation (see Zuraw 2000, Albright et al. 2001, Ernestus & Baayen 2003, Hayes & Londe 2006, among many others). The novel observation offered here, that only Universal trends are projected, does find support in previous work. In a study of voicing alternations in Dutch, Ernestus & Baayen (2003) show that speakers project the rate of alternation of different stops based on their place of articulation, just like the Turkish speakers. Ernestus & Baayen's (2003) report of the vowel effects is instructive: In the lexicon, stops alternate more following long vowels and less after short vowels. Following the high vowels of Dutch, which are all short, stops have an intermediate rate of alternation. In their experiment, however, speakers projected and strengthened the vowel length effect, preferring more alternations after long vowels. Speakers did not project the vowel height effect, choosing alternations equally frequently after short vowels that are either high or non-high. Given our proposal, this result is not surprising: As mentioned above, vowel height is universally not expected to interact with voicing. The preference for longer vowels before voiced consonants, however, is well-attested (Denes 1955; Peterson & Lehiste 1960; Chen 1970, among others). The absence of observed lengthening before voiced consonants in some languages lends support to the view that the lengthening is controlled by the grammar in terms of durational specifications (Keating 1985; Buder & Stoel-Gammon 2002), and thus can enter into speakers' learning of lexical trends.

The theoretical contribution of this work is two-fold: (a) It relates the projection of language-specific lexical trends to cross-linguistic patterns of phonological interactions, by deriving both from the inventory of universal constraints in CON, and (b) it offers an OT-based grammar that applies deterministically to known items, and projects lexical trends directly from those items onto novel nouns.

2. Turkish lexicon study

The distribution of voicing alternations in the lexicon of Turkish depends heavily on the phonological shape of nouns. For instance, while the final stop in most mono-syllabic nouns does not alternate (2a), the final stop in most poly-syllabic words does alternate with its voiced counterpart (2b). This section offers a detailed quantitative survey of the Turkish lexicon, based on information from the Turkish Electronic Living Lexicon (TELL, Inkelas et al. 2000).

(2)

	Bare stem	Possessive	
a.	at͡ʃʰ	at͡ʃʰ-ɨ	‘hunger’
b.	amat͡ʃʰ	amad͡ʒ-ɨ	‘target’

Several phonological properties of Turkish nouns will be discussed, showing that four of them correlate with stem-final alternations: (a) the noun’s size (mono-syllabic vs. polysyllabic), (b) the place of articulation of the stem-final stop, (c) the height of the vowel that precedes the stem-final stop, and (d) the backness of that vowel.

TELL lists a total of about 30,000 nouns, verbs, and adjectives. Nouns are listed with their bare citation forms and with four suffixed forms (1.SG possessive, accusative, professional, and 1.SG predicative). While the entries were collected from a variety of extant dictionaries, the listed forms were produced and transcribed by a native speaker.

Of the 3002 nouns in TELL whose bare stem ends in a voiceless stop, almost 90% are polysyllabic, and in most of those, the final stop alternates[2] (3). The rate of alternation is much lower for monosyllables, especially in those with a simplex coda.[3]

(3)

Size	n	% alternating
Monosyllabic, simplex coda (CVC)	137	11.7%
Monosyllabic, complex coda (CVCC)	164	25.9%
Polysyllabic (CVCVC and bigger)	2701	58.9%

The distribution of alternating stops also varies by the place of articulation of the word-final stop (4). Most word-final labials, palatals and dorsals[4] do alternate, but only a small proportion of the final coronals do.

---

[2] Some nouns in TELL are listed as both alternators and non-alternators. In calculating the percentage of alternating nouns, such nouns were counted as half alternators (although in reality it’s entirely possible that the actual rate of alternation is different from 50%). Therefore, the proportion of alternating nouns is calculated by adding the number of alternating nouns and half the number of vacillating nouns, and dividing the sum by the total number of nouns.

[3] Our discussion of alternation rates, here and throughout the paper, is based on type frequncies. Since we did not have access to lexical statistics in Turkish, we cannot confirm that there are no effects of token frequencies. It is, however, a well-established observation that novel word tasks are sensitive to the types in the lexicon, and ignore token frequencies (Bybee 1995; Albright & Hayes 2002; Hay et al. 2004).

[4] Dorsals delete post-vocalically, see §5.6 for discussion.

(4)

Place	n	% alternating
Labial (p)	294	84.0%
Coronal (t)	1255	17.1%
Palatal (t͡ʃ)	191	60.5%
Dorsal (k)	1262	84.9%

While longer words correlate with a higher proportion of alternating nouns, size does not affect all places equally (5). In all places, CVC words alternate less than CVCVC words, but the pattern of CVCC words is not uniform. For labials and palatals, a majority of CVCC words alternate, patterning with the CVCVC words. For the dorsals, the CVCC words pattern together with the shorter CVC words, showing a modest proportion of alternators. Finally, the coronals show a very minor place effect, with CVCC words actually having a slightly higher proportion of alternators than either longer or shorter words.

(5)

	CVC		CVCC		CVCVC	
Place	n	% alt	n	% alt	n	% alt
p	30	26.7%	16	75.0%	248	91.5%
t	41	6.1%	79	19.0%	1135	17.3%
t͡ʃ	23	17.4%	18	58.3%	150	67.3%
k	43	3.5%	51	9.8%	1168	91.2%

In other words, it is not the case that size and place each have a constant effect. Their effect on the distribution of voicing alternations cannot be accurately described separately. Anticipating the discussion in §3.2, it will be seen that indeed speakers treat each place/size combination separately.

Further study of TELL reveals a correlation between the quality of the vowel that precedes the word-final stop and the proportion of alternating nouns: high vowels correlate with a higher proportion of alternating stops relative to non-high vowels, and so do back vowels relative to front vowels. This correlation is rather surprising, since cross-linguistically, vowel quality is not known to influence the voicing of a neighboring obstruent[5].

---

[5] Vowel length does correlate with voicing, with long vowels correlating universally with voiced consonants and short vowels with voiceless consonants (Lisker & Abramson 1964; Ohala 1983; Volatis & Miller 1992).

A noun-final stop is about 30% more likely to alternate when following a high vowel than when following a non-high vowel (6).

(6)

Height of stem-final vowel	n	% alternating
−high	1690	41.7%
+high	1312	71.9%

The correlation with height, however, is not equally distributed among the different size and place combinations. The table in (7) shows that in most size/place combinations, there are only modest differences (less than 10%) between the proportions of alternating nouns given the height of the preceding vowel. A larger correlation in the opposite direction (53%) is seen for the CVCC *t͡ʃ*-final words, but this is limited to a mere 18 nouns, which explains its negligible impact on the overall size correlation. The correlation with height is concentrated at the longer *t*-final nouns, where several hundred nouns show 24% more alternating stops following a high vowel.

(7)

	CVC		CVCC		CVCVC	
	−high	+high	−high	+high	−high	+high
p	19	11	13	3	132	116
	26%	27%	77%	67%	85%	99%
t	24	17	55	24	796	339
	10%	0%	15%	29%	10%	34%
t͡ʃ	14	9	8	10	91	59
	18%	17%	88%	35%	66%	69%
k	31	12	33	18	474	694
	2%	8%	12%	6%	87%	94%

A fourth and final phonological property that significantly correlates with the distribution of voicing alternations is the backness of the stem-final vowel (8). When preceded by a back vowel, a stem-final stop is about 10% more likely to alternate compared to a stop preceded by a front vowel.

In some cases, such as that of Canadian Raising, the change in vowel length causes a concomitant change in vowel quality. See §5.2 below for discussion.

(8)

Backness of stem-final vowel	n	% alternating
−back	1495	49.5%
+back	1507	60.3%

Just like vowel height, the correlation with vowel backness is not uniformly distributed in the lexicon. As seen in (9), the correlation with backness is small (at most 13%) for labial-, coronal- and dorsal-final nouns. A robust correlation with backness is seen in *t͡ʃ*-final words of all sizes. Averaged over the 191 *t͡ʃ*-final nouns, the proportion of alternating nouns is 30% higher following a back vowel relative to a front vowel.

(9)

	CVC		CVCC		CVCVC	
	−back	+back	−back	+back	−back	+back
p	12	18	4	12	113	135
	33%	22%	75%	75%	96%	87%
t	18	23	34	45	673	462
	8%	4%	26%	13%	16%	19%
t͡ʃ	11	12	10	8	66	84
	14%	21%	40%	81%	50%	81%
k	19	24	25	26	510	658
	8%	0%	16%	4%	90%	92%

In contrast to the four properties that were examined until now (size, place, height and backness), a phonological property that has but a negligible correlation with the distribution of voicing alternations is the rounding of the stem's final vowel (10).

(10)

Rounding of stem-final vowel	n	% alternating
−round	2524	54.6%
+round	478	56.4%

A closer examination of vowel rounding is no more revealing, and the details are omitted here for lack of interest. Other phonological properties that were checked and found to be equally unrevealing are the voicing features of consonants earlier in the word, such as the closest consonant to the root-final stop, the closest onset consonant, and the closest obstruent.

To sum up the discussion so far, four phonological properties of Turkish nouns were seen to correlate with stem-final voicing alternations in Turkish:

- Size: mono-syllables alternate less than poly-syllables, and among the mono-syllables, roots with simplex codas alternate less than roots with complex codas.
- Place (of articulation): Stem-final coronals alternate the least, while labials and dorsals alternate the most.
- Vowel height: stem-final stops are more likely to alternate following a high vowel compared to a non-high vowel.
- Vowel backness: stem-final stops are more likely to alternate following a back vowel compared to a front vowel.

All of these properties allow deeper insight when considered in pairs: Size and place have a non-uniform interaction, with CVCC words behaving like CVC words when dorsal-final and like CVCVC words when labial- or palatal-final. Height and backness interact with place non-uniformly: the correlation with height is concentrated in the coronal-final nouns, while the correlation with backness is concentrated in the palatal-final nouns.

In statistical parlance, the aforementioned properties can be understood as predictors in a regression analysis. Since TELL makes a three-way distinction in stop-final nouns (nouns that don't alternate, nouns that do, and "vacillators"', i.e. nouns that allow either alternation or non-alternation), an ordinal logistic regression model was fitted to the lexicon using the *lrm()* function in R (R Development Core Team 2007). The dependent variable was a three-level ordered factor, with non-alternation as the lowest level, alternation as the highest level, and vacillation as the intermediate level.

Five independent variables were considered:

- Size: a three-level unordered factor, with levels corresponding to mono-syllables with a simplex coda (CVC), mono-syllables with a complex codas (CVCC), and poly-syllables (CVCVC). CVC was chosen as the base level.[6]

---

[6] We have also considered a less linguistically-informed size variable that was a simple raw count of the

- Place: a four-level unordered factor, with levels corresponding to coronal, palatal, labial and dorsal. Dorsal was chosen as the base level.
- High, back and round: each of the three features of the stem-final vowel was encoded as two-level unordered factor. The base levels chosen were non-high, front and unrounded.

First, each of these five predictors was tried in its own model, to assess each predictor's overall power in the lexicon (11). This power is measured by R^2 and by the model's likelihood ratio (Model L.R.), which comes with a number of degrees of freedom and a p-value. It turns out that *place*, *high*, *size*, and *back* are highly predictive of alternations, in that order, and *round* isn't[7].

(11)

	R^2	Model L.R.	df	p
place	.482	1469	3	$<$.001
high	.113	284	1	$<$.001
size	.078	193	2	$<$.001
back	.015	37	1	$<$.001
round	0	0	1	.489

While *high* has a larger R^2 than *size*, the interaction of *high* and *place* is less powerful than the interaction of *size* and *place*. The interaction of *place* with each of *size*, *high*, and *back* were tested in separate models, summarized in (12).

---

syllables of the stem. This variable was less informative than our size variable, producing lower R^2 and higher p values, so we excluded it from the following presentation. One reason raw size is less informative is that alternation rates don't keep going up as the word gets longer, but rather peak with di- and tri-syllables at 64% and 61% respectively, then go down to 40% and 41% for the tetra- and penta-syllables. The difference between the di- and tri-syllables is not significant generally, and only barely reaches significance for the labials ($p =$ .03). The difference between the tri- and tetra-syllables is significant only without place factored in – once the place variable is added, the difference goes away. The vowel effects that we report below come out essentially the same with either size variable.

[7] Another method for assessing the predictive power of each feature separately is a TiMBL simulation (Daelemans et al. 2002). Given the data in TELL, this system creates a number called "information gain" for every predictor that it is given. The system confirmed the verdict in (11), assigning the five predictors the following information gain values, respectively: .367, .071, .047, .009 and .0004.

(12)

	R^2	Model L.R.	df	p
place*size	.588	1920	11	$<$.001
place*high	.519	1621	7	$<$.001
place*back	.488	1496	7	$<$.001

When a base model that has *place*size* as a predictor is augmented with *place*high*, R^2 goes up to .616. Augmenting the base model with *place*back* only brings R^2 up to .594. Finally, model with all three of the interactions in (12) as predictors reaches an R^2 of .622, with a model L.R. of 2078 for 19 degrees of freedom. This final model is given in (13)[8].

The model in (13) contains few surprises, as it confirms the validity of the observations made earlier in this section. It simply restates the numerical observations as differences in the propensity to alternate relative to the arbitrarily chosen baseline levels of the predictors, namely CVC size, dorsal place, non-high vowels and front vowels. The size effect is mostly limited to the difference between CVC and CVCVC, with none of the CVCC levels reaching significance relative to CVC. In the CVCVC size, the coronal and palatal places alternate significantly less than the baseline dorsal, and labial place only approaches significance at this size. The vowel features reach significance for the interaction of high and coronal, and for the interaction of back and palatal.

---

[8] The model in (13) was validated with the fast backwards step-down method of the *validate()* function, and the predictor *back* was the only one deleted. Since the interaction of *back* with *place* was retained, we did not remove *back* from the model, so as not to leave an interaction in the model without its components. In 200 bootstrap runs, seven factors were considered: the three interaction factors, and the four basic factors they were made of. At least 5 of the 7 factors were retained in 197 of the runs, and in the vast majority of the runs, the three interaction factors were among the ones retained. The R^2 of the model was adjusted slightly from .6213 to .6117. An additional step of model criticism was taken with the *pentrace()* function, which penalizes large coefficients. With a penalty of .3, The penalized model was left essentially unchanged from the original model in (13), with slight improvements of the p-values of the vowel-place interactions at the fourth decimal place.

(13)

	Coefficient	SE	Wald z	p
(y>=vacillator)	−3.502	0.745	−4.70	>0.001
(y>=alternating)	−3.822	0.746	−5.13	>0.001
COR	−0.102	0.976	−0.10	0.917
LAB	2.201	0.954	2.31	0.021
PAL	1.249	0.950	1.31	0.189
CVCC	0.783	0.869	0.90	0.367
CVCVC	5.488	0.735	7.47	0.000
high	0.874	0.205	4.27	0.000
back	0.288	0.204	1.41	0.158
CVCC:COR	0.703	1.102	0.64	0.523
CVCC:LAB	2.022	1.157	1.75	0.081
CVCC:PAL	1.269	1.129	1.12	0.261
CVCVC:COR	−4.011	0.959	−4.18	>0.001
CVCVC:LAB	−1.737	0.901	−1.93	0.054
CVCVC:PAL	−3.110	0.919	−3.38	0.001
COR:high	0.620	0.254	2.45	0.014
LAB:high	0.533	0.539	0.99	0.323
PAL:high	−0.754	0.387	−1.95	0.051
COR:back	0.077	0.254	0.30	0.762
LAB:back	−0.755	0.490	−1.54	0.123
PAL:back	1.136	0.386	2.95	0.003

The quantitative analysis of the proportions of alternating nouns, in the form of a logistic regression, revealed four factors that are predictive of whether voicing alternation will occur: the phonological *size* of the word, the *place* of articulation, the *height* of the preceding vowel, and the *backness* of the preceding vowel. The first two of these have been previously identified as having an influence on voicing alternation in Turkish (Inkelas & Orgun 1995; Inkelas et al. 1997), and indeed the first two of these, from a crosslinguistic perspective are more likely than the other two to have a causal relationship with stop voicing.

One characterization of different types of phonotactics makes a distinction between first-order and second-order phonotactics (Warker & Dell 2006): first-order phonotactics regulate the distribution of a particular (set of) phonological feature(s) within a particular position in a syllable or word, whereas second-order phonotactics relate the distribution of a phonological feature in a particular position to some *other* property of the syllable or word, such as a feature of a neighboring segment. While it is not the case that across the board, first-order phonotactics are more widespread than second-order (for example, vowel harmony is a second-order phonotactic), with respect to the case at hand, namely the distribution of voicing in stops, it is generally the case that only first-order phonotactics matter.

