# The Surfeit of the Stimulus: Analytic biases filter lexical statistics in Turkish devoicing neutralization

Michael Becker
UMass Amherst

Nihan Ketrez
Yale University

Andrew Nevins
Harvard University

**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. This finding can be interpreted 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.

## Introduction

### *Constraints on Learning of Statistical Regularities*

It is clear that 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 (see Saffran (2003), Hay and Baayen (2005), Chater and 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, Peña, Nespor, and Mehler (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 and 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 the current 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 child and 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.

## *Background on Turkish Phonotactics*

Turkish has a rich set of phonotactic restrictions, yielding altenations in consonants and vowels in morphologically-derived environments. These phonotactic restrictions include vowel harmony and consonantal devoicing.

Turkish has an eight vowel system (Table 1), with all vowels allowed in word-initial syllables. Principles of vowel harmony restrict the distribution of non-initial vowels in the native vocabulary (Lees, 1961). Backness harmony applies for all vowels: each vowel must agree in its feature of [± back] with the vowel that precedes it.

In addition, [+high] vowels are required to agree in rounding with preceding vowels. Non-high round vowels can are restricted largely to the initial syllable of a root, and non-high vowels do not undergo rounding harmony with preceding vowels. This effectively restricts the vowels [o,ö] to word-initial positions in the native vocabulary, and in addition provides evidence that the phonology of Turkish treats high vowels and non-high vowels as two distinct natural classes.

| | −back | | +back | |
|---|---|---|---|---|
| | −round | +round | −round | +round |
| +high | i | ü | ı | u |
| −high | e | ö | a | o |

Table 1: Turkish vowel system

Stress in Turkish is largely word-final, with certain classes of exceptions, mostly in place names (Sezer (1981b), Inkelas (1999)). The phonological size of words is measured not in terms of number of segments, but rather in moras – units of phonological weight that count both vowels and coda consonants (Hayes, 1989). For example, Turkish nouns are required to minimally have either a short vowel followed by a consonant, with each of those contributing one mora, or a long vowel that contributes two moras (Inkelas & Orgun, 1995). The irrelevance of onset consonants to weight means that a CVC word is 'just as long' from the point of view of phonology as a VC or CCVC word, all being bimoraic.

Turkish exhibits a contrast between the voiced stops [b, d, ǰ, g] and the voiceless stops [p, t, č, k] in onset position, e.g. *ter* 'sweat' vs. *der* 'give-aorist'. In coda position, however, the contrast is lost, with voiced stops becoming devoiced through complete phonetic neutralization (Kopkallı (1993); Wilson (2003)). This restriction on the distribution of voiced stops applies productively to loanwords, e.g. *rop* '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 suffixed with a vowel-initial morpheme, the final stop may surface with its voiced counterpart, e.g. *jop* 'club' vs. the possessed form *jo**b**-u* 'club.3SG'; however, when suffixed with a consonant-initial morpheme, the final stop remains in coda position and is thus subject to devoicing: *jop-lar* 'club.plural'. This alternation is occurs in 54% of the nouns of the language (Inkelas, Küntay, Lowe, Orgun, & Sprouse, 2000), and applies productively to loanwords, e.g. *gurup* 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. *sop*, *sop-u* 'clan.3SG', *sop-lar* 'clan.plural'. While it is unpredictable, or at least partly so (as we discuss below) whether a noun will alternate or not, it is a fact that for the nouns that show voicing alternations, all vowel-initial suffixes will trigger the alternation.

The velar stops [k,g] may contrast in onset positions, e.g. *so.**k**ak* 'street' vs. *ga.**g**a* 'beak' . In word-final position, they neutralize to the voiceless stop [k]. While alternations such as *renk* vs. *ren**g**-i* 'color.3SG' display the general process of voicing alternation, there is an additional process affecting intervocalic velar stops undergoing suffixation: when nouns ending in *postvocalic* velar stop are suffixed with a vowel-initial morpheme, the velar stop deletes e.g. *ete**k**/ete-i* 'skirt.3SG' (Zimmer and Abbott (1978), Sezer (1981a)). It is possible to analyze intervocalic *g*-deletion as a subcase of the general process of voicing alternation discussed above; in fact, in the orthography, the deleted stop is represented by a 'soft ğ', e.g. <eteği>; in addition, not all dialects contain this additional process of intervocalic g-deletion. As will be shown below, whether a noun stem shows the $k/\emptyset$ 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.

While the stem-final stop of some nouns alternates between voiceless and voiced under affixation (e.g. *jop* vs. *job-u* 'club'), others do not, remaining voiceless throughout, e.g. *sop*, *sop-u* 'clan'. This distinction is traditionally captured within generative phonology as the difference between an underlying voiced stem-final stop in the case of *job-u* and an underlying voiceless stem-final stop in the case of *sop-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)), 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.

An important question throughout studies of phonological alternation is the extent to which the membership of a noun in the alternating or non-alternating class is predictable. For instance, in Dutch, a very similar voicing alternation holds for root-final obstruents, due to the existence of coda-devoicing. Ernestus and Baayen (2003) looked at the sublexical factors governing the process of apparently unpredictable alternation of stop voicing between word-final and intervocalic positions in Dutch verbs, e.g. vɛrʋɛit 'widen-3sg.' vs. vɛrʋɛid-ən 'widen-inf'. Ernestus and Baayen found a number of factors that were statistically reliable predictors of voicing alternation: the identity of the obstruent (i.e. its place of articulation), the number of syllables in the word, the stress placement of the word, the length of the vowel preceding the obstruent, and the identity of the onset consonant beginning the final stem syllable. In the Turkish literature, a number of trends influencing whether a noun will be in the alternating or non-alternating class for stop-voicing have been identified. These include the number of syllables (Lewis (1967, p.10), Inkelas and Orgun (1995), Inkelas, Orgun, and Zoll (1997)), a factor that plays out in the adaptation of English loanwords into Turkish (Vaux, 2005), as well as the place of articulation of the stop (Inkelas & Orgun, 1995). We conducted a quantitative study of the phonological factors correlating with the alternation rates of Turkish stem-final stops, in order to assess the relative strength of these correlation and to examine these effects in light of phonological typology.

## Lexicon Analysis of Voicing Alternation

### *Materials and method*

To estimate the effect of a number of factors of interest in the lexicon, we examined the rates of alternation as the dependent variable in the 2899 stop-final nouns in TELL, an electronic lexicon of Turkish (Inkelas et al., 2000). We extracted the complete paradigms of stop-final nouns in TELL and classified them as alternating or non-alternating based on whether the stop became voiced in the possessed form of the noun, which is formed by the addition of a vowel-initial suffix to the noun stem. We annotated a variety of phonological properties for each noun, including phonological size, the place of articulation of the final stop, the height of the final vowel of the stem, the roundness of the final vowel of the stem, the backness of the final vowel of the stem, and the identity of the onset consonant of the final syllable. Within the category of phonological size, in addition to the basic monosyllabic/polysyllabic distinction, CVC and CVCC words were counted separately because of the observation in Inkelas and Orgun (1995) that this size difference may condition the voicing alternation. We discuss the number of alternating vs. non-alternating nouns (henceforth the *alternation rate*) for each of these factors.

### *Results*

The effect of size on alternation rates is considerable. In monosyllabic words of either the CVC or CVCC type, the dominant pattern is non-alternation, while in polysyllabic words, the dominant pattern is alternation, as seen in Table 2. In addition, within monosyllabic words, the presence of two coda consonants makes the alternation rate twice as likely; the interaction of this factor with place of articulation will be discussed below. The percentage counts for the three categories of phonological size provide quantitative confirmation of the trend noticed in earlier studies that *size* (e.g. monosyllabic or not and two coda consonants or not) is a relevant factor in conditioning alternation rates.

The place of articulation of the stop whose voicing alternation is in question has a decidedly relevant effect on alternation rates. Within the four places of articulation for

| Shape | $n$ | % alternating |
|---|---|---|
| Monosyllabic, simple coda (CVC) | 137 | 11.7% |
| Monosyllabic, complex coda (CVCC) | 164 | 25.9% |
| Polysyllabic (CVCVC and longer) | 2701 | 58.9% |

Table 2: Voicing alternation rates by Size for all nouns in TELL.

Turkish stops, the ordering of alternation is dorsal > labial > palatal > coronal, with coronal as the only place of articulation with an overall rate of alternation below 50%, as shown in Table 3.

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

Table 3: Alternation rates by Place for all nouns in TELL, as ordered from front to back by Place of Articulation.