The phonological size of a word, as measured here, is a proxy for a fact about the location of the potentially alternating stem-final stop: whether it occurs in the *initial syllable* of the word or not. Indeed, as mentioned in the discussion of Turkish phonotactics above, one notorious locus of exceptions to otherwise persistent coda devoicing is in the coda of the initial syllable, as evidenced by words such as *ad* 'name' and *abla* 'older sister'. This resistance to alternations in monosyllabic words is a result of the fact that in monosyllabic words, the stem-final syllable *is* the initial syllable. As a consequence, in a word such as *so**p**^{h}-u* 'clan' (as opposed to *guru**b**-u* 'group') the fact that the stop does not alternate is precisely because of a general resistance to alternations for segments in the initial syllable. Cross-linguistically, initial syllables enjoy greater faithfulness, or resistance to alternation (Beckman 1998). The *size* variable is thus a first-order phonotactic, as it relates the occurrence of particular features (voicing and aspiration) to a particular position in the word (the initial syllable).

The effect of the place of articulation on a stop that potentially undergoes alternation has crosslinguistic support as well. Different places are known to interact differently with laryngeal features (Lisker & Abramson 1964; Ohala 1983; Volatis & Miller 1992), and different relative proportions of alternation rates for different places of articulation were found by Ernestus & Baayen (2003) in their study of the Dutch lexicon. While the relative ranking of alternation rates across places of articulation may differ from language to language, it is a fact that languages exhibit phonotactics in manner features and laryngeal features that are gradient and differential specifically depending on place of articulation. The *place* variable is thus a first-order phonotactic, as it relates the occurrence of a particular set of features (voicing, aspiration, and place).

The effect within the Turkish lexicon of vowel quality (in particular, height and backness) on consonant voicing alternation is, on the other hand, unexpected given crosslinguistic phonological typology. Interactions between vowel timber (height, backness, rounding) and the laryngeal features of consonants are infrequent, and the handful of documented cases show a causal influence in the opposite direction: the consonant's laryngeal features can affect the height of a preceding vowel (Kingston 2002), but not vice versa. Consonant voicing and aspiration have been argued to affect vowel height in various languages (e.g. in diphthong centralization before voiceless consonants in North American dialects of English, known as "Canadian Raising" (Chambers 1973; Moreton & Thomas 2007); in Polish (Gussmann 1980); in Madurese (Stevens 1968) and vowel backness in Northern Sarawak (Blust 2000), but there is no documented case of a phonological process wherein vowel quality induces a change in consonant voicing or aspiration. Given the fact that interactions of vowel quality and consonantal laryngeal features are second-order phonotactics with little to no crosslinguistic attestation, their existence in Turkish is expected to be accidental rather than principled in nature.

These data therefore raise the question of whether Turkish speakers themselves will take the correlation between vowel quality and consonant voicing to be accidental or, whether they will take it to reflect an active generalization over their lexicon that they will reproduce. Given that all four of the factors of *size*, *place*, *high* and *back* are statistically reliable predictors of voicing alternations in the lexicon, we sought to determine whether speakers actually track and extend these patterns in experimental tasks with novel words.[9]

To summarize the study of the Turkish lexicon, it was found that both *size* and *place* are excellent predictors of the alternation status of nouns. Larger nouns are more likely to alternate, and coronal-final nouns are less likely to alternate. In addition, the *height* and *backness* of final stem vowels are also good predictors in combination with place: High vowels promote the alternation of coronals, and back vowels promote the alternation of palatals. All of these generalizations were confirmed to be highly statistically significant in a logistic regression model. In other words, the size of nouns, the place of their final stop, and the height and backness of their final vowels all strongly correlate with voicing alternations in a way that is statistically unlikely to be accidental.

---

[9] Our study assumes that TELL is a good model of the lexica of our speakers. The native speaker who supplied the judgments for TELL is about fifty years older than the average participant in our experiment, but they share a comparably high level of education and socio-economic background. Voicing alternations are known to vary with socio-economic levels, but not with age. Additionally, the *validate()* function that we applied to the model in (13) assures that the effects of the predictors are strong and reliable even in lexica that are different from TELL by as much as 37%. We conclude that we have little reason to doubt the usefulness of comparing the TELL data with data from highly educated younger speakers.

3. Speakers' knowledge of the lexicon

In the previous section, the distribution of voicing alternations in the Turkish lexicon was examined and shown to be rather skewed. The distribution of alternating and non-alternating noun-final stops is not uniform relative to other phonological properties that nouns have: *size*, *place*, *height*, and *backness* were identified as statistically powerful predictors of alternation.

What the humans who are native speakers of Turkish know about the distribution of voicing alternations, however, is a separate question, which is taken on in this section. It will turn out that native speakers identify generalizations about the distribution of voicing alternations relative to the *size* of nouns and the *place* of articulation of their final stops. However, speakers ignore, or fail to reproduce, correlations between the voicing of final stops and the quality of the vowels that precede them.

A novel word task (Berko 1958) was used to find out which statistical generalizations native speakers extract from their lexicon. This kind of task has been shown to elicit responses that, when averaged over several speakers, replicate distributional facts about the lexicon (e.g. Zuraw 2000 and many others).

3.1. Materials and method

3.1.1. Speakers

Participants were adult native speakers of Turkish ($n = 24$; 13 males, 11 females, age range: 18-45) living in the United States. Some of the speakers were paid $5 for their time, and others volunteered their time. The experiment was delivered as a web questionnaire, with some speakers doing the experiment remotely. For those speakers, reaction times were indicative of the speakers taking the questionnaire in one sitting, with no discernible distractions or pauses.

3.1.2. Materials

A native speaker of Turkish (male, mid-20s) recorded the bare form and two possible possessive forms for each noun, repeated three times. Each stimulus was normalized for peak intensity and pitch and inspected by a native speaker to be natural and acceptable. One of the possessive forms was completely faithful to the base, with the addition of a final high vowel that harmonized with the stem, following the regular vowel harmony principles of the language. In the other possessive form, the stem final stop was substituted with its voiced counterpart, except for post-vocalic *k*'s, which were deleted.

Creating stimuli that exemplify all size, place and vowel quality combinations would have come up to 96 (four places * three sizes * eight vowel qualities). Since the lexical distribution of voicing alternations among palatals and labials is fairly similar, and in the interest of reducing the number of trials, the palatal and labial categories were collapsed into one category, using 12 words of each place, compared to 24 for the coronal- and dorsal-final words. The total number of stimuli, then, was 72 (three place categories * three sizes * eight vowel qualities).

Additionally, native Turkish nouns disallow the round nonhigh vowels $\{o, ø\}$ in non-initial position. To make the stimuli more Turkish sounding, non-high round vowels in the second syllable of the CVCVC words were replaced with the corresponding high vowels $\{u, y\}$. The nouns that were used are presented in (14).

The non-final consonants were chosen such that the resulting nouns all sounded plausibly native, with neighborhood densities equalized among the stimuli as much as possible.[10]

---

[10] To evaluate our choice of test items, we made a post-hoc comparison between the items' neighborhood density and the experimental results. We concluded that neighborhood density did not have any measurable effect on speakers' behavior. We should note that by necessity, longer items have lower neighborhood density, and since the participants preferred more alternations with longer items, neighborhood density was negatively correlated with our experimental results ($r(70) = -.363$, $p < .005$). However, the correlation between neighborhood density and alternation rates is mediated by size, and indeed, size is the better predictor of alternation rates: Adding neighborhood density as a predictor into the analysis in (18) made no noticeable change, as confirmed by an ANOVA model comparison ($\chi^2(1) = .260$, $p > .1$).

(14)

			CVC		CVCC		CVCVC	
			−high	+high	−high	+high	−high	+high
p/t͡ʃ	−round	−back	gep^h	yit͡ʃ^h	t^{h}elp^h	gint͡ʃ^h	hevet͡ʃ^h	d͡ʒisip^h
		+back	dap^h	nɨt͡ʃ^h	p^{h}ant͡ʃ^h	dɨrp^h	yɨyap^h	ma.ɨt͡ʃ^h
	+round	−back	k^{h}öt͡ʃ^h	züp^h	yönt͡ʃ^h	k^{h}ürp^h		bölüt͡ʃ^h t^{h}ürüt͡ʃ^h
		+back	p^{h}ot͡ʃ^h	t^{h}up^h	solp^h	munt͡ʃ^h		k^{h}onup^h guyup^h
t	−round	−back	p^{h}et^h	hit^h	zelt^h	t͡ʃ^{h}int^h	nik^{h}et^h	gevit^h
		+back	fat^h	mɨt^h	hant^h	ʃɨrt^h	ya.at^h	p^{h}ɨsɨt^h
	+round	−back	söt^h	d͡ʒüt^h	gönt^h	nürt^h		sölüt^h bünüt^h
		+back	yot^h	nut^h	d͡ʒolt^h	bunt^h		t͡ʃ^{h}orut^h muyut^h
k	−round	−back	vek^h	zik^h	helk^h	t^{h}ink^h	mesek^h	p^{h}erik^h
		+back	d͡ʒak^h	p^{h}ɨk^h	vank^h	nɨrk^h	t^{h}at^{h}ak^h	banɨk^h
	+round	−back	hök^h	sük^h	sönk^h	p^{h}ürk^h		nönük^h düyük^h
		+back	mok^h	nuk^h	bolk^h	dunk^h		zoruk^h yuluk^h

Finally, 36 fillers were included. All the fillers ended in either fricatives or sonorant consonants. To give speakers a meaningful task to perform with the fillers, two lexically-specific processes of Turkish were chosen: vowel-length alternations (e.g. *ruh* ∼ *ruːh-u* ‘spirit’) and vowel-∅ alternations (e.g. *burun* ∼ *burn-u* ‘nose’). Eighteen fillers displayed vowel-length alternations with a CVC base, and the other eighteen displayed vowel-∅ alternations with a CVCVC base. All of the fillers were chosen from a dictionary of Turkish, some of them being very familiar words, and some being obsolete words that were not familiar to the speakers we consulted.

The materials were recorded in a sound attenuated booth into a Macintosh computer at a 44.1 KHz sampling rate. Using Praat (Boersma & Weenink 2008), the token judged best of each suffixed form was spliced and normalized for peak intensity and pitch. Peak intensity was normalized using Praat's "scale peak" function set to 0.6. For pitch normalization, three points were manually labeled in each affixed form: the onset of the word, the onset of the root's final segment (the onset of the burst in the case of stops), and the offset of the word. Then, a reversed V-shaped pitch contour was superimposed on the materials, with a pitch of 110 Hz at the onset of the word, 170 Hz at the onset of the root-final segment, and 70 Hz at the offset of the word. These values were chosen in order to best fit most of the speaker's actual productions, such that changes would be minimal.

Finally, for each stimulus, two .wav files were created by concatenating the two suffixed forms with a 0.8-second silence between the two, once with the voiceless form followed by the voiced form, and once with the voiced followed by the voiceless. A linguist who is a native speaker of Turkish verified that the final materials were of satisfactory quality. While she had some concerns about stress being perceived non-finally in a few of the filler items, no problems were found with the stimuli.

3.1.3. Procedure

Before the beginning of the experiment, speakers were reminded that voicing alternations are lexically-specific by presenting a familiar non-alternating paradigm (*t^{h}op^h* ∼ *t^{h}op^h-u* 'ball') next to a familiar alternating paradigm (*d͡ʒep^h* ∼ *d͡ʒeb-i* 'pocket'). Then, speakers were asked to choose the possessive form of two familiar alternating nouns (*dolap^h* 'cupboard' and *a.at͡ʃ^h* 'tree'), and feedback was given on their choices.

The stimuli were presented in a self-paced forced-choice task. The base form (e.g. *fet^h*) was presented in Turkish orthography (e.g. ⟨fet⟩) which reflects the relevant aspects of the phonology faithfully. The participants saw an overt possessor with genitive case followed by a blank, to provide the syntactic context for a possessive suffix, e.g. *Ali'nin* ________ "Ali's ________", and they heard two possible possessed forms, e.g. *fe**t**^{h}-i* and *fe**d**-i*. Speakers pressed "F" or "J" to choose the first or the second possessive form they heard. Most speakers took 15-20 minutes to complete the experiment.

The order of the stimuli and the order of the choices were randomized. Additionally, the fillers were randomly distributed among the first three quarters of the stimuli.

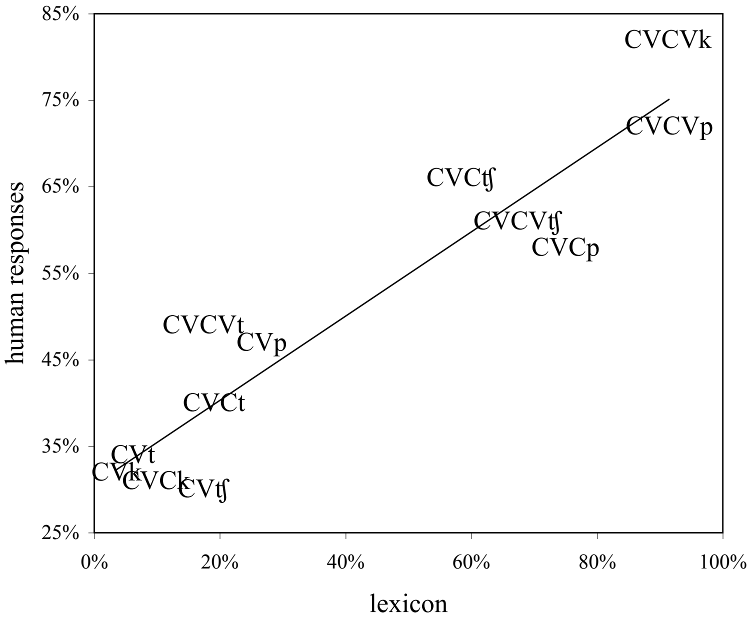
3.2. Results

The experimental results are plotted in (15), grouped by size and place, plotted against the percent of alternating words in the lexicon with the matching size and place. The correlation is excellent (Spearman's rank correlation test, $S = 46$, $\rho = .839$, $p < .005$), showing that speakers have accurately matched the percentages of alternating words in the lexicon. On average, the proportion of alternating responses ranges from 30% to 82%, as opposed to a wider range of 6% to 92% in the lexicon. Nevertheless, this compressed range of responses correlates with the lexicon very well.[11]

---

[11] The source of the compression of the human results comes both from between-speaker and within-speaker sources. Some participants showed a strong preference for alternating responses, and some showed the opposite preference, resulting in at least 3 and at most 22 alternating responses per item, thus covering only 79% of the range of 0 to 24 alternating responses possible with 24 participants. Additionally, individual participants varied as to how strong the size and place effects were in their responses, with weak-effect participants causing further compression. The strength of these effects did not correlate with participants' overall preference for alternation or non-alternation.

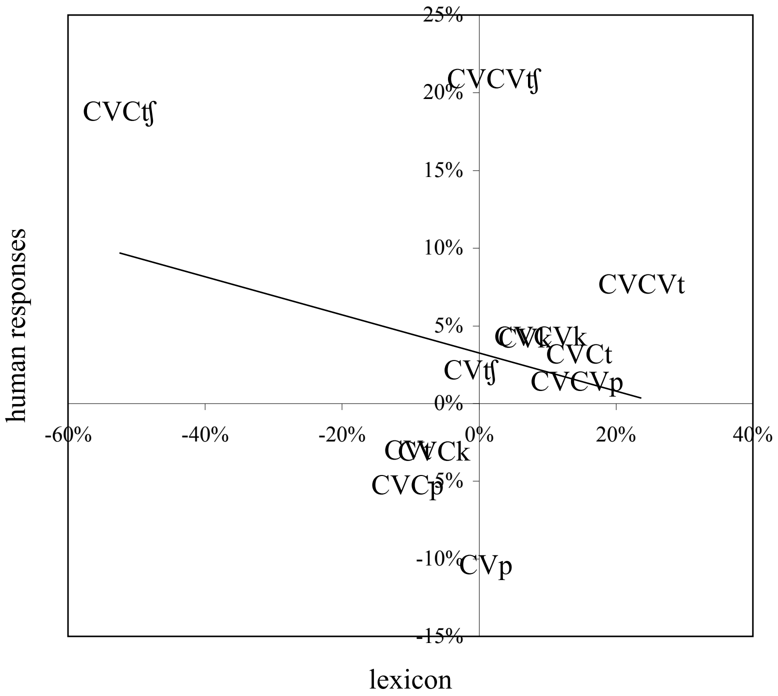
(15) Proportions of nouns with voicing alternations in the lexicon vs. the percent of alternating choices in the experiment, by size and place.



In stark contrast to the tight correlation between the experimental results and the lexicon for place and size effects, as seen in (15), there is no pattern when the height or backness effects are considered. The chart in (16) shows the results of the height factor. Each point in this chart shows the difference in rates of alternation between high and non-high vowels, by size and place. Positive values indicate more alternations with [+high] vowels, and negative values indicate more alternations with [−high] vowels.