As discussed above, the factors of place and size are predictive on their own. However, considering them jointly reveals some trends of interaction, as shown in Figure 1. The asymmetry between polysyllabic and monosyllabic nouns holds for all places of articulation. The lower rate of alternation for coronals as compared to non-coronal places of articulation holds for all three phonological sizes. The category of CVCC words shows an interesting split across places of articulation, however. CVCC words pattern with polysyllabic words in labial and palatal places of articulation, while patterning together with CVC words in the dorsal place of articulation. The existence of this interaction necessitates the treatment of CVCC as an independent phonological size for purposes of alternation rate calculations.

Turning to effects of the quality of the stem-final vowel, i.e. the vowel preceding the stop in question, vowel quality shows differential correlations for the three vowel features of height, backness, and roundness. A decidedly predictive factor in the lexicon is the *height* of the preceding vowel: when the vowel that precedes the stem-final stop is [+high], there is an overwhelming trend for alternation, as shown in Table 4.

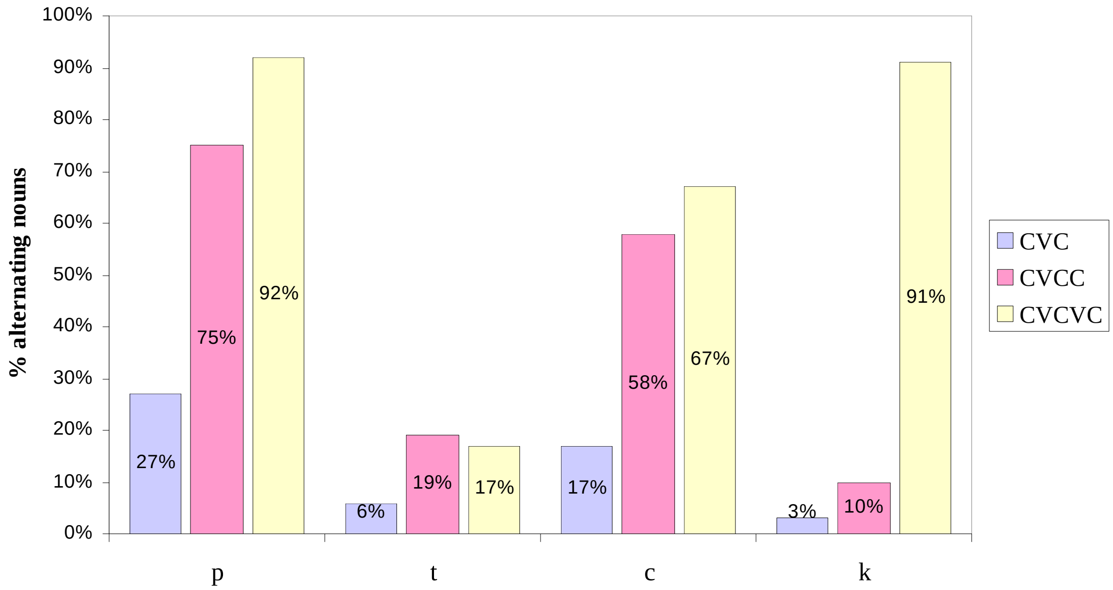


*Figure 1*. Lexicon: Place x Size for all stop-final words in TELL.

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

Table 4: Alternation rates by preceding vowel Height for all nouns in TELL.

The effect of *backness* of the preceding vowel on alternation rates is more nuanced; as opposed to the 30.2% difference in alternation rates for high and non-high vowels, there is only a 10.8% difference in alternation rates between back and front vowels, as shown in Table 5. However, a more dramatic difference in alternation rates is observed when one considers the interaction of backness with place: a difference in alternation rate of 30% emerges specifically within the palatal place of articulation, as shown in Table 6.

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

Table 5: Alternation rates by preceding vowel backness for all nouns in TELL.

| | $n$ | −back | +back |
|---|---|---|---|
| Labial (p) | 294 | 90% | 79% |
| Coronal (t) | 1255 | 16% | 18% |
| Palatal (č) | 191 | 44% | 74% |
| Dorsal (k) | 1262 | 84% | 86% |

Table 6: Alternation rates by Backness of preceding vowel and Place, for all nouns in TELL.

Finally, the two other factors that were considered in the counts of alternation rates had negligible effects on alternation rates: roundness of the preceding vowel (difference of 1.8%) and the identity of the onset consonant.

Having computed the raw percentages for alternation as organized by size, place, height, backness, roundness, and their interactions in the case of size-place and backness-place, we assessed the statistical power of these factors as predictors of alternation rates using a logistic regression analysis for all the 2899 stop-final nouns extracted from TELL. The logistic regression was performed with independent variables organized in the following order: *Place, Size, Roundness, Backness, Height*. *Place* was a four level variable, with a range from 0 to 3, corresponding to coronal, palatal, labial and dorsal. *Size* was a three level variable, with range from 0 to 2, corresponding to mono-syllables with a simplex coda (CVC), mono-syllables with a complex codas (CVCC), and poly-syllables. *Height* was a binary variable with high and non-high, *Back* was a binary variable with back and non-back, and *Roundness* was a binary variable with round and non-round. Using the forward stepwise (likelihood ratio) method of SPSS, four of these variables were entered with high significance , as shown in Table 7. *Roundness* did not reach significance, and was therefore not entered as a variable into the stepwise logistic regression.

| | $\beta$ | $\chi^2$ | $R^2$ | $p$ |
|---|---|---|---|---|
| place | 1.221 | 1379 | .506 | <.001 |
| size | 1.813 | 261 | .578 | <.001 |
| height | .525 | 87 | .601 | <.001 |
| backness | 1.151 | 22 | .606 | <.001 |

Table 7: Logistic regression results for TELL.

The results of this logistic regression confirm that the phonological factors of Place, Size, Height and Backness each independently have a statistically significant power in the prediction of whether a given stem will alternate or not.

### *Discussion*

The results of quantitative analysis of the proportions of alternating nouns and a logistic regression revealed four factors that are predictive of whether voicing alternation will occur: Phonological Size of the word, Place of Articulation, Height of the preceding vowel, and Backness of the preceding vowel. The first two of these have been previously identified as having an influence on voicing alternation in Turkish, 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**-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 a particular feature (voicing) 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 voicing (Lisker & Abramson, 1964; Ohala, 1983; Volatis & Miller, 1992), and different relative proportions of alternation rates for different places of articulation were found by Ernestus and 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 and voicing 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 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 quality and consonant voicing are infrequent, and the handful of documented cases show a causal influence in the opposite direction: the consonant's voicing can affect the height of a preceding vowel (Kingston (2002)), but not vice versa. Consonant voicing has 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 and Thomas (2005)); 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. Given the fact that vowel quality-consonant voicing interactions are second-order phonotactics with little to no crosslinguistic attestation, their existence in Turkish is unexpected and may even 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, Height and Backness 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.

## Experiment 1

This experiment sought to determine the sensitivity of the nonce-word methodology in finding correlations that one would expect in speakers' behavior based on the factors that have been identified to be relevant in predicting whether a noun will alternate in the existing phonological literature on Turkish (and confirmed quantitatively in our lexicon analysis), namely Size and Place effects. In addition, we pursued the question of whether this sensitivity would be exclusive to adult speakers or whether evidence for knowledge and use of statistical patterns from the lexicon could be found in children's productions as well.

### *Method*

#### *Participants*

Participants were native speakers of the standard (north-western) variety of Turkish and they were all residents of Istanbul. Children (n=10; 6 females, 4 males; age range: 3;6-4;3; mean age: 4;0) were recruited at a kindergarten and received a color sticker for their participation. Adults (n=12; 6 females, 6 males) were university students and volunteered to participate.

*Materials*

Materials were constructed from four places of articulation (Labial, Coronal, Palatal, Dorsal) for the potentially alternating stop and two phonological sizes (CVC and CVCVC), with 3 items in each category. The stem-final vowel preceding the potentially alternating stop was evenly distributed among the eight Turkish vowels, with an aim to avoid the vowels [o, ö] in non-initial syllables, thereby respecting phonotactic constraints of Turkish round vowel distribution discussed in the introduction. All materials were judged as Turkish-sounding by two native speakers. Each item had at least one phonological neighbor within the existing Turkish lexicon, according to the one-edit-distance phonological neighbor calculation (Luce & Pisoni, 1998). Table 8 shows the 24 experimental items. The experimental items were mixed with 12 nonce-word fillers, 6 of which were CVC and 6 of which were CVCVC. Additionally, the experiment included 16 actual Turkish stop-final words, to estimate the children's development with respect to this aspect of the adult lexicon as a control.

| | CVC | CVCVC |
|---|---|---|
| Labial | tup | yıyap |
| | dap | ǰisip |
| | zip | kunup |
| Coronal | nut | gevit |
| | hit | muyut |
| | fet | niket |
| Palatal | nıč | bölüč |
| | yüč | heveč |
| | peč | marıč |
| Dorsal | sük | banık |
| | pık | nönük |
| | vek | mesek |