There is no correlation between the lexicon and speakers' performance when vowel height is considered (Spearman's rank correlation test, $S = 196.8$, $\rho = .312$, $p > .1$). The chart in (16) shows that speakers' behavior was essentially random with respect to vowel height.

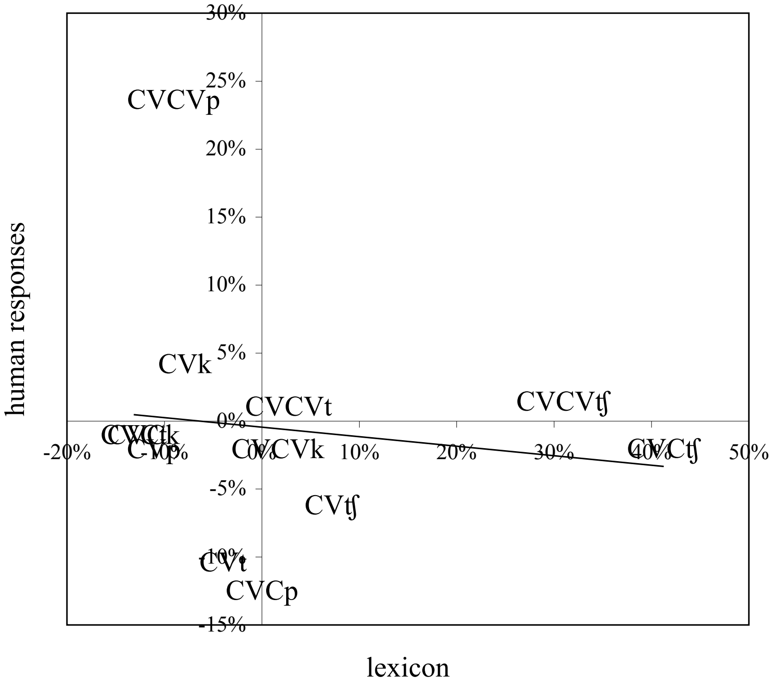
(16) Differences between high and non-high stem-final vowels in the lexicon vs. the differences between high and non-high vowels in the experiment, by size and place.



The lack of correlation in (16) is probably only due to a subset of the points, most noticeably CVCt͡ʃ, CVCVt͡ʃ, and CVp. There is no sense, however, in which these are "outliers", as they represent a sizable proportion of the data. The data for the CVCt͡ʃ point, for instance, comes from 18 lexical items and from 96 experimental responses (4 items * 24 participants).

When vowel backness is considered (17), the result is essentially the same: There is no correlation between the lexicon and speakers' responses when the results are categorized by size, place and backness (Spearman's rank correlation test, $S = 326.1$, $\rho = -.140$, $p > .1$). Each point in (17) shows the difference in rates of alternation between back and front vowels, by size and place. Positive values indicate more alternations with back vowels, and negative values indicate more alternations with front vowels.

(17) Differences between back and front stem-final vowels in the lexicon vs. the differences between back and front vowels in the experiment, by size and place.



The contrast between the strong correlation in (15) and the lack of correlation in (16-17) shows that speakers' behavior is best understood as replicating the lexicon's size and place effects, but not replicating its height or backness effects. This contrast is seen in the statistical analysis below.

The results were analyzed with a mixed-effects logistic regression in R (R Development Core Team 2007) using the *lmer()* function of the LME4 package, with *participant* and *item* as random effect variables. The fixed effect variables were the same ones used in the analysis of the lexicon: *size*, *place*, *high*, *back* and *round*.

An initial model was fitted to the data using only *size* and *place* as predictors. Adding their interaction to the model made a significant improvement (sequential ANOVA model comparison, $\chi^2(6) = 50.58$, $p < .001$). The improved model with the interaction term is given in (18). This model shows that labial place and CVCVC size are more conducive to alternating responses than the baseline dorsal place and CVC size, respectively. As for interactions, for the CVCC size, palatal place is more conducive to voicing than the baseline dorsal place with the same CVCC size. Additionally, in the CVCVC size, all places are less conducive to alternating responses than the baseline dorsal place eith the same CVCVC size. All of these effects mirror the lexical effects as presented in §2. The model stays essentially unchanged when validated by the *pvals.fnc()* function (Baayen 2008).

(18)

	Estimate	SE	z	p
(Intercept)	−0.864	0.283	−3.056	0.002
COR	0.111	0.256	0.434	0.665
LAB	0.744	0.304	2.451	0.014
PAL	−0.119	0.320	−0.372	0.710
CVCC	−0.089	0.260	−0.341	0.733
CVCVC	2.694	0.285	9.469	< 0.001
CVCC:COR	0.385	0.361	1.065	0.287
CVCC:LAB	0.641	0.431	1.487	0.137
CVCC:PAL	1.867	0.447	4.173	< 0.001
CVCVC:COR	−1.936	0.377	−5.142	< 0.001
CVCVC:LAB	−1.436	0.455	−3.154	0.002
CVCVC:PAL	−1.126	0.457	−2.463	0.014

The addition of any vowel feature to the baseline model (*high*, *back* or *round*) made no significant improvement ($p > .1$). No vowel feature approached significance, either on its own or by its interaction with *place*. For example, adding the interaction *place*high* to the model in (18) gives a new model where the interaction of coronal place and *high* is almost exactly at chance level ($p = .981$). Adding *place*back* the to baseline model gives an interaction of palatal place and *back* that is non-significant ($p = .661$) and its coefficient is negative, i.e. going in the opposite direction from the lexicon, where palatal place and backness are positively correlated.

In other words, *size* and *place* had statistically significant power in predicting the participants' choice of alternation vs. non-alternation of stem-final stops. Crucially, however, none of the vowel features had an effect on the participants' choices, either significantly or as a mere trend.

To summarize the findings, Turkish speakers reproduced the distribution of voicing alternations in the lexicon by paying attention to the size of the nouns and the place of the final stops, while ignoring the quality of the vowel that precedes the stem-final stop.

3.3. Discussion

The experimental results show that Turkish speakers generalize their knowledge of the voicing alternations in their lexicon. Not contenting themselves with memorizing the alternating or non-alternating status of single nouns, speakers have access to the relative proportion of alternating nouns categorized by size and place. Using size and place as factors, speakers must somehow project their lexical statistics onto novel items. Although the height and backness of stem-final vowels are strongly correlated with alternations in the lexicon, speakers' treatment of stem-final vowels in novel words is random, showing no significant interaction with their choice of alternating or non-alternating forms.

Speakers failed to reproduce the correlation between vowels and voicing alternations in spite of an abundance of overt evidence, while learning the size and place effects even with very little evidence. For instance, the difference in alternation rates between *t͡ʃ*-final CVC and CVCC nouns was successfully reproduced in the experiment results, even though the evidence comes from 23 and 18 words, respectively. The evidence for the vowel effects, however, comes from hundreds of words.

The proposal advanced here is that the results are best understood in light of a theory of universally possible phonological interactions, as encoded in a set of universal constraints. Only factors that can be expressed in terms of constraint interaction can be identified by language learners, with other lexical generalizations going unnoticed. This model is contrasted with general-purpose statistical learners that can learn any robust distributional generalization, as discussed in §6.

4. Turkish voicing alternations and Underlying Representations

Before we present our analysis of Turkish in §5, which uses an Optimality Theoretic grammar with lexically-specific rankings, we review the phonetics and phonology of laryngeal contrasts in Turkish (§4.1). We then show why the difference between alternating and non-alternating nouns must not be encoded in the underlying representation of roots (§4.2) if one is to formulate a grammatical explanation for our experimental results.

4.1. Laryngeal Contrasts in Turkish

The literature on Turkish (at least since Lees 1961) agree that Turkish contrasts two stops in each place of articulation on the surface (19), but that stem-final stops display three kinds of behavior under affixation: They are either pronounced the same in the base and in the affixed form (20a-b), or they alternate (20c). It is also known that final voiced stops, as in (20b), are rare in the language.

(19) Two-way surface distinction in roots

	initially		inter-vocalically	
a.	t^{h}in	'soul'	at^{h}a	'ancestor'
b.	din	'religion'	ada	'island'

(20) Three different contrasts finally

	bare stem	possessive	
a.	at^{h}	at^{h}-ɨ	'horse'
b.	ad	ad-ɨ	'name'
c.	t^{h}at^{h}	t^{h}ad-ɨ	'taste'

In Turkish orthography, the surface distinction is represented by the letters ⟨p, t, ç, k⟩ and ⟨b, d, c, g⟩, and the distinction was taken to be one of voicing by much of the literature on Turkish (Lees 1961; Inkelas & Orgun 1995; Inkelas et al. 1997, and many others).

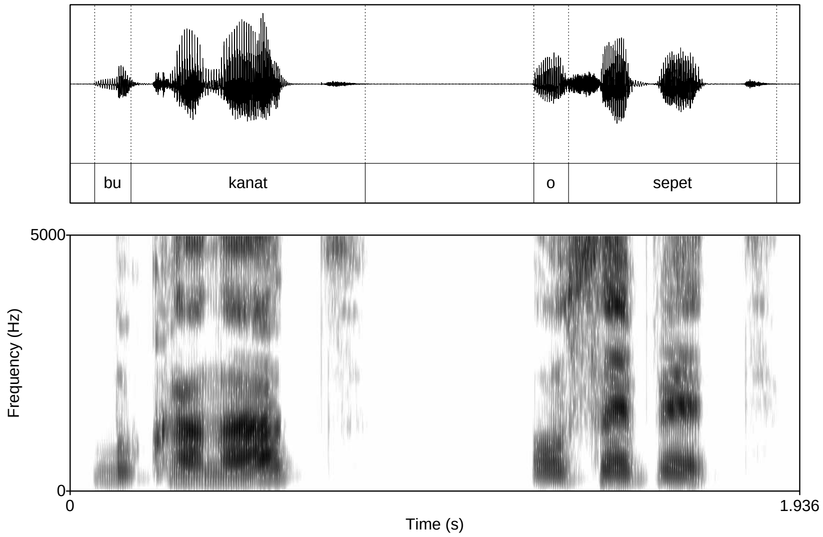
More recently, Kallestinova (2004) and Petrova et al. (2006) have shown that the voiceless stops of Turkish are in fact aspirated in onset position.[12] While these authors do not commit to the surface realization of word-final stops, it is known that final stops are consistently released with an audible voiceless burst. Crucial evidence for considering this audible release as aspiration comes from Kopkallı (1993), who shows that the release of word-final stops is as long as the duration of aspiration on intervocalic voiceless stops, suggesting that speakers treat these as a consistent phonetic category. For further discussion of laryngeal features in Turkish, see Jannedy (1995), and for a broader perspective, see Avery (1996), Beckman & Ringen (2004), and Vaux & Samuels (2005).

---

[12] The aspiration is consistent in roots. In affixes that show voicing alternations, such as the locative *–ta / –da* and the ablative *–tan / –dan*, the voiceless variant is unaspirated. In affixes that don't alternate, like the relativizer *–k^{h}i* and adverbial *–k^{h}en*, voiceless stops are aspirated just like root stops.

The spectrogram in (21) exemplifies the finding in Kopkallı (1993), showing a clear, voiceless burst at the end of both the alternating $k^{h}anat^{h}$ and the non-alternating $sep^{h}et^{h}$. In fact, this token, spoken by a 30 year old male speaker from Istanbul, happens to have an even stronger burst for $k^{h}anat^{h}$, although Kopkallı (1993) shows that there is no significant difference in the duration of the final burst between alternating and non-alternating nouns.

(21) [bu k$^{\text{h}}$anat$^{\text{h}}$ o sep$^{\text{h}}$et$^{\text{h}}$] “This is a wing; that is a basket” (lit. this wing; that basket)



For the purposes of the analysis we offer in §5, the exact details of Turkish laryngeal features are not crucial. What is crucial is that all stop-final nouns fall into one of two groups: In one group, the suffixed form is faithful to the base (such that faithfulness to laryngeal features ranks over any relevant markedness constraints), and in the other group, the suffixed form is unfaithful (and markedness ranks over any relevant faithfulness constraints). As we will show, the inconsistent ranking arguments allow the speaker to build lexical information into their grammar, and thus learn the distribution of the voicing alternations in grammatical terms. In this paper, we use the more accurate transcription, which marks aspiration.

Under this view, Turkish stops surface either voiced or aspirated. Any hypothetical underlyingly voiceless unaspirates map unfaithfully either to voiced or to aspirated stops due to high ranking constraint that requires a laryngeal specification on every stop (Petrova et al. 2006). Additionally, barring a few exceptional native words and some loanwords, word-final stops are regularly required to be aspirated, as was shown for German, Kashmiri, and

Klamath (Iverson & Salmons 2007).

4.2. Encoding (Non-)Alternation with Constraint Rankings instead of Underlying Representations

The existing analyses of Turkish voicing alternations, either in terms of voicing (Lees 1961; Inkelas & Orgun 1995; Inkelas et al. 1997) or in terms of aspiration (Avery 1996; Kallestinova 2004; Petrova et al. 2006), share the same architecture that attributes the different behavior of final stops to different underlying representations of laryngeal features. In this section, we argue specifically against this analysis, showing that it prevents speakers from learning the distribution of voicing alternations in grammatical terms.

The traditional analysis along the lines of (Inkelas et al. 1997) is shown in (22). In this analysis, nouns that surface with a voiceless (aspirated) stop throughout the paradigm have a voiceless (aspirated) stop underlyingly, while stops that alternate have an underlying stop that is unspecified for laryngeal features. Identity to larnygeal features assures that underlyingly specified stops surface faithfully in all positions, while a constraint against intervocalic voiceless stops causes alternation when faithfulness is not at issue.

(22) a. The UR's of [at$^{\mathrm{h}}$] and [t$^{\mathrm{h}}$at$^{\mathrm{h}}$] are /at$^{\mathrm{h}}$/ and /t$^{\mathrm{h}}$aD/
b. The UR of the possessive is /I/ (a high vowel)
c. /at$^{\mathrm{h}}$ + I/ → [at$^{\mathrm{h}}$-ɨ] **requires** IDENT(lar) ≫ *VtV

at$^{\mathrm{h}}$ + I	IDENT(lar)	*VtV
a. ☞ at$^{\mathrm{h}}$-ɨ		*
b. ad-ɨ	*!	

d. /t$^{\mathrm{h}}$aD + I/ → [t$^{\mathrm{h}}$ad-ɨ] is **consistent** with IDENT(lar) ≫ *VtV

t$^{\mathrm{h}}$aD + I	IDENT(lar)	*VtV
a. t$^{\mathrm{h}}$at$^{\mathrm{h}}$-ɨ		*!
b. ☞ t$^{\mathrm{h}}$ad-ɨ		

In this theory, Ident(lar) dominates any relevant markedness constraints, and alternating stops have under-specified underlying representations that escape faithfulness. The deletion of dorsals can be encoded using another representational mechanism, that of “floating segments”, or segments whose absence from the output does not violate the regular Max (as in, e.g., Zoll 1996).

The crucial element of this analysis is that both rankings in (22) are consistent. In other words, the behavior of alternating nouns like t^hat^h and non-alternating nouns like at^h do not require different grammatical factors that point to their alternation, and thereby cannot situate alternation itself as something specifically interacting with the phonological grammar of the language. Rather, the behavior of different nouns is encoded in the lexicon, outside the purview of grammar. The same is true of Avery (1996); Kallestinova (2004) and Petrova et al. (2006).

We propose that the status of a word as alternating or non-alternating must be represented by lexically-specific grammatical rankings, instead of in terms of an underlying difference. In essence, our argument is that only by including the alternating or non-alternating status of a word as a *grammatical* rather than lexically memorized phenomenon can one make sense of the grammatical biases against extending all lexical statistics.

Since the experiment in §3 shows that speakers have detailed grammatical knowledge about the propensity of final stops to alternate, it is not clear how speakers could encode this knowledge if they had allowed it to escape the grammar. Relegating information about voicing alternations to the lexicon would force speakers to look for generalizations directly in the lexicon, where nothing would prevent them from finding the vowel quality effects that they didn’t exhibit in §3.

The analysis offered in §5, summarized in (23) below, posits the bare forms of nouns as their underlying representations, and it is exactly this move that forces the speaker to find conflicting ranking arguments, and then encode lexical statistics in the grammar.

(23) a. The UR’s of [at^{h}] and [$\text{t}^{\text{h}}\text{at}^{\text{h}}$] are /$\text{at}^{\text{h}}$/ and /$\text{t}^{\text{h}}\text{at}^{\text{h}}$/
b. The UR of the possessive is /I/ (a high vowel)
c. /at^{h} + I/ → [at^{h}-ɨ] **requires** Ident(Lar) $\gg$ *VtV
/$\text{t}^{\text{h}}\text{at}^{\text{h}}$ + I/ → [t^{h}ad-ɨ] **requires** *VtV $\gg$ Ident(Lar)

The distribution of voicing alternation in Turkish is available to speakers: They know how many words have alternating stops and how many have non-alternating stops, and they keep this information separately for the stops in the different places of articulation, and within each place, for mono-syllablic nouns separately from poly-syllabic nouns. The availability of this knowledge is predicted by an approach that partitions the lexicon based on phonological principles, and it is left unexplained by the UR-based analysis in (22).

In the UR-based analysis, the grammar (IDENT(LAR) $\gg$ *VtV) is consistent for all the words of the language, and therefore the learner is left without a way to build lexical statistics into their grammar. In principle, one could imagine that a speaker will find the relevant lexical statistics by going directly to the lexicon and extracting the relevant information from it. When going to the lexicon directly, however, the speaker will not be biased by UG to find only grammatically-principled generalizations. Any kind of regularity in the lexicon could be discovered and projected onto novel items, contrary to fact: In the Turkish lexicon, there is a trend for more voicing alternations after high vowels than after low vowels, yet speakers show no sign of having learned this trend.

Assuming the base form of a noun as its underlying representation means that any additional aspects of the noun's pattern that are not directly observable in the base form will have to be attributed to other aspects of the linguistic system. Given an OT framework that uses underlying representations of roots and affixes and a constraint ranking, if hidden properties of roots are blocked from being attributed to those roots, hidden properties can only be attributed to the underlying representations of affixes or to the grammar.

Seeing that encoding the hidden pattern of lexical items in the underlying representations of either roots or suffixes leaves the learner with no way or reason to identify lexical trends, encoding such patterns in the grammar is left as the only logical option. Capturing hidden patterns in terms of cloned constraints assures that lexical trends are identified in terms of constraints, i.e. it assures that trends are captured in phonological terms, using the variety of phonological primitives that constraints are sensitive to, such as marked combinations of features, preferred alignments of phonological elements, positional faithfulness, etc.

Contrasted with traditional generative analyses, the proposal made here "reverses" the effect of the phonology. Instead of assigning the hidden aspects of bases to their underlying representation, and then neutralizing them in the unaffixed form, as is done traditionally, we propose that the surface forms of bases are assumed as their underlying form, and any properties of the base that emerge only in suffixed forms are achieved by constraint interaction. In the simple case of Turkish, where the only hidden property of nominal roots is the voicing of their final stop, the analysis in terms of cloned constraints is not only clearly feasible, it is also the only analysis that allows speakers to capture the variety of lexical trends that the language has.

The idea that the surface form of the base, rather than some abstract underlying form, is preferred in phonological systems has been argued for in Hayes (1995, 1999). Assuming the base form as the underlying representation has the added benefit of obviating the search for non-surface-true underlying representations. This search requires a significant amount of computation, as shown by Tesar (2006) and Merchant (2008), and in parallel lines of work, also by Boersma (2001) and by Jarosz (2006), who specifically look at final-devoicing languages, i.e. languages like Turkish, where the pattern of root-final stops is hidden in the bare form of the root. A full comparison of the computational complexity of these approaches and our approach, however, goes beyond the scope of this paper.

5. Analysis with cloned constraints

Turkish speakers evidence a detailed knowledge of trends in their lexicon that regulate the choice of alternation or non-alternation of stem-final stops. Furthermore, speakers are biased by Universal Grammar to learn only lexical trends that can be captured in terms of cross-linguistically observed interactions between phonological elements. This section shows how an OT-based model can be used to learn the trends the humans learn. The model reads in the lexicon of Turkish and projects a probabilistic grammar from it, a grammar that can in turn be used to derive novel words in a way that correlates with the experimental results shown in §3.

Given a stop-final novel noun and asked to choose a possessive form for it, Turkish speakers consult a subset of their lexicon: For instance, given the noun *dap*^{h}, speakers identify it as a mono-syllabic *p*-final simplex-coda noun, and they compare it to the other mono-syllabic *p*-final simplex-coda nouns in their lexicon. If they have 30 such nouns, of which 8 alternate and 22 don't alternate, as in TELL, then the likelihood that *dap*^{h} will exhibit a voicing alternation is 8 out of 30, or 27%.

In other words, Turkish speakers partition their lexicon based on phonological principles. The mass of stop-final nouns is partitioned by the size of each noun (mono- vs. poly-syllabic), by the place of articulation of the final stop (p, t, t͡ʃ, k), and by the complexity of the final coda, and within each such group, alternating nouns are separated from non-alternating nouns. This creates a total of 2 * 4 * 2 * 2 = 32 partitions. Nouns that don't end in a stop are all lumped together in the "elsewhere" partition.

Constraint cloning is a mechanism for partitioning the lexicon and listing the words that belong in each partition. The partitions are defined by the set of universal constraints in CON, which ensures that nouns are only categorized based on universal grammatical principles.

5.1. Constraint cloning

The OT-based model proposed here makes crucial use of the concept of Inconsistency Resolution, offered by Pater (2006, 2008), which relies on the Recursive Constraint Demotion Algorithm (RCD, Prince & Tesar 1999).

In RCD, the speaker learns from "errors", or mismatches between the words of the language they are exposed to and the words that are produced by their current grammar. Suppose the learner hears the adult form [k^{h}anat^h] 'wing', but their grammar produces [k^{h}ana], because the markedness constraint *CODA is ranked above faithfulness in their grammar (24).

(24)

[k^{h}anat^h]	*CODA	MAX
a. ☹ k^{h}anat^h	*!	
b. ☞ k^{h}ana		*

Since the current winner, [k^{h}ana], is different from the adult form, the speaker constructs a winner-loser pair, as in (25). The tableau in (25) is a comparative tableau (Prince 2002), where W means "winner-preferring" (i.e. the constraint assigns less violations to the winner) and L means "loser-preferring (i.e. the constraint assigns less violations to the loser).

(25)

	*CODA	MAX
a. k^{h}anat^h $\succ$ k^{h}ana	L	W

RCD takes winner-loser pairs such as the one in (25) and extracts a grammar from them by identifying columns that don't have L's in them and "installing" them. In this simple case, MAX can be installed, meaning that it is added to the grammar below any other previously installed constraints (which would be at the top of the grammar in this case, since no constraints were previously installed), and winner-loser pairs that MAX assigns a W to are removed from the tableau. Once MAX is thus installed, the tableau is emptied out, and the remaining constraints, in this case just *CODA, are added at the bottom of the grammar. The resulting grammar is now MAX ≫ *CODA, which allows codas to be produced, as in adult Turkish.

There is no guarantee, however, that RCD will always be able to install any constraints and remove all of the winner-loser pairs from the tableau. If all of the available columns have L's in them, RCD will stall. This situation arises when the language provides the learner with conflicting data, as in (26). In some words, a stem-final stop is voiceless aspirated throughout the paradigm (26a-b), and in others, a final stop shows up voiceless aspirated in the bare stem and voiced in the possessive (26c-d).

(26)

	bare stem	possessive	
a.	at͡ʃʰ	at͡ʃʰ-ɨ	'hunger'
b.	anat͡ʃʰ	anat͡ʃʰ-ɨ	'female cub'
c.	tat͡ʃʰ	tad͡ʒ-ɨ	'crown'
d.	amat͡ʃʰ	amad͡ʒ-ɨ	'target'

Assuming the bare stem with its voiceless aspirated stop as the underlying form, as discussed in §4, the non-alternating forms rank faithfulness to the underlying representations above the markedness pressure against intervocalic voiceless stops (27), while alternating forms require ranking faithfulness below markedness (28).

(27)

/ anat͡ʃʰ + ɨ /	IDENT(LAR)	*Vt͡ʃV
a. ☞ anat͡ʃʰ-ɨ		*
b. anad͡ʒ-ɨ	*!	

(28)

/ amat͡ʃʰ + ɨ /	*Vt͡ʃV	IDENT(LAR)
a. ☞ amad͡ʒ-ɨ		*
b. amat͡ʃʰ-ɨ	*!	

With this understanding of the situation, the ranking between the faithfulness constraint IDENT(LAR) and the markedness constraint *Vt͡ʃV cannot be determined for the language as a whole. Pairing the winners in (27) and (28) with their respective losers allows the ranking arguments to be compared, as in (29).

(29)

	IDENT(LAR)	*Vt͡ʃV
a. anat͡ʃʰ-ɨ ≻ anad͡ʒ-ɨ	W	L
b. amad͡ʒ-ɨ ≻ amat͡ʃʰ-ɨ	L	W

Since the ranking arguments in (29) are inconsistent, there are no rows with no L's in them, and therefore no constraints can be installed, and a grammar cannot be found using RCD. Pater (2006) proposes a mechanism for resolving such inconsistencies by cloning. In cloning, the speaker replaces a universal constraint of general applicability with two copies, or clones, of the universal constraint that are lexically-specific, with each clone listing the lexical items it applies to.[13]

Given the situation in (29), the speaker can clone IDENT(LAR), making one clone specific to the root *anat͡ʃʰ* (and any other lexical items that IDENT(LAR) assigns a W to), and the other clone specific to the root *amat͡ʃʰ* (and any other lexical items that IDENT(LAR) assigns an L to). The resulting grammar is no longer inconsistent:

---

[13] Pater (2006) suggests a slightly different mechanism, where one clone is lexically specific and the other clone stays general. We argue in §5.2 below that both clones must be lexically specific to account for the behavior of Turkish speakers.

(30)

	IDENT (voice)$_{anat͡ʃʰ}$	IDENT (voice)$_{amat͡ʃʰ}$	*Vt͡ʃV
a. anat͡ʃʰ-ɨ ≻ anad͡ʒ-ɨ	W		L
b. amad͡ʒ-ɨ ≻ amat͡ʃʰ-ɨ		L	W

Now RCD can be successfully applied to (30): First, IDENT(LAR)$_{anat͡ʃʰ}$ is installed, and the first winner-loser pair is removed. This leaves the column of *Vt͡ʃV with no L's in it, so *Vt͡ʃV is installed below IDENT(LAR)$_{anat͡ʃʰ}$, and the second winner-loser pair is removed. The remaining constraint, IDENT(LAR)$_{amat͡ʃʰ}$ is added to the ranking below *Vt͡ʃV. The resulting grammar is IDENT(LAR)$_{anat͡ʃʰ}$ ≫ *Vt͡ʃV ≫ IDENT(LAR)$_{amat͡ʃʰ}$, which correctly blocks the voicing alternation in *anat͡ʃʰ-ɨ* but allows it in *amad͡ʒ-ɨ*. In the case of (29), choosing to clone IDENT(LAR) solved the inconsistency, but cloning *Vt͡ʃV would have been equally useful. The question of which constraint to clone is beyond the scope of this paper, and it is addressed more systematically in Becker (2009)

The cloning of IDENT(LAR), and the listing of lexical items with its clones, divided the lexicon into three partitions: One partition contains the items listed with the high-ranking clone of IDENT(LAR), another partition contains the items listed with the low-ranking clone of IDENT(LAR), and a third partition contains all the lexical items that are not listed with either clone. These partitions are not arbitrary, but rather determined by the the mark that IDENT(LAR) assigns to each winner-loser pair: W, L, or none.

Once a constraint is cloned, its clones accumulate lists of the morphemes they apply to. This approach allows for two sub-grammars to coexist in a language, while keeping track of the number of lexical items that belong to each sub-grammar. Since the number of lexical items of each kind becomes available in the grammar, the speaker can estimate the likelihood of each pattern.

The rest of this section shows how constraint cloning creates a grammar of Turkish that reflects speakers' knowledge of the lexicon, as determined by the experimental findings in §3.

5.2. The place effect

As discussed in §2, all stops are not equally likely to alternate: While the stops in most *t͡ʃ*-final and *p*-final nouns alternate, the stops in most *t*-final nouns do not. The table in (31), repeated from (4) above, lists the numbers of alternating and non-alternating (faithful) paradigms by the place of articulation of the final stop, as found in TELL (Inkelas et al. 2000).

(31)

Place	Alternating	Faithful	Total	% alternating
p	247	47	294	84%
t	214	1041	1255	17%
t͡ʃ	117	74	191	61%
k	1071	191	1262	85%

To replicate the effect that place has over the distribution of voicing alternations, the language learner must separately keep track of words that end in different stops. The fact that laryngeal features affects stops of different places of articulation differently is well documented (e.g. Lisker & Abramson 1964; Ohala 1983; Volatis & Miller 1992). Additionally, the lenition of voiceless stops to voiced stops between vowels is also very well documented (for an overview, see Kirchner 1998). These effects quite plausibly give rise to a family of constraints that penalize voiceless stops between vowels: *VpV, *VtV, *Vt͡ʃV, *VkV. The interaction of each of these constraints with IDENT(LAR) will allow the speaker to discover the proportion of the stop-final nouns of Turkish that alternate in each place of articulation.

Note that for each place of articulation, the speaker has to keep track of both the number of words that alternate and the number of words that do not. Simply keeping a count of words that alternate leads to a wrong prediction: Compare, for instance, *t*-final words and *t͡ʃ*-final words. There are 214 *t*-final words that alternate, but only 117 *t͡ʃ*-final words that do. If the speaker were to only keep a count of alternating words, they would reach the conclusion that *t*-final words are more likely to alternate. But in fact, speakers choose alternating responses with *t͡ʃ*-final words more often than they do with *t*-final words, reflecting the relative proportions of alternating and non-alternating nouns, not the absolute number of alternating nouns.

Similarly, keeping track of just the non-alternating nouns will also make the wrong prediction. Comparing *t͡ʃ*-final words and *k*-final words, we see that there are more than twice as many *k*-final non-alternators than there *t͡ʃ*-final non-alternators. Speakers, however, choose non-alternating responses with *k*-final words less often than they do with *t͡ʃ*-final words. In order to match the proportion of alternating stops in each place, both alternating and non-alternating words will need to be tracked.

Imagine a learner that has learned just two paradigms, *amat͡ʃʰ* ∼ *amad͡ʒ-ɨ* and *sepʰetʰ* ∼ *sepʰetʰ-i*. While one alternates and the other doesn't, no inconsistency is detected yet, since IDENT(LAR) interacts with two different markedness constraints (32).

(32)

	IDENT(LAR)	*VtV	*Vt͡ʃV
a. amad͡ʒ-ɨ ≻ amat͡ʃʰ-ɨ	L		W
b. sepʰetʰ-i ≻ sepʰed-i	W	L	

Running RCD on (32) yields the clone-free grammar *Vt͡ʃV ≫ IDENT(LAR) ≫ *VtV. If the speaker learns the word *anat͡ʃʰ* ∼ *anat͡ʃʰ-ɨ*, however, the grammar becomes inconsistent (33).

(33)

	IDENT(LAR)	*VtV	*Vt͡ʃV
a. amad͡ʒ-ɨ ≻ amat͡ʃʰ-ɨ	L		W
b. anat͡ʃʰ-ɨ ≻ anad͡ʒ-ɨ	W		L
c. sepʰetʰ-i ≻ sepʰed-i	W	L	

Since there are no columns in (33) that don't have L's in them, RCD stalls. Cloning either *Vt͡ʃV or IDENT(LAR) can resolve the inconsistency. In this case, *Vt͡ʃV is chosen since its column has the least number of non-empty cells. The result of cloning *Vt͡ʃV is shown below:

(34)

	ID(voice)	*VtV	*Vt͡ʃV$_{\textit{amat͡ʃʰ}}$	*Vt͡ʃV$_{\textit{anat͡ʃʰ}}$
a. amad͡ʒ-ɨ ≻ amat͡ʃʰ-ɨ	L		W	
b. anat͡ʃʰ-ɨ ≻ anad͡ʒ-ɨ	W			L
c. sepʰetʰ-i ≻ sepʰed-i	W	L		

Installing *Vt͡ʃV$_{\textit{amat͡ʃ}}$ removes the first winner-loser pair. This leaves IDENT(LAR) with no L's in its column, so it is installed, and the last two winner-loser pairs are removed. Then, *VtV and *Vt͡ʃV$_{\textit{anat͡ʃ}}$ are installed, yielding the ranking in (35).

(35) *Vt͡ʃV$_{\textit{amat͡ʃ}}$ ≫ IDENT(LAR) ≫ *VtV, *Vt͡ʃV$_{\textit{anat͡ʃ}}$

The resulting grammar has successfully partitioned the data available to the learner: Lexical items that end in *t͡ʃ* are listed with the two clones of *Vt͡ʃV, and the *t*-final noun was not listed, since *t*-final nouns behave consistently in this limited set of data.

Cloning of *VtV will only become necessary once the speaker encounters a word with an alternating *t*, e.g. *kʰanatʰ* ∼ *kʰanad-ɨ* 'wing', as in (36). Note that whenever the speaker learns a new paradigm, information about constraint conflicts may change; therefore, constraint cloning always starts from square one with the addition of a new winner-loser pair.