Table 8: Experiment 1 Materials

*Procedure*

An elicited production task was used to elicit nouns with the suffix for possessed items, *-i/-ü/-u/-ı* (the choice among these determined by vowel harmony), used for 3rd person possessors. Test items were presented in randomized order. The experiment lasted approximately 20 minutes. A native speaker experimenter conducted the test. On each trial, the participant was shown a novel object, created by the experimenters, and was told, for example, *bak, bu bir fet* 'Look, this is a fet'. They were further told that the 'fet' belonged to Kermit (e.g. *bu Kermit'in* 'It's Kermit's). The participant was then asked to complete the sentence *bu Kermit'in*. . . "This is Kermit's . . .", which in Turkish requires the possessed form of a noun. Addition of this vowel-initial suffix provides a "deneutralizing" environment, in which hypothetically the participant could answer either *fet-i* or *fed-i*, i.e. with a *non-alternating* or an *alternating* response.

### *Results*

The children's and adults' responses were analyzed with a logistic regression model using the same variables described in the lexicon analysis above: *size*, *place*, *height*, *backness* and *roundness*, and an additional binary variable for *age-group* (0 for the children and 1 for the adults). Since the experiment only tested two sizes, the *size* variable was binary (CVC and CVCVC) in this analysis. The results in terms of percentages of alternations for each place and size are shown in Table 9.

| | CVC | | CVCVC | |
|---|---|---|---|---|
| | children | adults | children | adults |
| Labial | 17% | 34% | 20% | 53% |
| Coronal | 7% | 6% | 10% | 31% |
| Palatal | 3% | 28% | 40% | 53% |
| Dorsal | 0% | 3% | 60% | 95% |

Table 9: Experiment 1 results

Using the forward stepwise (likelihood ratio) method of SPSS, *size*, *age-group* and *place* were entered with high significance, as shown in Table 10. The factors of *size* and *place* had statistically significant power in predicting the choice of alternation vs. non-alternation in children's and adults' performance in Experiment 1. Vowel features did not have a significant effect on the participants' choices. Adults chose alternating responses significantly more often than children. In the control stimuli, children gave adult-like responses in the vast majority of real Turkish words (96%), with no single child giving more than two non-adult responses.

| | $\beta$ | $\chi^2$ | $R^2$ | $p$ |
|---|---|---|---|---|
| size | 1.915 | 75.8 | .191 | <.001 |
| age-group | 1.070 | 24.0 | .246 | <.001 |
| place | .276 | 8.4 | .264 | <.005 |

Table 10: Logistic regression results for Experiment 1.

The logistic regression analysis facilitated comparison between the role of size, place, height and backness for the dependent variable of alternation rate, in the lexicon and in the experimental results. An ANOVA over the experimental results with *Size* and *place* used as within-subject factors, and *age-group* as a between-subject factor showed that *size*, *place* and their interaction came out significant: size [F(1,20) = 65.7, p<.001], place [F(3,18) = 8.92, p=.001] and size X place [F(3,18) = 13.2, p<.001]. Additionally, *age-group* was weakly significant, with [F(1) = 5.19, p=.034].

Figure 2 shows the percent of alternating responses from children (grey) and adults (black), grouped by size and place, plotted against the percent of alternating words of the relevant sizes and places in the lexicon. While children were somewhat more conservative than adults, erring on the side of less alternations, responses from both groups match the lexicon very well (correlation coefficients r(6)=.82, p<.001 children; r(6)=0.90, p<.001 adults).

## *Discussion*

The results of the production task demonstrate that Turkish speakers are sensitive to lexical effects of place of articulation and size in determining the likelihood that a stop will alternate. These two factors are crosslinguistically plausible as determinants of voicing, and their effects on voicing alternation rates are reproduced by speakers in novel words. Although we found a quantitative difference in rates of alternation between children and adults, the overall qualitative pattern was the same, namely more alternations in disyllabic