(36)

	ID(voice)	*VtV	*Vt͡ʃV
a. amad͡ʒ-ɨ ≻ amat͡ʃ-ɨ	L		W
b. anat͡ʃ-ɨ ≻ anad͡ʒ-ɨ	W		L
c. kʰanad-ɨ ≻ kʰanatʰ-ɨ	L	W	
d. sepʰetʰ-i ≻ sepʰed-i	W	L	

Given (36), cloning *Vt͡ʃV will not suffice to make the grammar consistent. If *Vt͡ʃV is cloned first, the learner will install *Vt͡ʃV$_{\textit{amat͡ʃʰ}}$ and remove the first winner-loser pair, but then they will still have a tableau with no columns that have no L's in them. Cloning *VtV as well will solve the inconsistency, and the resulting grammar would be as in (37).

(37) *Vt͡ʃV$_{amat͡ʃ^h}$, *VtV$_{k^hanat^h}$ ≫ IDENT(LAR) ≫ *VtV$_{sep^het^h}$, *Vt͡ʃV$_{anat͡ʃ^h}$

The resulting grammar in (37) successfully partitioned the lexicon: *t*-final nouns are listed with clones of *VtV, and *t͡ʃ*-final nouns are listed with clones of *Vt͡ʃV. These partitions are defined by the constraints that distinguish winners from losers. The language learner's ability to treat each place separately is a consequence of the availability of universal constraints that relate voicing and place of articulation. These constraints let the speaker detect inconsistency in each place separately, and create lists of lexical items in each place.

5.3. The size effect

Both the lexicon (§2) and the experimental results (§3) show a higher preference for alternations in poly-syllabic nouns relative to mono-syllabic, in every place of articulation. The size effect is not equal across the different places, however. Mono-syllabic nouns generally don't alternate, regardless of the place of articulation of their final stop. Poly-syllabic nouns usually do alternate if they are *p*-final or *t͡ʃ*-final, but not if they are *t*-final. Speakers have replicated this pattern of differential treatment of poly-syllabic nouns. In statistical terms, the size and place affect have a significant interaction, and the implication for the learner is that the proportion of alternating nouns is learned separately in each place-size combination.

The proposed account of this size effect relies on the position of the alternating final stop relative to the initial syllable of the root. In a mono-syllabic noun, the unfaithful mapping from a voiceless stop to a voiced one affects the initial syllable of the base, while a voicing alternation in a poly-syllablic noun doesn't affect the initial syllable. Initial syllables are known to enjoy greater faithfulness cross-linguistically, as formalized by Beckman (1997).[14] The availability of a faithfulness constraint that protects only mono-syllabic roots allows the speaker to partition the lexicon along this dimension, putting mono-syllables in one partition, and leaving the other nouns, which are therefore poly-syllabic, in another partition. The formalization of initial-syllable faithfulness in Beckman (1997) refers to the

---

[14] In a separate line of work, Dresher & van der Hulst (1998) derive similar results by using head/dependent asymmetries.

initial syllable of the derived form, not the output, but the use of positional faithfulness defined over the base is not without precedent, e.g. Kager (1999b).

The role of the word-initial syllable in the distribution of laryngeal features in Turkish is not limited to voicing alternations. Generally in the language, a coda stop followed by an onset stop will surface with the laryngeal specification of the onset stop (e.g. _is.tʰ_**_ib_**_.datʰ_ ‘despotism’, *_is.tʰ_**_ip_**_ʰ.datʰ_), but a coda stop in the initial syllable may surface with its independent voicing specification (e.g. _ma_**_k_**_ʰ.bul_ ‘accepted’, **_eb_**_.kʰem_ ‘mute’).

For concreteness, this section focuses on learning the _t͡ʃ_-final nouns of Turkish with simple codas. The relevant lexical counts are in (38).

(38)

	CVt͡ʃ	CVCVt͡ʃ	Total
Faithful	18	44	62
Alternating	3	96	99
Total	21	140	161

Given both mono-syllabic and poly-syllabic nouns that do and do not alternate, as in (39), the learner can successfully separate mono-syllabic roots from poly-syllablic ones by cloning the specific IDENT(LAR)$_{\sigma 1}$ first.

(39)

	IDENT	IDENT$_{\sigma 1}$	*Vt͡ʃV
a. sat͡ʃʰ-ɨ ≻ sad͡ʒ-ɨ	W	W	L
b. tʰad͡ʒ-ɨ ≻ tʰat͡ʃʰ-ɨ	L	L	W
c. anat͡ʃʰ-ɨ ≻ anad͡ʒ-ɨ	W		L
d. amad͡ʒ-ɨ ≻ amat͡ʃʰ-ɨ	L		W

IDENT(LAR)$_{\sigma 1}$ can be identified as more specific than IDENT(LAR) by examining the number of W’s and L’s in each column, since the more specific constraint will necessarily assign a subset of the W’s and L’s that the general constraint assigns (Tessier 2007). The result of cloning IDENT(LAR)$_{\sigma 1}$ is in (40). Since only mono-syllabic stems are assigned W’s or L’s by IDENT(LAR)$_{\sigma 1}$, only mono-syllables get listed by clones at this point.

(40)

	Ident	Ident$_{\sigma1_{sat͡ʃ^h}}$	Ident$_{\sigma1_{t^hat͡ʃ^h}}$	*Vt͡ʃV
a. sat͡ʃʰ-ɨ ≻ sad͡ʒ-ɨ	W	W		L
b. tʰad͡ʒ-ɨ ≻ tʰat͡ʃʰ-ɨ	L		L	W
c. anat͡ʃʰ-ɨ ≻ anad͡ʒ-ɨ	W			L
d. amad͡ʒ-ɨ ≻ amat͡ʃʰ-ɨ	L			W

The column of Ident(Lar)$_{\sigma1_{sat͡ʃ^h}}$ has no L's in it, so it can be installed, and the first winner-loser pair can be removed from the tableau. While the mono-syllabic *t͡ʃ*-final nouns were successfully listed by clones of Ident(Lar)$_{\sigma1}$, the learner is not quite ready to discover the rest of the *t͡ʃ*-final nouns. Given the tableau in (40), there are no constraints to install after the installation of Ident(Lar)$_{\sigma1_{sat͡ʃ^h}}$, so either Ident(Lar) or *Vt͡ʃV will need to cloned. Once either of them is cloned, *tʰat͡ʃʰ* and *amat͡ʃʰ* will be listed with one clone, and *anat͡ʃʰ* will be listed with the other. Assuming it is Ident(Lar) that is cloned, the resulting grammar will be the one in (41).

(41) Ident(Lar)$_{\sigma1_{sat͡ʃ^h}}$ $\gg$ Ident(Lar)$_{anat͡ʃ^h}$ $\gg$ *Vt͡ʃV $\gg$
Ident(Lar)$_{\sigma1_{t^hat͡ʃ^h}}$, Ident(Lar)$_{t^hat͡ʃ^h,\ amat͡ʃ^h}$

The problem with the grammar in (41) is that the lexicon is not neatly partitioned in the way the learner needs it to be: The specific Ident(Lar)$_{\sigma1}$ correctly lists all and only the mono-syllables, but the general Ident(Lar), in addition to correctly listing all the poly-syllabic *t͡ʃ*-final nouns, also incorrectly lists the mono-syllabic *t͡ʃ*-final alternators.

The problem is that the general Ident(Lar) assigns W's and L's to all nouns, regardless of size, potentially allowing some nouns to "double dip", as seen in (41). To ensure that nouns are not listed multiple times, the learner needs to make sure that when they clone a specific constraint and list words with the clones, they also ignore any W's or L's that a more general constraint assigns to these listed words. In the case of (40), the learner needs to notice that Ident(Lar) is more general than Ident(Lar)$_{\sigma1}$ (as determined by the fact that Ident(Lar) assigns a superset of the W's and L's that Ident(Lar)$_{\sigma1}$ assigns), and ignore (or "mask") the W's and L's that Ident(Lar) assigns to the nouns that are listed by

IDENT(LAR)$_{\sigma 1}$.[15] The correct tableau, with the masking of the W that IDENT(LAR) assigns to *sat͡ʃʰ-ɨ* and the L that it assigns to *tʰad͡ʒ-ɨ*, is in (42).

(42)

	IDENT	IDENT$_{\sigma 1\,\textit{sat͡ʃ}}$	IDENT$_{\sigma 1\,\textit{tat͡ʃ}}$	*Vt͡ʃV
a. sat͡ʃʰ-ɨ ≻ sad͡ʒ-ɨ	~~Ⓦ~~	W		L
b. tʰad͡ʒ-ɨ ≻ tʰat͡ʃʰ-ɨ	~~Ⓛ~~		L	W
c. anat͡ʃʰ-ɨ ≻ anad͡ʒ-ɨ	W			L
d. amad͡ʒ-ɨ ≻ amat͡ʃʰ-ɨ	L			W

Given the tableau in (42), the column of IDENT(LAR) has the fewest W's and L's, so IDENT(LAR) will be chosen for cloning. The learner will clone IDENT(LAR) and successfully list just the poly-syllables with it. The resulting grammar will be the one in (43). This grammar achieves the intended partitioning of the lexicon: The *t͡ʃ*-final nouns are divided into mono-syllables and poly-syllables, and within each category, the nouns are further divided into alternators and non-alternators.

(43) IDENT(LAR)$_{\sigma 1\,\textit{sat͡ʃʰ}}$ ≫ IDENT(LAR)$_{\textit{anat͡ʃʰ}}$ ≫ *Vt͡ʃV ≫
IDENT(LAR)$_{\sigma 1\,\textit{tʰat͡ʃʰ}}$, IDENT(LAR)$_{\textit{amat͡ʃʰ}}$

To summarize, the analysis of the size effect in Turkish relies on the availability of a specific version of IDENT(LAR) that only assesses voicing alternations in mono-syllables. The speakers uses the specific IDENT(LAR)$_{\sigma 1}$ to list the mono-syllables, leaving the poly-syllables to the care of the general IDENT(LAR). The intended result relies on two principles: (a) the selection of the constraint to clone by identifying the column with the fewest non-empty cells, and (b) the masking of W's and L's from general constraints upon the listing of items with a specific constraint.

---

[15] The masking operation can also be defined to operate only on L's, since the W's will be removed by the installation of a clone of the specific constraint, and masking of W's will turn out to be vacuous.

5.4. Combining place and size

The distribution of the voicing alternations in Turkish is analyzed here as affected by two factors: The place of articulation of the final stop, which was attributed to the markedness of different stops between vowels, and the size, which was attributed to specific faithfulness to voicing in mono-syllables. The two effects have a significant interaction, where the size effect is strong in labials and palatals and much smaller for coronals. This section will show how the learner can model this interaction by using pairs of constraints to list lexical items.

The tableau in (44) shows the full range of possible winner-loser pairs given two places (t and t͡ʃ), two sizes (mono-syllabic and poly-syllabic) and two alternation patterns (faithful and alternating). The intended result is for the speaker to partition their lexicon by size and place, making four partitions, and within each of the four, further partition and list alternating and non-alternating items separately. Using the cloning technique that was offered in §5.2 and §5.3 above, no constraint will lead to the correct partitioning: For instance, cloning IDENT(LAR)$_{\sigma 1}$ will separate the alternating mono-syllabic nouns from the non-alternating mono-syllabic nouns, so *sat͡ʃʰ* and *atʰ* will be listed with one clone and *tʰat͡ʃʰ* and *tʰatʰ* will be listed with the other clone. But this listing collapses the place distinction, putting *t͡ʃ*-final nouns and *t*-final nouns in the same partition.

(44)

	IDENT	IDENT$_{\sigma 1}$	*Vt͡ʃV	*VtV
a. sat͡ʃʰ-ɨ ≻ sad͡ʒ-ɨ	W	W	L	
b. tʰad͡ʒ-ɨ ≻ tʰat͡ʃʰ-ɨ	L	L	W	
c. anat͡ʃʰ-ɨ ≻ anad͡ʒ-ɨ	W		L	
d. amad͡ʒ-ɨ ≻ amat͡ʃʰ-ɨ	L		W	
e. atʰ-ɨ ≻ ad-ɨ	W	W		L
f. tʰad-ɨ ≻ tʰatʰ-ɨ	L	L		W
g. sepʰetʰ-i ≻ sepʰed-i	W			L
h. kʰanad-ɨ ≻ kʰanatʰ-ɨ	L			W

The mechanism of cloning must be made sensitive to the various sources of conflict in the data: The column of IDENT(LAR)$_{\sigma1}$ indeed contains W's and L's, but these conflict with different constraints. Some W's that IDENT(LAR)$_{\sigma1}$ assigns are offset by L's from *VtV, and some are offset by L's from *Vt͡ʃV. Similarly, the L's that IDENT(LAR)$_{\sigma1}$ assigns are offset by W's from *VtV and from *Vt͡ʃV.

To capture the different sources of conflict in the data, lexical items that are listed with clones of IDENT(LAR)$_{\sigma1}$ must also mention which constraint they conflict with: If a lexical item gets a W from IDENT(LAR)$_{\sigma1}$, this W must be offset by an L from some other constraint, and vice versa. The clones of IDENT(LAR)$_{\sigma1}$ don't simply list lexical items, but rather list lexical items by the constraint they conflict with, or more formally, clones list ⟨constraint, {lexical items}⟩ pairs. This is shown in (45). As before, the listing of items with clones of the specific IDENT(LAR)$_{\sigma1}$ causes the masking of W's and L's from the column of the more general IDENT.

(45)

	IDENT	IDENT$_{\sigma1}$ ⟨*Vt͡ʃV, *sat͡ʃʰ*⟩, ⟨*VtV, *atʰ*⟩	IDENT$_{\sigma1}$ ⟨*Vt͡ʃV, *tʰat͡ʃʰ*⟩, ⟨*VtV, *tʰatʰ*⟩	*Vt͡ʃV	*VtV
a. sat͡ʃʰ-ɨ ≻ sad͡ʒ-ɨ	$\cancel{W}$	W		L	
b. tʰad͡ʒ-ɨ ≻ tʰat͡ʃʰ-ɨ	$\cancel{L}$		L	W	
c. anat͡ʃʰ-ɨ ≻ anad͡ʒ-ɨ	W			L	
d. amad͡ʒ-ɨ ≻ amat͡ʃʰ-ɨ	L			W	
e. atʰ-ɨ ≻ ad-ɨ	$\cancel{W}$	W			L
f. tʰad-ɨ ≻ tʰatʰ-ɨ	$\cancel{L}$		L		W
g. sepʰetʰ-i ≻ sepʰed-i	W				L
h. kʰanad-ɨ ≻ kʰanatʰ-ɨ	L				W

Next, the learner is ready to clone IDENT(LAR), which will again list items by the constraints they conflict with. The resulting grammar is in (46).

(46)
$$\textsc{Ident(Lar)}_{\sigma1\,\substack{\langle \text{*Vt͡ʃV, sat͡ʃ}^{\text{h}}\rangle \\ \langle \text{*VtV, at}^{\text{h}}\rangle}} \gg \textsc{Ident(Lar)}_{\substack{\langle \text{*Vt͡ʃV, anat͡ʃ}^{\text{h}}\rangle \\ \langle \text{*VtV, sep}^{\text{h}}\text{et}^{\text{h}}\rangle}} \gg \text{*Vt͡ʃV, *VtV} \gg$$
$$\textsc{Ident(Lar)}_{\sigma1\,\substack{\langle \text{*Vt͡ʃV, t}^{\text{h}}\text{at͡ʃ}^{\text{h}}\rangle \\ \langle \text{*VtV, t}^{\text{h}}\text{at}^{\text{h}}\rangle}},\ \textsc{Ident(Lar)}_{\substack{\langle \text{*Vt͡ʃV, amat͡ʃ}^{\text{h}}\rangle \\ \langle \text{*VtV, k}^{\text{h}}\text{anat}^{\text{h}}\rangle}}$$

This grammar correctly partitions the lexicon: Clones of Ident(Lar)$_{\sigma1}$ list all the mono-syllabic stop-final nouns that the speaker has, and those are further divided by markedness constraints into *t*-final and *t͡ʃ*-final nouns. Of course, the full grammar also lists *p*-final nouns under *VpV, and those *k*-final nouns that show a voicing alternation are listed under *VkV (for more on *k*-final nouns, see §5.6). The nouns that were assessed neither W's nor L's by Ident(Lar)$_{\sigma1}$, which are therefore poly-syllabic, are listed by clones of the general Ident(Lar). These again are listed by the markedness constraint that Ident(Lar) conflicts with, correctly separating the poly-syllabic nouns according to the place of articulation of their final stop.

This grammar allows the speaker to learn the proportion of alternating nouns in each size and place combination, with these combinations made available by listing lexical items with pairs of constraints.