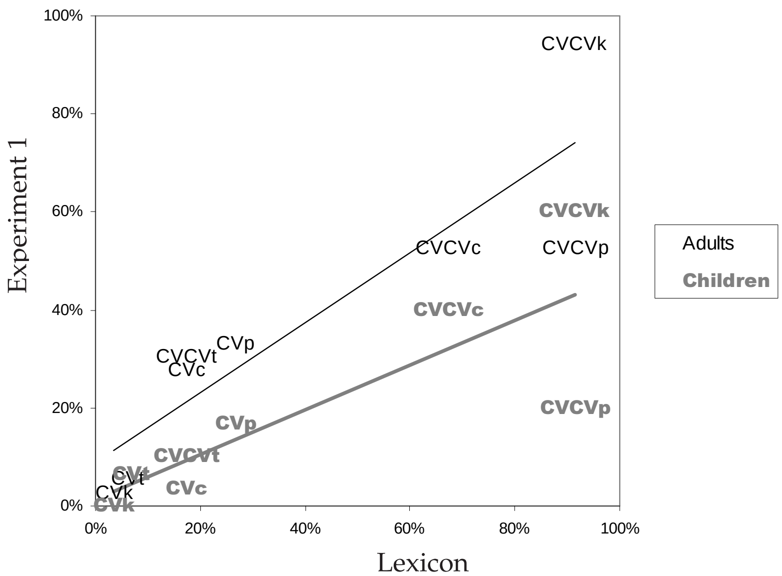


*Figure 2*. Rates of alternation in the lexicon, by place and size, plotted against the percentage of alternating responses in Experiment 1.

words, and a scale of alternations across place that followed the pattern dorsal > labial > palatal > coronal.

Having established that the two factors that we expected, based on lexicon analysis, to play a role in conditioning active generalization to novel words in fact did so, and that the method of nonce word experimentation is sensitive enough to detect such effects for the alternation in question in Turkish, we turned to a direct investigation of whether vowel height would or would not be used in conditioning voicing alternation. If a set of analytic biases about possible and impossible phonological interactions constrains what types of lexical statistics are used and employed, there should be no effect of vowel height on speakers' choices of alternating vs. non-alternating responses.

## Experiment 2

Experiment 2 examined the effects of vowel quality on voicing alternation rates for nonce word stimuli, directly comparing them with size and place effects, and thereby increased the number of experimental items drastically. As the qualitative pattern of both children and adults was essentially the same in Experiment 1, we experimented only with adults in Experiment 2, since adults are more tolerant of longer experiments.

## *Method*

### *Participants*

Participants were adult native speakers of Turkish (n = 24; 13 males, 11 females, age range: 18-45) living in the United States, and participated in the experiment on a volunteer basis. None of the subjects in Experiment 2 were participants in Experiment 1.

### *Materials*

A speaker of Turkish recorded the bare form and two possible possessive forms for each noun. Each stimulus was normalized for peak intensity and pitch and inspected by a native speaker to be natural and acceptable. As the primary purpose of this experiment was to investigate the effect of vowel height, labial and palatal place of articulation were collapsed as a single category, based on the lexical statistics and the results of Experiment 1, in which these show similar rates of alternation. For each of these three place categories (labial/palatal, coronal, dorsal), there were three phonological sizes: CVC, CVCC, and CVCVC. Recall that for dorsal-final words, the relevant alternation is k/∅ in CVC and CVCVC words and k/g in CVCC words; both of these cases instantiate an alternation, while non-alternating k-final words remain k-final in all three cases of CVC,CVCVC, and CVCVC.

Within the CVC and CVCC words, there were 4 items with high vowels and 4 items with non-high vowels. Within CVCVC, the second vowel has to be more constrained given the phonotactics of Turkish, so there were 2 items with non-high final vowels, and 6 items with high final vowels. Recall that this latter unequal distribution is due to the fact that native Turkish disyllabic words cannot have the two round nonhigh vowels [o, ö] in their final syllable. There were thus a total of 72 experimental items, shown in Table 11. The experimental items were mixed with 36 fillers, all of which ended in either fricatives or sonorant consonants.

| | | | CVC | | CVCC | | CVCVC | |
|---|---|---|---|---|---|---|---|---|
| | | | −high | +high | −high | +high | −high | +high |
| p/č | −rd | −back | gep | yič | telp | ginč | heveč | ǰisip |
| | | +back | dap | nɨč | panč | dɨrp | yɨyap | ma.ɨč |
| | +rd | −back | köč | züp | yönč | kürp | | bölüč<br>türüč |
| | | +back | poč | tup | solp | munč | | konup<br>guyup |
| t | −rd | −back | pet | hit | zelt | čint | niket | gevit |
| | | +back | fat | mɨt | hant | šɨrt | ya.at | pɨsɨt |
| | +rd | −back | söt | ǰüt | gönt | nürt | | sölüt<br>bünüt |
| | | +back | yot | nut | ǰolt | bunt | | čorut<br>muyut |
| k | −rd | −back | vek | zik | helk | tink | mesek | perik |
| | | +back | ǰak | pɨk | vank | nɨrk | tatak | banɨk |
| | +rd | −back | hök | sük | sönk | pürk | | nönük<br>düyük |
| | | +back | mok | nuk | bolk | dunk | | zoruk<br>yuluk |

Table 11: Experiment 2 Materials

*Procedure*

We employed a forced-choice decision task, presented as a computerized experiment. The base form, e.g. *fet* was presented in Turkish orthography, which reflects the relevant aspects of the phonology faithfully. On each trial, a participant 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 were auditorally presented with two possible possessed forms, e.g. non-alternating *fe**t**-i* and alternating *fe**d**-i*. For filler items, participants were presented with a forced-choice between vowel-length alternations (e.g. *ruh*; *Ali'nin ruh-u* or *Ali'nin ru:h-u*) or vowel/zero alternations in the final syllable (e.g. *burun*; *Ali'nin burun-u* or *Ali'nin burn-u*). There were four real stop-final words as warm-up items, two of which were alternating, and two of which were non-alternating. The order of nonce items were randomized for each subject, as was the order between the two auditorily presented possessed forms between which the forced-choice was to be made. The experiment was self-paced and lasted approximately 15-20 minutes.

## *Results*

The results, in terms of percentages of alternation for size, place, and each vowel feature, are shown in Table 12.

| | | | CVC | | CVCC | | CVCVC | |
|---|---|---|---|---|---|---|---|---|
| | | | −high | +high | −high | +high | −high | +high |
| p/č | −rd | −back | 63% | 33% | 67% | 75% | 46% | 54% |
| | | +back | 42% | 29% | 54% | 50% | 71% | 63% |
| | +rd | −back | 33% | 33% | 58% | 63% | | 63%<br>75% |
| | | +back | 25% | 50% | 54% | 85% | | 88%<br>75% |
| t | −rd | −back | 46% | 25% | 38% | 29% | 33% | 46% |
| | | +back | 29% | 21% | 54% | 46% | 54% | 54% |
| | +rd | −back | 42% | 46% | 42% | 54% | | 50%<br>66% |
| | | +back | 29% | 38% | 21% | 38% | | 33%<br>59% |
| k | −rd | −back | 33% | 25% | 21% | 42% | 79% | 83% |
| | | +back | 21% | 33% | 38% | 29% | 79% | 92% |
| | +rd | −back | 33% | 29% | 50% | 13% | | 92%<br>79% |
| | | +back | 33% | 50% | 21% | 33% | | 75%<br>79% |

Table 12: Experiment 2 results

Participants' responses were analyzed with a logistic regression model that had the same variables as the ones described in the analysis of the lexicon: *size*, *place*, *height*, *backness* and *roundness*. Using the forward stepwise (likelihood ratio) method of SPSS, *size*, and *place* were entered with high significance, as shown in Table 13.

| | $\beta$ | $\chi^2$ | $R^2$ | $p$ |
|---|---|---|---|---|
| size | .654 | 113 | .085 | <.001 |
| place | .249 | 28 | .105 | <.001 |

Table 13: Logistic regression results for Experiment 2.

Additionally, an ANOVA was performed with *size*, *place*, *high* and *back* as within-subject factors. *Size*, *place* and their interaction were significant: [F(2,22) = 29.54, p<.001], [F(2,22) = 6.55, p=.006] and [F(4,20) = 13.68, p<.001], respectively. *Height* and *backness* did not reach significance: [F(1,23) = 1.43, p = .244] and [F(1,23) = .05, p = .827] respectively.

Figure 3 shows the results from Experiment 2, grouped by size and place, plotted against the percent of alternating words in the lexicon with the matching size and place. We see an excellent correlation between the two, showing that speakers have accurately matched the percentages of alternating words in the lexicon.

In contrast to the tight correlation between the lexicon and Experiment 2 that is shown in Figure 3 for place and size effects (r(10)=.937, p<.001), there is no pattern in the correlation between the lexical statistics and the experimental results for height effects, as shown in Figure 4 (r(10)= .259, p>.1). In other words, even though the lexicon demonstrates a pattern of predictability of alternation based on vowel height, participants did not reproduce this trend in their experimental results.