5.5. The complex coda effect

As discussed in §2 and §3, stop-final CVC nouns have a lower proportion of alternators relative to CVCC nouns. The complexity of the coda does not have the same effect in all places of articulation, e.g. CVCC nouns have a proportion of alternators that's similar to the proportion of alternators among the poly-syllables when *p*-final and *t͡ʃ*-final nouns are considered, but *k*-final CVCC nouns pattern with the mono-syllabic *k*-final nouns, which have a low proportion of alternators.

Of the 354 stop-final nouns in TELL that have a complex coda, 244 have a sonorant before the final stop, and 39% of those 244 nouns alternate. Of the 110 nouns that have an obstruent before their final stop, only 3% alternate. Since only sonorants lead to a non-negligible proportion of alternators, only sonorants were used in the experiment in §3, and hence only nouns with a sonorant before their final stop will be considered below.

The alternation of nouns with simple codas was attributed in §5.2 to a family of markedness constraints that penalize intervocalic voiceless stops: *VpV, *VtV, *Vt͡ʃV, and *VkV.

Similarly, the alternations of nouns with complex codas is attributed here to markedness constraints that penalize voiceless stops between a sonorant consonant and a vowel, namely *RpV, *RtV, *Rt͡ʃV, and *RkV. This formulation of the constraints collapses the distinction between the nasal sonorants {m, n} and the oral sonorants {l, λ, r, y}, which might be an over-simplification. In the lexicon, stops are more likely to alternate following nasals than following oral sonorants (47.6% vs. 29.3%), a tendency that was also found in the experimental results (49.0% vs. 39.6%).

The pattern of alternating and non-alternating *t͡ʃ*-final nouns with final complex codas is shown in (47). The markedness constraint *Rt͡ʃV prefers alternation, while the familiar IDENT(LAR) and IDENT(LAR)$_{\sigma1}$ prefer a faithfully voiceless root-final stop.

(47)

	IDENT	IDENT$_{\sigma1}$	*Rt͡ʃV
a. gønt͡ʃʰ-y ≻ gønd͡ʒ-y	W	W	L
b. gend͡ʒ-i ≻ gent͡ʃʰ-i	L	L	W
c. gylynt͡ʃʰ-y ≻ gylynd͡ʒ-y	W		L
d. gyvend͡ʒ-i ≻ gyvent͡ʃʰ-i	L		W

With different markedness constraints regulating voicing alternations in nouns with simplex codas and complex codas, the learner can easily partition the lexicon by the complexity of the final coda. Adding the nouns with complex codas in (47) to the grammar in (46) gives rise to the more complete grammar in (48).

(48) IDENT(LAR)$_{\sigma1}$ ⟨*Vt͡ʃV, sat͡ʃʰ⟩ ⟨*VtV, atʰ⟩ ⟨*Rt͡ʃV, gønt͡ʃʰ⟩ ≫ IDENT(LAR) ⟨*Vt͡ʃV, anat͡ʃʰ⟩ ⟨*VtV, sepʰetʰ⟩ ⟨*Rt͡ʃV, gylynt͡ʃʰ⟩

≫ *Rt͡ʃV, *Vt͡ʃV, *VtV ≫

IDENT(LAR)$_{\sigma1}$ ⟨*Vt͡ʃV, tʰat͡ʃʰ⟩ ⟨*VtV, tʰatʰ⟩ ⟨*Rt͡ʃV, gent͡ʃʰ⟩ , IDENT(LAR) ⟨*Vt͡ʃV, amat͡ʃʰ⟩ ⟨*VtV, kʰanatʰ⟩ ⟨*Rt͡ʃV, gyvent͡ʃʰ⟩

The grammar in (48) allows the speaker to partition their *t͡ʃ*-final nouns by their mono- or poly-syllabicity, and within each length, by the complexity of their coda. Within each of the four kinds of t͡ʃ-final nouns, alternators are separated from non-alternators, giving the speaker access to the relative proportion of alternating nouns in each partition. The stimuli with complex codas that were used in the experiment in §3 were all mono-syllabic, and for those nouns, speakers successfully replicated the proportion of alternators from the lexicon.

Poly-syllabic nouns with complex codas were not treated separately in the statistical analyses in §2 due to their small number relative to the poly-syllabic nouns with simple codas. Of the 301 mono-syllabic nouns in TELL, the 164 nouns that have a complex coda make a respectable 54.5%. However, the 190 poly-syllabic nouns with a complex coda make a mere 7% of the 2701 poly-syllabic nouns in TELL. Consequently, poly-syllabic nouns with complex codas are not very representative of the Turkish lexicon as a whole, nor are they representative of the poly-syllabic nouns of Turkish, and therefore they were not tested in the experiment in §3. They are included in the analysis here for the sake of completeness.

5.6. Voicing alternations and k∼∅ alternations

The discussion of voicing alternations in §2 and §3 abstracted away from the fact that post-vocalic dorsals delete, rather than become voiced. The crucial observation in this context is that the voicing of stem-final stops and the deletion of stem-final dorsals are in *complementary distribution*. This is seen in (49) below, where post-vocalic dorsals either surface faithfully in the possessive (a-b) or delete (c-d), whereas post-consonantal dorsals either surface faithfully (e-f) or voice (g-h).

(49)

	bare stem	possessive	
a.	okʰ	okʰ-u	‘arrow’
b.	t͡ʃʰekʰikʰ	t͡ʃʰekʰikʰ-i	‘slanting’
c.	gøkʰ	gø-y	‘sky’
d.	t͡ʃʰilekʰ	t͡ʃʰile-i	‘strawberry’
e.	mylkʰ	mylkʰ-y	‘real estate’
f.	mehenkʰ	mehenkʰ-i	‘measure’
g.	renkʰ	reng-i	‘color’
h.	kʰepʰenkʰ	kʰepʰeng-i	‘rolling shutter’

Given a *k*-final noun in Turkish, it is not predictable whether it will surface faithfully or unfaithfully, but if it is known to surface unfaithfully, it is predictable whether the final [k] will voice (following a consonant) or delete (following a vowel). If dorsal deletion were in some sense an independent process of Turkish, its complementary distribution with respect to voicing would be left unexplained.

Both the voicing and the deletion of final dorsals show a size effect in TELL (50). While the size effect is dramatic for the post-vocalic dorsals (3% vs. 93%), there is also a noticeable size effect for the post-consonantal dorsals (10% vs. 41%).[16]

(50)

	Size	Faithful	Alternating	% alternating
Deletion	mono-syllabic	42	1	3%
	poly-syllabic	79	1048	93%
Voicing	mono-syllabic	45	5	10%
	poly-syllabic	19	13	41%

The deletion of a final dorsal does not violate IDENT(LAR), but rather violates MAX, a faithfulness constraint that penalizes deletion. To learn the size effect, the learner will need to use the general MAX and the specific MAX$_{\sigma 1}$, which penalizes the deletion of material from the initial syllable of the stem.

The complementary distribution of voicing alternation and dorsal deletion is apparent from the summary of the ranking arguments, exemplified with mono-syllabic nouns in (51). There is a conflict between IDENT(LAR)$_{\sigma 1}$ and *RkV, and there is a separate conflict between MAX$_{\sigma 1}$ and *VkV. The learner is free to discover each conflict separately.

---

[16] The size effect is highly significant in both cases, as determined by the Fisher exact test. For the post-vocalic dorsals: odds ratio = 542, $p < .0001$; for the post-consonantal dorsals: odds ratio = 6, $p < .005$.

(51)

	Ident$_{\sigma1}$	*RkV	Max$_{\sigma1}$	*VkV
a. mylk^{h}-y $\succ$ mylg-y	W	L		
b. reng-i $\succ$ renk^{h}-i	L	W		
c. ok^{h}-u $\succ$ o-u			W	L
d. gø-y $\succ$ gøk^{h}-y			L	W

If Ident$_{\sigma1}$ is cloned first, Ident(Lar)$_{\sigma1\,\mathit{mylk}^{h}}$ will be installed, followed by the installation of *RkV. Then, either Max$_{\sigma1}$ or *VkV will need to be cloned. If Max$_{\sigma1}$ is cloned, the resulting grammar will be as in (52).

(52) $\text{Ident}_{\sigma1\,\langle\text{*RkV, mylk}^{\text{h}}\rangle} \gg \text{*RkV} \gg \text{Max}_{\sigma1\,\langle\text{*VkV, ok}^{\text{h}}\rangle} \gg \text{*VkV}$
$\gg \text{Ident}_{\sigma1\,\langle\text{*RkV, renk}^{\text{h}}\rangle}, \text{Max}_{\sigma1\,\langle\text{*VkV, gøk}^{\text{h}}\rangle}$

Equivalently, If Max$_{\sigma1}$ is cloned first, followed by the cloning of Ident(Lar)$_{\sigma1}$, the resulting grammar, in (53), is just as good as the grammar in (52) in accounting for the available data.

(53) $\text{Max}_{\sigma1\,\langle\text{*VkV, ok}^{\text{h}}\rangle} \gg \text{*VkV} \gg \text{Ident(Lar)}_{\sigma1\,\langle\text{*RkV, mylk}^{\text{h}}\rangle} \gg \text{*RkV}$
$\gg \text{Max}_{\sigma1\,\langle\text{*VkV, gøk}^{\text{h}}\rangle}, \gg \text{Ident(Lar)}_{\sigma1\,\langle\text{*RkV, renk}^{\text{h}}\rangle}$

Since the deleting dorsals and the voicing dorsals are in complementary distribution, and controlled by separate constraints, it doesn't matter which trend leads to cloning first.

6. General-purpose learning with the MGL

The Minimal Generalization Learner (MGL) of Albright & Hayes (2002, 2003, 2006) is an information-theoretic algorithm that generalizes patterns over classes of words that undergo similar alternations. MGL provides a reflection of trends in the lexicon and has the potential to generalize them to novel outputs. The MGL has been shown to successfully model humans' experimental results in novel word-formation tasks with the past tense in English and with similar tasks in other languages, and is thus a good representative of a class of models that access lexical patterns without any bias against generalizing from phonologically unnatural trends.

The MGL works by reading in pairs of surface forms that are morphologically related, such as a bare noun and its possessive form in Turkish, creating a rule for each pair, and then generalizing over those rules to make more general rules. These more general rules can be applied to novel bare nouns, giving a set of possible derived forms with a confidence score assigned to each. The MGL's operation is exemplified in (54) below. Two alternating nouns, *k^hebap^h* 'kebab' and *ʃarap^h* 'wine' are read, and a rule is projected from each (54a,b). The MGL identifies the structural change in each paradigm ([p^h] becomes [bɨ]), and the environment in which this change occurs (which in its most specific instance is the entire remainder of each word). Each rule has a narrow scope, as it applies to the paradigm of a single alternation word. In order to make a generalization, the MGL compares all the rules it has and finds pairs of rules that share the same structural change. Given a set of rules with the same structural change, the algorithm compares the immediate environments for the change, and projects a new, more general rule (54c). The new rule has a wider scope (as can be seen in the example, where it will apply to any polysyllabic noun that ends in *ap^h*) but its success rate is lower, since it will mistakenly apply to non-alternating nouns that end in *ap^h*. This tradeoff between scope and accuracy is balanced by calculating adjusted confidence scores for each postulated rule.

(54) A minimal generalization in the MGL. The final subscript "2" annotates the number of syllables in the base.

	paradigm	change	environment
a.	ʃarap$^h{}_2$ ∼ ʃarabɨ$_2$	p^h → bɨ /	ʃ a r a $\underline{\quad}_2$
b.	k^{h}ebap$^h{}_2$ ∼ k^{h}ebabɨ$_2$	p^h → bɨ /	k^h e b a $\underline{\quad}_2$
c.		p^h → bɨ /	X a $\underline{\quad}_2$

As the MGL begins with a separate rule for every alternating word in the language and gradually collapses these into a more general rule based on their reliability, the question is whether it would converge upon general rules of alternation based on size, place, and vowel quality factors.

6.1. Materials and method

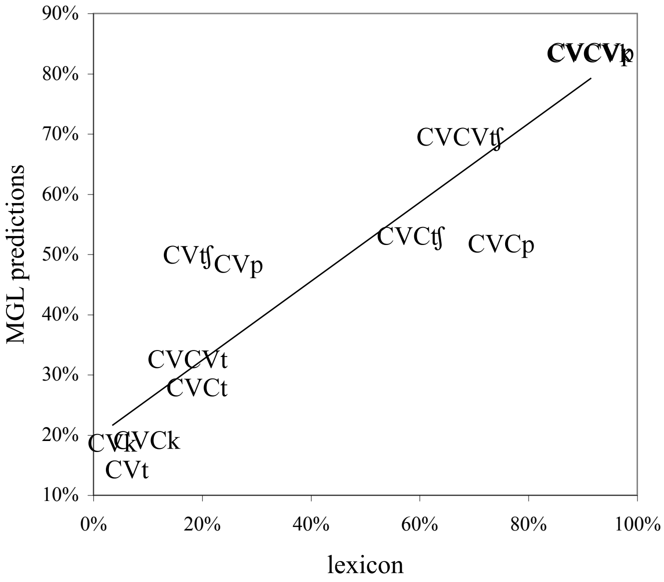
To simulate the behavior of the human participants as described in the experiment in §3, the MGL was provided with all the stop-final words in TELL as training data, and with the stimuli of the experiment as test items. Since the MGL is built to discover generalizations locally over a small span of segments, bases and their possessive forms were annotated at their right edge with the mono-syllabic status of the base, to allow the MGL to discover the size effect locally, at the site of affixation. In addition, the MGL received a feature matrix of the consonants and vowels of Turkish, which it uses to find natural classes. The results reported here were obtained by running the MGL at the 75% confidence level, which is the level that generated the results that most closely matched the human results.

For each test item, the MGL generated alternating and non-alternating possessive forms, each form associated with a confidence score, which represents the likelihood of getting that response from a human. To calculate the proportion of alternating responses that the MGL predicts, the confidence score of each alternating response was divided by the sum of the confidence scores of the alternating and non-alternating responses. For example, given the noun *fat*^{h}, the MGL produced the form *fat*^{h}*-ɨ* with a confidence of 87% and the form *fad-ɨ* with a confidence of 23%. The predicted alternation rate for *fat*^{h} was calculated as 23%/(23%+87%) = 21%. Thus, the MGL predicted alternation rates for each of the 72 test items of the experiment.

6.2. Results

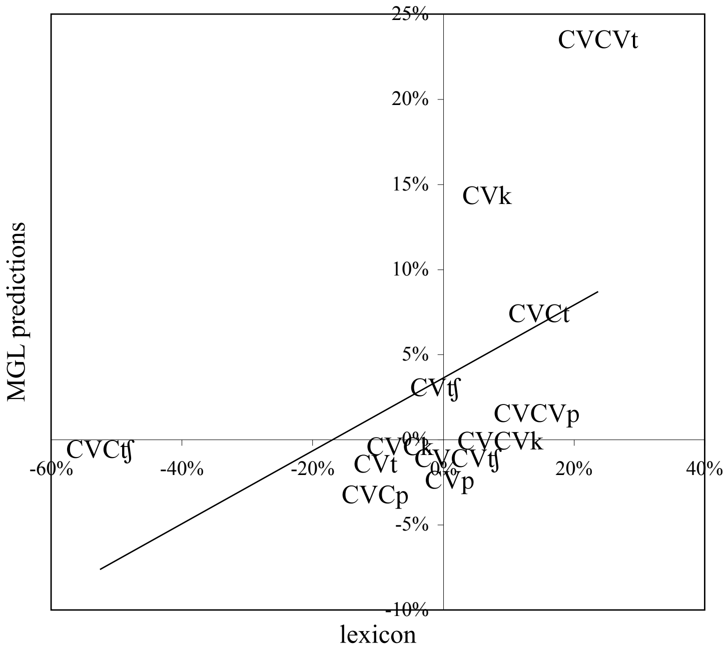
The chart in (55) shows MGL's prediction for the nonce words used in the experiment, grouped by size vs. place, plotted against the proportion of alternating words in TELL in the corresponding size and place. The MGL predictions matches the lexicon very well (Spearman's rank correlation test, S = 18, ρ = .937, p < .001). In fact, the MGL's correlation with the lexicon is a little better than the correlation of the experimental results with the lexicon (compare with 15 above).

(55) Rates of alternation in the lexicon, by place and size, plotted against the percentage of alternating responses predicted by the Minimal Generalization Learner.



The MGL prediction match the lexicon for the height effect as well, as shown in (56), with significant correlation (Spearman's rank correlation test, $S = 92$, $\rho = .678$, $p < .05$). This contrasts sharply with the lack of correlation between the lexical statistics and the experimental results (see 16 above).

(56) The difference in rates of alternation between high and non-high vowels, by size and place, in the lexicon and in the MGL results.



6.3. Discussion

The MGL's impressive performance in matching the lexical trends of Turkish voicing alternations were to its detriment. In out-performing the participants of the experiment described in §3, it failed to mimic human behavior.

The MGL is a powerful learner for phonological patterns. Given nothing but a list of paradigms and the natural classes that the segments in it form, it learned that Turkish has voicing alternations and that there are factors that are correlated with their distribution. However, since the MGL lacks a theory of possible interactions between phonological elements, it could not ignore the predictive power of vowel height and backness in determining the alternating or non-alternating status of attested nouns, and it used all the correlations it found in predicting the status of novel forms.

Humans, we argue, are biased to ignore any effect that vowel quality might have on the voicing of a neighboring consonant. This one and the same bias is observed in two domains of linguistic investigation: In the cross-linguistic study of regular phonological phenomena, and in the language-specific study of the distribution of lexically-determined phonological processes.

The MGL results are representative of a wider range of learning algorithms, such as CART (Breiman et al. 1984), C4.5 (Quinlan 1993), or TiMBL (Daelemans et al. 2002), which use purely distributional properties of a lexicon to model human behavior. The MGL's advantage over these other models is that it isn't given a list of possible generalizations to explore in advance, but rather generates its own set of hypotheses. With models other than the MGL, the lack of vowel effect could be hard-wired by not supplying the model with information about vowel quality. Since these models are not specific to language and therefore don't have any information about natural phonological interactions, such an exercise would offer little insight into the problem at hand. The MGL simulation is informative specifically because it is given whole words to deal with, without additional information about which generalizations to attend to.

The MGL results show that a model that isn't equipped with a set of biases that determine the universal range of phonological interactions will be unable to successfully mimic human behavior and ignore accidental regularities in a lexicon.

7. Conclusions

This paper presented a study of Turkish voicing alternations that contrasted trends found in the Turkish lexicon with the knowledge that speakers have about it, showing that speakers are biased to reproduce certain trends but not others. The experimental finding, that speakers do not adopt an omnivorous model of statistical generalization when it comes to vowel-consonant interactions, fall under a more general set of conclusions about the phonetic basis for phonotactic interactions. Taken together, these results suggest a more general implication for realistic models of inductive generalization from linguistic regularities: the need for a balanced interaction between the power of tracking statistical information and the constraints of linguistically-specific filters that guide the learner's analysis and acquisition of phonotactic patterns.

7.1. Use of base rate information in deneutralization

Whether or not a stop-final noun will fall into the alternating or non-alternating class of words in Turkish is seemingly unpredictable: the unsuffixed noun stem *sop*^{h} does not alternate when a vowel-initial suffix is added, as in the possessed form *sop*^{h}*-u*, but the noun stem *d͡ʒop*^{h} does: its possessed form is *d͡ʒo**b**-u*. Given a nonce word like *zop*^{h}, in which the stem-final consonant appears at the end of the word in coda position, the distinction between alternating and non-alternating stops is neutralized, due to the process of coda devoicing in Turkish.

When a speaker is presented with the novel form *zop* and asked to form the possessive, they have to undo the neutralization caused by final devoicing, and decide whether the final stop is of the alternating or non-alternating kind. This *deneutralization* task shows a number of parallels with more general schema of *backwards blocking* inference, discussed in the literature on causal reasoning and inductive inference. In studies on backwards blocking, participants observe an outcome occurring in the presence of two potential causes (A and B). Participants observe that event A independently causes the outcome. Participants are then often less likely to judge B as the cause of the outcome. One example task in which backwards blocking inferences arise is in the “blicket detector” task of Sobel et al. (2004), in which children were introduced to a blicket-detecting machine that lights up and plays music when certain objects (blickets) are placed on it and were told that “blickets make the machine go”. In the blicket-detector backward-blocking task at hand, A and B are two blocks placed on the blicket detector together which result in the machine activating. Subsequently, object A is put on the detector alone, again resulting in activation of the machine. Children were then asked whether B was a blicket. As the detection of B’s blickethood is neutralized in the presence of A, a known blicket, the “logical” response rate of whether B is a blicket should have been a 50% rate of guesses that it was. Nonetheless, in Sobel et. al’s Experiment 3, they showed that 4-year old children were remarkably sensitive to the *base rates* of whether something was likely to be a blicket, and made use of this information in the face of the logical uncertainty of backward blocking. In this experiment, they exposed and familiarized children to a number of nonce objects before introducing them to the blicket detector. There were two conditions. In the “rare blicket” condition, 1 out of 10 of the objects that the participants were exposed to beforehand were blickets. In the “common blicket” condition, 9 out of 10 objects were blickets. The children were then presented with the same task described above: seeing two objects, A and B, seeing that A lights up the blicket detector, and seeing that A and B together light up the blicket detector. The children were then asked if B was a blicket or not. The 4-year olds categorized B as a blicket on average 25% of the time in the rare blicket setup, but 81% of the time in the common blicket setup, showing that they actively employed base rate information in the deneutralized context of B alone.

The backwards-blocking blicket detector task is highly similar in structure to the coda deneutralization task we performed with nonce words in Turkish. Participants observed an outcome (e.g. [p^{h}] in final position) which occurs in the presence of two potential causes. One potential cause is the process of coda-devoicing, and a second potential cause is if this noun falls into the non-alternating class of words with a final [p^{h}]-throughout their noun paradigm. Once it is known that the presence of A alone is sufficient to trigger the outcome (in this case, that coda devoicing exists as a regular process in Turkish), then the likelihood that B is playing any role in the outcome should logically be 50%. When Turkish speakers are presented with a word like *zop*^{h} and asked whether to judge whether the deneutralized form should be *zo**p***^{h}*-u* or *zo**b**-u*, however, they take into account the overall likelihood that a word of this shape is in the alternating class. For monosyllabic nouns with a final labial stop, there is only a 30% base rate that it will be in the alternating class. The results of the experiment reported here show that Turkish speakers can and do use this information in reasoning whether a word like *zop*^{h} should be in the alternating class.

Turkish speakers thus track and consult the base rates of alternating nouns in their lexicon that match the size and place of the noun under consideration. Similarly to the findings of Ernestus & Baayen (2003), speakers appear to be highly sensitive to lexical statistics that can aid them in informed guesses in “predicting the unpredictable” to determine how to deneutralize a potentially alternating word. Despite this sensitivity to generalizations about the effects of word size and shape on voicing, however, speakers did not consider the vowel that precedes the stem-final stop, even though their lexicon contains a statistically significant generalization about the effect of final vowels, one that a machine learning simulation had no hesitation in aggressively extending to nonce word formations.

Recall that just like the Turkish speakers, the Dutch speakers in Ernestus & Baayen (2003) ignored vowel height, but they did not ignore vowel length. Vowel length, unlike vowel height, is universally correlated with the voicing of a following stop, and thus should be learned by speakers who are biased by Universal Grammar.

7.2. Phonetic features as a basis for second-order phonotactics

We claim that speakers are attuned to certain factors and ignore others, and furthermore, that the choice is based on a principled inventory of universally possible phonological interactions. Among these are the fact that the size of a word and the place of articulation of an alternating stop are reasonable determinants of phonotactic distributions to consider in whether a stop will undergo a voicing alternation or not, but that the height or backness of a preceding vowel are factors that learners are biased against considering in tracking phonotactic generalizations.

The size effect can be traced to a well-known initial syllable effect. Cross-linguistically, initial syllables enjoy greater faithfulness, or resistance to alternation (Beckman 1998). The initial syllable plays a central role in Turkish phonology: Native Turkish nouns allow voiced codas only in the initial syllable (e.g. *ab.la* 'elder sister', *ad* 'name'), and initial syllables serve as starting points for vowel harmony. Napikoğlu & Ketrez (2006) find that children quickly master suffixal allomorphy for the aorist, which is based on syllable-count. Ketrez (2007) finds that children's metathesis errors involving labials (e.g. $k^hit^hap^h \rightarrow k^hip^hat^h$ 'book') do not occur with monosyllables (e.g. yap^h) and attributes this to protection of initial-syllable. In addition, Barnes (2001) finds significantly longer duration for initial syllables in Turkish. Hence, a predicate such as "within initial syllable" is likely to be a salient factor for Turkish learners, and thus biases attention to alternation rates correlated with this factor.

The place of articulation of stem-final stops is also very likely to influence alternation rates. Different places are known to interact differently with voicing Lisker & Abramson (1964); Ohala (1983); Volatis & Miller (1992). Specifically in Turkish, dorsal stops delete rather than undergo voicing intervocalically, supplying a cue to learners that the behavior of at least one place must be learned separately. Indeed, Nakipoğlu & Üntak (2006) show that Turkish-learning children are sensitive to the differential behavior of the different places of articulation.

By contrast to size and place, the vowel that precedes the stem-final stop is not likely to play any causal role in stop alternations, and hence we argue that learners ignore this factor. Although consonant voicing has been argued to affect vowel height in various languages, as in Canadian Raising (Chambers 1973; Moreton & Thomas 2007) and Polish (Gussmann 1980) — in many cases due to the historical development of quality alternations from a pre-existing vowel length contrast in closed syllables — there is no report of vowel height or backness inducing a change in voicing in a following obstruent.

We argue that this typological gap reflects a principled lacuna in the inventory of possible phonological interactions, and specifically that phonological grammars lack any constraint-based or rule-governed process of vowel quality affecting adjacent consonantal voicing. In fact, Moreton (2008), in an attempt to teach an artificial language pattern with height-voicing interactions (i.e. in which VC sequences were always high vowel followed by voiced consonant or nonhigh vowel followed by voiceless consonant), found that participants were biased against generalizing this pattern. Importantly, Moreton's subjects were able to learn a comparably complex vowel-to-vowel interaction, suggesting that the failure to learn the height-voicing pattern was truly due to an analytic bias.

While studies of phonotactic typology and the predictions of phonological theory make clear that relations between vowel height or vowel backness and the voicing of a following stop are not possible phonological interactions, it is not the case that all vowel-consonant interactions are disfavored in natural language; on the contrary, such interactions can be quite commonplace. For example, front high vowels force a change of the place of articulation in an adjacent obstruent consonant in a number of languages, leading to phonotactic bans against sequences such as *ti*, *si*, or *ki* as opposed to *t͡ʃi* or *ʃi*; such palatalization processes are found in Japanese, Italian, Finnish, and Korean, among many other languages (Bhat 1978; Hall & Hamann 2006). Similarly, consonants can affect the distribution of adjacent vowels, as in the case of nasalization in Brazilian Portuguese, in which a stressed vowel must be nasalized before a nasal consonant, leading to phonotactic bans against sequences such as *ana* as opposed to *ãna* (Wetzels 1997). Importantly, these cases of consonant-vowel assimilatory interactions are mediated by the fact that the phonetic feature in the consonant that triggers the change is identical to the changed feature on the vowel (or vice-versa): for example, the palatal place of articulation of high front vowels is identical to the palatal place of articulation of the consonant affected by palatalization, and the phonological representation of the Place of Articulation of [i] and [č] has been argued to be identical (Hume 1994). Similarly, nasal consonants and nasalized vowels share a common phonetic articulation, [+nasal], required in the production of sounds that allow airflow through the nose (Cohn 1993).

The cases of palatalization and nasalization discussed above are processes in which vowel-consonant interaction is mediated by a common supralaryngeal phonetic feature. There are also, in fact, cases of vowel-consonant assimilatory interactions involving laryngeal features. One such phonotactic restriction involves voicing of obstruents, in which a high tone on a vowel can affect the voicing of an adjacent the consonant (i.e. a high tone on a vowel implies voiceless consonants, or vice versa), as found in Shanghainese or Jabem (Poser 1981). However, this vowel-consonant phonotactic interaction involves a common phonetic feature in both the trigger and target as well: high tone in vowels and voicelessness in obstruents are both controlled by the laryngeal property of stiffened vocal folds (Halle & Stevens 1971).

Phonotactic interactions between vowels and consonants are thus possible and indeed quite common when the nature of the phonotactic restriction involves a phonetic feature shared by the vowel and consonant. The phonetic basis for this phonotactic interaction can be either a laryngeal feature that both the vowel and consonant share, such as stiffened vocal folds, or a supralaryngeal feature that the vowel and consonant share, such as place of articulation in the vocal tract. However, the putative interaction of vowel height with consonant voicing does not even remotely fit within this rubric: vowel height is a supralaryngeal feature, consonant voicing is a laryngeal feature, and the two have thus virtually nothing to do with each other either phonetically or in terms of their phonological representations.

The *same-feature constraint* on vowel-consonant interactions is thus an “overhypotheses” in the sense of Goodman (1955) and Kemp et al. (2007): a meta-level hypotheses that constrains the form of possible specific hypotheses and generalizations induced from the data. Whether or not the same-feature constraint on vowel-consonant phonotactics is innate, or perhaps itself induced in parallel, e.g. through use of a hierarchical Bayesian model (Good 1980; Kemp et al. 2007), is not something that our experimental results speak to directly, but is an important question for modeling how it is that the vowel-quality / obstruent voicing phonotactic of Turkish is ignored.

7.3. Prior analytic biases filter statistical regularities

A number of current phonological theories adopt a constrained theory of possible phonological processes. Optimality Theory posits a universal inventory of possible phonological interactions that can be expressed as the result of the interactions among a universal set of constraints (see Kager 1999a; McCarthy 2002). Parametric models of phonological rules express constraints on what can be a possible phonological interaction as a property of the space created by a given parametric system (e.g. Dresher & Kaye 1990; Archangeli & Pulleyblank 1994; Cho 1999). Both the theories of universal constraint families and the theories of parameterized rules of assimilation require that the feature dictating a vowel-consonant interaction must be shared by both the consonant and the vowel. These models thus adopt a specific set of analytic biases, often called Universal Grammar, that the language learner brings to the task of extracting phonotactic generalizations from the lexicon, and that constrain possible generalizations that learners will make. The possibility of consonant voicing being determined or affected by vowel height or vowel backness is excluded, or highly disfavored to the point that even significant evidence for such a relationship in the lexicon is not enough. Computational modeling studies of phonological rule induction have converged on the conclusion that abstract learning biases lead to more compact, more accurate, and more general finite-state transducers for generating morphophonemic alternations (Gildea & Jurafsky 1996).

If these phonetically-unmotivated patterns are never used and in fact excluded or disfavored by learning biases, why do exist in the Turkish lexicon in the first place? The existence of a statistically significant trend for high vowels or for back vowels to be followed by alternating voiced stops in the Turkish lexicon is arguably tied to the fact that the Turkish lexicon represents an accumulation of several centuries worth of language contact. Many of the lexical trends that identified in our quantitative lexicon analysis are ultimately traceable to extensive lexical borrowing from Arabic, to much the same degree that many of the lexical trends found in English phonotactics, such as the existence of more words that begin with [d͡ʒ] than [ʒ], are ultimately traceable to lexical borrowing from French centuries ago, when Old French had [d͡ʒ] but not [ʒ] word-initially. In Turkish borrowings of words with voiced stops in the source language, final devoicing in the bare stem but not in the forms with vowel-initial suffixes causes a noun to become alternating (e.g. Arabic *burd͡ʒ* ‘sign’ > Turkish *burt͡ʃʰ* ∼ *burd͡ʒ-u*), whereas source words that end in a voiceless stop are non-alternating across the paradigm. Arabic lacks the consonants [p] and [t͡ʃ] and has many nouns that end in [b] and [d͡ʒ], and as a consequence, the lexicon’s overall alternation rates are boosted for those places of articulation. On the other hand, the existence of many Arabic nouns with feminine suffix *-at/-et* boosted the number of non-alternating, non-high vowel, coronal-final nouns. Ultimately, however, the historical explanation for these lexical trends is completely inaccessible to speakers that are not experts in historical linguistics, many of whom (like the English speakers who know the word *judge* but not its origin), do not even know that there was a source language that provided this borrowed word, well-integrated into the phonotactics for centuries.

If indeed the skewed distribution of the Turkish voicing alternations is largely due to massive borrowing from Arabic, it is instructive that Turkish speakers synchronically generalize the historically accidental place effect, but discard the equally accidental height effect. History has dealt Turkish speakers a certain hand, and they use Universal Grammar to pick the cards they want to keep. This view contrasts with the proposal in Hayes (1999), who claims that when history creates non-Universal patterns, speakers are able to complement their Universal Grammar with arbitrary generalizations.

In Turkish, the distribution of voicing alternations is not known to correlate with the native or borrowed status of roots (and as mentioned in the introduction, loanwords such as *group* $>$ *guru**b**-u* conform to the polysyllabic-as-alternating generalization). Thus, the sources of some of the unprincipled statistical regularities are arguably historical in nature, yielding phonetically-ungrounded synchronic patterns that are simply ignored.

The result that Turkish speakers reliably extend base rates for voicing alternations based on place of articulation and size of the word, but not based on preceding vowel quality, arguably due to an analytic bias against learning such arbitrary interactions, strengthens the finding of Moreton (2008) that English speakers were less successful learning an artificial language pattern with height-voicing interactions, and more successful learning non-adjacent V-V interactions, in which high vowels were followed by high vowels in the adjacent syllable. In Turkish, the case is even more striking: a lexical generalization is staring Turkish speakers in the face, but they do not generalize it productively in experimental contexts. The results provide support for an analytically-biased mechanism of filtering lexical statistics, one in which phonologically-implausible interactions are not actively incorporated into phonotactic knowledge. There is by now a general consensus that statistical information is indispensable in arriving at phonotactic generalizations, a fact which our experimental results confirm. At the same time, accurate models of the acquisition of phonological knowledge need to build in a set of linguistically-specific priors that constrain and restrict the learning of statistical patterns. Apparently, given a surfeit of the stimulus, not every statistical fact about the lexicon is used or kept track of.

We proposed a learning model which consists of identifying conflicting lexical patterns in the lexicon, resolving the conflict by cloning constraints. Once constraints are cloned, each clone keeps a list of the words it governs, assuring that existing words behave consistently. At the same time, the clones can be used in a generalized way, referring only to the *pro-*

portion of words that are governed by each clone, to project the lexical trend onto novel words.

The resulting learner simulated the process of learning a lexicon without relying on general-purpose pattern matching. Rather, such statistical patterns were filtered though a proposed set of universal constraints that were augmented by the ability to clone constraints. These 'priors' on what data is to be used in forming grammatical hypotheses implicate an analytic bias that, in this case, ignored the correlation between vowel quality and consonant voicing thanks to the absence of constraints that relate the two, thus closely modeling the pattern produced by native speakers.

References

Albright, Adam (2008). A Restricted Model of UR Discovery: Evidence from Lakhota. Ms. MIT.

Albright, Adam, Argelia Andrade & Bruce Hayes (2001). Segmental environments of Spanish diphthongization. In Adam Albright & Taehong Cho (eds.) *UCLA Working Papers in Linguistics 7 (Papers in Phonology 5)*, UCLA. 117–151.

Albright, Adam & Bruce Hayes (2002). Modeling English past tense intuitions with minimal generalization. In Michael Maxwell (ed.) *Proceedings of the sixth meeting of the ACL special interest group in computational phonology*. Philadelphia: ACL, 58–69.

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

Albright, Adam & Bruce Hayes (2006). Modeling productivity with the gradual learning algorithm: The problem of accidentally exceptionless generalizations. In Gisbert Fanselow, Caroline Féry, Matthias Schlesewsky & Ralf Vogel (eds.) *Gradience in Grammar*, Oxford University Press. 185–204.

Anttila, Arto (2002). Morphologically conditioned phonological alternations. *Natural Language and Linguistic Theory* **20**. 1–42.

Archangeli, Diana & Douglas Pulleyblank (1994). *Grounded Phonology*. MIT Press.

Avery, Peter (1996). *The Representation of Voicing Contrasts*. Ph.D. dissertation, University of Toronto.

Baayen, Harald (2008). *Analyzing Linguistic Data: A practical introduction to statistics*. Cambridge University Press.

Bailey, Todd & Ulrike Hahn (2001). Determinants of wordlikeness: phonotactics or lexical neighborhoods. *Journal of Memory and Language* **44**. 568–591.

Barnes, Jonathan (2001). Domain-initial strengthening and the phonetics and phonology of positional neutralization. Paper presented at the Northeast Linguistics Society Meeting, CUNY.

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

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

Beckman, Jill (1998). *Positional Faithfulness*. Ph.D. dissertation, UMass, Amherst.

Beckman, Jill N. & Catherine O. Ringen (2004). Contrast and redundancy in OT. In Vineeta Chand, Ann Kelleher, Angelo J. Rodríguez & Benjamin Schmeiser (eds.) *Proceedings of WCCFL 23*, Cascadilla Press. 85–98.

Berko, Jean (1958). The child's learning of english morphology. *Word* **14**. 150–177.

Bhat, D.N.S. (1978). A general study of palatalization. In Joseph H. Greenberg (ed.) *Universals of human language, vol. 2: phonology,*, Stanford, CA: Stanford University Press. 47–92.

Blust, Robert (2000). Low-vowel fronting in northern sarawak. *Oceanic Linguistics* **39**. 285–319.

Boersma, Paul (2001). Phonology-semantics interaction in ot, and its acquisition. In Robert Kirchner, Wolf Wikeley & Joe Pater (eds.) *Papers in Experimental and Theoretical Linguistics*, University of Alberta, vol. 6. 24–35.

Boersma, Paul & David Weenink (2008). Praat: Doing phonetics by computer (version 5.0.24). [Computer program].

Bonatti, Luca, Marcela Peña, Marina Nespor & Jacques Mehler (2005). Linguistic constraints on Statistical Computations. *Psychological Science* **16.6**. 451–459.

Breiman, Leo, Jerome Friedman, R. A. Olshen & Charles J. Stone (1984). *Classification and regression trees*. Belmont, CA: Wadsworth International Group.

Buder, E. & C. Stoel-Gammon (2002). American and swedish children's acquisition of vowel duration: Effects of vowel identity and final stop voicing. *Journal of the Acoustic Society of America* **111**. 1854–1864.

Burzio, Luigi (2002). Surface-to-Surface Morphology: when your representations turn into constraints. In P. Boucher (ed.) *Many Morphologies*, Cascadilla Press. 142–177.

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

Chambers, J.K. (1973). Canadian Raising. *Canadian Journal of Linguistics* **18**. 113–135.

Chater, Nick & Christopher D. Manning (2006). Probabilistic models of language processing and acquisition. *Trends in Cognitive Sciences* **10.7**. 335–344.

Chen, M. (1970). Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* **22**. 125–159.

Cho, Young-mee Yu (1999). *Parameters of Consonantal Assimilation*. Munich: Lincom Europa.

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

Cohn, Abigail (1993). *Phonetic and Phonological Rules of Nasalization*. Ph.D. dissertation, UCLA.

Coleman, John & Janet Pierrehumbert (1997). Stochastic phonological grammars and acceptability. In *Computational Phonology, Third Meeting of the ACL Special Interest Group*, Somerset, NJ: Association for Computational Linguistics. 49–56.

Daelemans, Walter, Jakub Zavrel, Ko Van Der Sloot & Antal Van Den Bosch (2002). Timbl: Tilburg memory based learner reference guide, version 4.2. Tilburg: Computational Linguistics, Tilburg University.

Denes, P. (1955). Effect of duration on the perception of voicing. *Journal of the Acoustic Society of America* **27**. 761–764.

Dresher, B. Elan & Jonathan Kaye (1990). A computational learning model for metrical phonology. *Cognition* **34.2**. 137–195.

Dresher, B. Elan & Harry van der Hulst (1998). Head-dependent asymmetries in phonology: Complexity and visibility. *Phonology* **15**. 317–352.

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

Flack, Kathryn (2007). *The sources of phonological markedness*. Ph.D. dissertation, UMass Amherst.

Gildea, Daniel & Daniel Jurafsky (1996). Learning Bias and Phonological-Rule Induction. *Computational Linguistics* **22.4**. 497–530.

Good, I.J. (1980). Some history of the hierarchical Bayesian methodology. In J.M. Bernardo, M.H. DeGroot, D.V. Lindley & A.F.M. Smith (eds.) *Bayesian statistics*,, Valencia University Press. 489–519.

Goodman, Nelson (1955). *Fact, Fiction, and Forecast*. Harvard University Press.

Gussmann, Edmund (1980). *Studies in Abstract Phonology*. Cambridge, MA: MIT Press.

Hall, Tracy & Silke Hamann (2006). Towards a typology of stop assibilation. *Linguistics* **44.6**. 1195–1236.

Halle, Morris & Kenneth N. Stevens (1971). A note on laryngeal features. *MIT Quarterly Progress Report* **11**. 198–213.

Hay, Jennifer & Harald Baayen (2005). Shifting paradigms: gradient structure in morphology. *Trends in Cognitive Sciences* **9**. 342–348.

Hay, Jennifer, Janet Pierrehumbert & Mary Beckman (2004). Speech perception, well-formedness and the statistics of the lexicon. In J. Local, R. Ogden & R. Temple (eds.) *Phonetic Interpretation: Papers in Laboratory Phonology VI*, Cambridge University Press. 58–74.

Hayes, Bruce (1995). On what to teach the undergraduates: Some changing orthodoxies in phonological theory. In Ik-Hwan Lee (ed.) *Linguistics in the Morning Calm 3*, Seoul: Hanshin. 59–77.

Hayes, Bruce (1999). Phonological Restructuring inYidiɲ and its Theoretical Consequences. In Ben Hermans & Marc van Oostendorp (eds.) *The derivational residue in phonology*, Amsterdam: Benjamins. 175–205.

Hayes, Bruce (2004). Phonological acquisition in optimality theory: the early stages. In René Kager, Joe Pater & Wim Zonneveld (eds.) *Constraints in Phonological Acquisition*, Cambridge: Cambridge University Press. 158–203.

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

Hayes, Bruce & Colin Wilson (2008). A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* **39**. 379–440.

Hume, Elizabeth (1994). *Front Vowels, Coronal Consonants and their Interaction in Nonlinear Phonology*. New York: Garland.

Inkelas, Sharon, Aylin Kuntay, John Lowe, Orhan Orgun & Ronald Sprouse (2000). Turkish electronic living lexicon (TELL). Website, http://socrates.berkeley.edu:7037/.

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

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

Itô, Junko & Armin Mester (1995). The core-periphery structure in the lexicon and constraints on re-ranking. In Jill Beckman, Laura Walsh Dickey & Suzanne Urbanczyk (eds.) *Papers in Optimality Theory*, GLSA. 181–210.

Iverson, Gregory & Joseph Salmons (2007). Domains and directionality in the evolution of german final fortition. *Phonology* **24**. 121–145.

Jannedy, Stefanie (1995). Ohio State University Working Papers in Linguistics. *Journal of Linguistics* **45**. 56–84.

Jarosz, Gaja (2006). *Rich Lexicons and Restrictive Grammars - Maximum Likelihood Learning in Optimality Theory*. Ph.D. dissertation, Johns Hopkins University.

Kager, René (1999a). *Optimality Theory*. Cambridge University Press.

Kager, René (1999b). Surface opacity of metrical structure in optimality theory. In Ben Hermans & Marc van Oostendorp (eds.) *The Derivational Residue in Phonological Optimality Theory*, Amsterdam: John Benjamins. 207–245.

Kallestinova, Elena (2004). Voice and aspiration of stops in Turkish. In Grzegorz Dogil (ed.) *Folia Linguistica 38: Special Issue on Voice*, Berlin: Mouton de Gruyter. 117–143.

Keating, Patricia (1985). Universal phonetics and the organization of grammars. In V. A. Fromkin (ed.) *Phonetic Linguistics*, Academic Press. 115–132.

Kemp, Charles, Amy Perfors & Joshua Tenenbaum (2007). Learning overhypotheses with hierarchical Bayesian models. *Developmental Science* **10.3**. 307–321.

Ketrez, Nihan (2007). Alignment versus linearity constraints in a Turkish child's speech. *Dilbilim Araştırmaları* .

Kingston, John (2002). Keeping and losing contrasts. In *Proceedings of the 28th Annual Meeting*, Berkeley Linguistics Society. 155–176.

Kirchner, Robert (1998). *An Effort-Based Approach to Consonant Lenition*. Ph.D. dissertation, UCLA.

Kopkallı, Handan (1993). *A phonetic and phonological analysis of final devoicing in Turkish*. Ph.D. dissertation, University of Michigan.

Lees, Robert (1961). *The Phonology of Modern Standard Turkish*. Bloomington: Indiana University Press.

Lisker, Leigh & Arthur Abramson (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word* **20**. 384–422.

McCarthy, John J. (2002). *A Thematic Guide to Optimality Theory*. Cambridge University Press.

Merchant, Nazarré (2008). *Discovering Underlying Forms: Contrast Pairs and Ranking*. Ph.D. dissertation, Rutgers University.

Moreton, Elliott (2008). Analytic bias and phonological typology. *Phonology* **25**. 83–127.

Moreton, Elliott & Erik Thomas (2007). Origins of Canadian Raising in voiceless-coda effects: a case study in phonologization. In Jennifer Cole & José I. Hualde (eds.) *Papers in Laboratory Phonology 9*, Mouton de Gruyter. 37–64.

Nakipoğlu, Mine & Aslı Üntak (2006). What does the acquistion of stems that undergo phonological alternation reveal about rule application. Paper presented at the International Conference on Turkish Linguistics, Uppsala, Sweden.

Napikoğlu, Mine & Nihan Ketrez (2006). Children's overregularizations and irregularizations of the Turkish aorist. In *BUCLD 30: Proceedings of the Boston University Conference on Language Development, Volume 2*, Somerville, MA: Cascadilla Press. 399–410.

Ohala, John (1983). The origin of sound patterns in vocal tract constraints. In *The production of speech*, New York: Springer-Verlag. 189–216.

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

Pater, Joe (2008). Morpheme-specific phonology: Constraint indexation and inconsistency resolution. In Steve Parker (ed.) *Phonological Argumentation: Essays on Evidence and Motivation*, Equinox. 1–33.

Peterson, G. E. & I. Lehiste (1960). Duration of syllable nuclei in english. *Journal of the Acoustic Society of America* **32**. 693–703.

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

Poser, William J. (1981). On the directionality of the tone-voice correlation. *Linguistic Inquiry* **12**. 483–488.

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

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

Prince, Alan & Bruce Tesar (1999). Learning phonotactic distributions. Technical Report RuCCS-TR-54, Rutgers Center for Cognitive Science, Rutgers University, New Brunswick [Available as ROA-353].

Quinlan, J. R. (1993). *C4.5: Programs for Machine Learning*. San Mateo, CA: Morgan Kaufman.

R Development Core Team (2007). R: A language and environment for statistical computing. Vienna, Austria. (http://www.R-project.org).

Saffran, Jenny R. (2003). Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science* **12.4**. 110–114.

Saffran, Jenny R. & Erik D. Thiessen (2003). Pattern induction by infant language learners. *Developmental Psychology* **39.3**. 484–494.

Sezer, Engin (1981). The k/∅ alternation in Turkish. In *Harvard Studies in Phonology*, Bloomington: Indiana University Linguistics Club. 354–382.

Sobel, David, Joshua Tenenbaum & Alison Gopnik (2004). Children's causal inferences from indirect evidence: Backwards blocking and Bayesian reasoning in preschoolers. *Cognitive Science* **28**. 303–333.

Stevens, Alan M. (1968). *Madurese Phonology and Morphology. American Oriental Series, #52*. New Haven: American Oriental Society.

Tesar, Bruce (1998). Using the mutual inconsistency of structural descriptions to overcome ambiguity in language learning. In P. Tamanji & K. Kusumoto (eds.) *Proceedings of NELS 28*. Amherst, MA: GLSA, 469–483.

Tesar, Bruce (2006). Learning from paradigmatic information. In *Proceedings of NELS 36*.

Tesar, Bruce & Alan Prince (2006). Using phonotactics to learn phonological alternations. In *CLS 39*. Available as ROA-620.

Tesar, Bruce & Paul Smolensky (1998). Learnability in optimality theory. *Linguistic Inquiry* **29**. 229–268.

Tesar, Bruce & Paul Smolensky (2000). *Learnability in Optimality Theory*. Cambridge, MA: MIT Press.

Tessier, Anne-Michelle (2007). *Biases and stages in phonological acquisition*. Ph.D. dissertation, University of Massachusetts, Amherst.

Vaux, Bert & Bridget Samuels (2005). Laryngeal markedness and aspiration. *Phonology* **22**. 395–436.

Volatis, Lydia & Joanne Miller (1992). Phonetic prototypes: Influence of place of articulation and speaking rate on the internal structure of voicing categories. *J. Acoustic Soc. Am.* **92.2**. 723–735.

Warker, Jill & Gary Dell (2006). Speech errors reflect newly learned phonotactic constraints. *Journal of Experimental Psychology: Learning, Memory and Cognition* **32.2**. 387–398.

Wetzels, Leo (1997). The Lexical Representation of Nasality in Brazilian Portuguese. *Probus* **9.2**. 203–232.

Wilson, Stephen (2003). A phonetic study of voiced, voiceless and alternating stops in Turkish. *CRL Newsletter* **15**. 3–13. URL http://stephenw.bol.ucla.edu/papers/turkishphon.pdf.

Zimmer, Karl & Barbara Abbott (1978). The k/∅ alternation in Turkish: Some experimental evidence for its productivity. *Journal of Psycholinguistic Research* **7**. 35–46.

Zoll, Cheryl (1996). *Parsing below the segment in a constraint-based framework*. Ph.D. dissertation, UC Berkeley.

Zuraw, Kie (2000). *Patterned Exceptions in Phonology*. Ph.D. dissertation, UCLA.