### *Discussion*

The results of this experiment replicated the findings of Experiment 1 that speakers extend and generalize the statistical trend in the lexicon that place of articulation and word-length are determining factors in whether a root-final stop will be alternating or not. In addition, the method was sensitive enough to find a reliable difference in alternation rates between CVC and CVCC palatal-final nouns, even though there is only a handful of such words available for speakers as evidence (23 and 18 words in TELL, respectively). By contrast, speakers did not make use of the vowel height or vowel backness correlations, even though these are strong predictiors of alternation in the lexicon. The overt

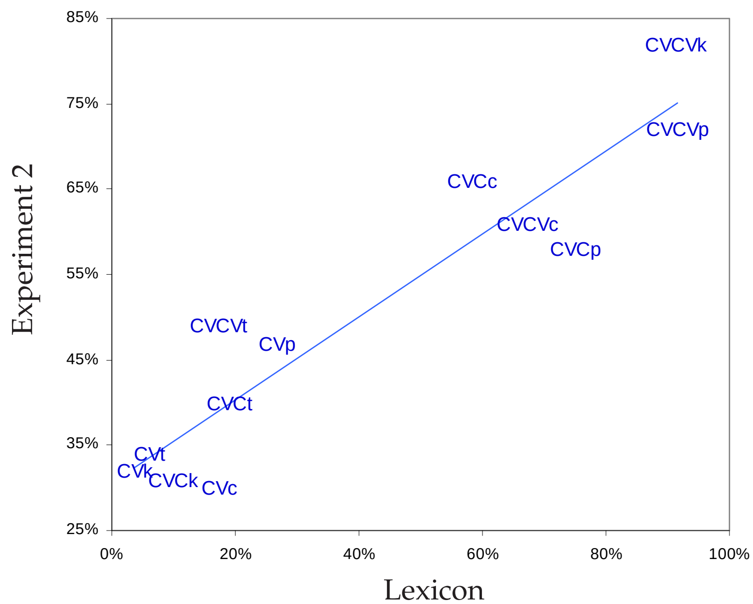


*Figure 3*. Rates of alternation in the lexicon, by place and size, plotted against the percentage of alternating responses in Experiment 2.

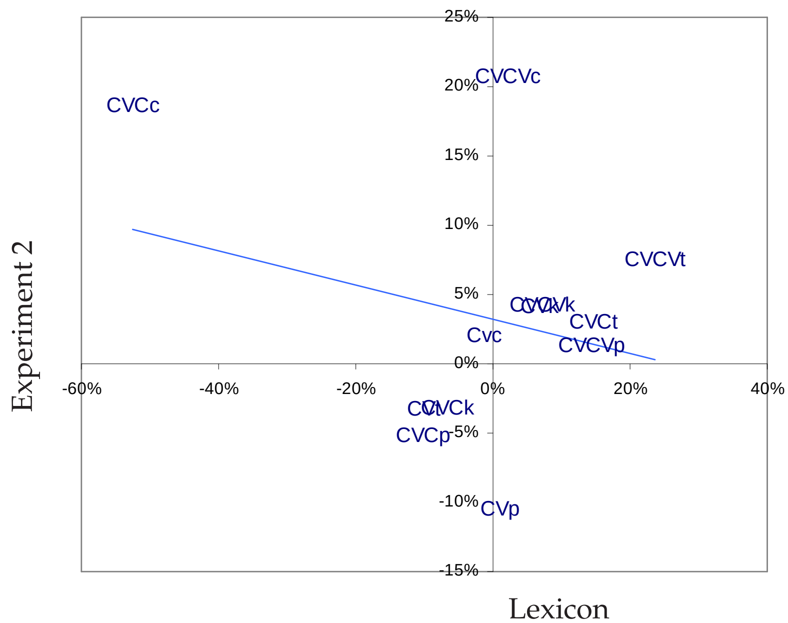


*Figure 4*. The difference in rates of alternation between high and non-high vowels, by size and place, in the lexicon and in Experiment 2. Positive values indicate more alternations with [+high] vowels, negative values indicate more alternations with [−high] vowels.

evidence available for speakers to learn the vowel height correlation includes potentially all the stop-final words in the lexicon, or more conservatively, the 1255 coronal-final words where the height effect is strongest. Nonetheless, the results of Experiment 2 demonstrate that this informative evidence about vowel quality from the lexicon was not extended by speakers in the nonce word task.

These results suggest that language speakers cannot literally learn any correlation, but rather are constrained by a set of biases that favor (and disfavor) attention to certain types of phonological generalizations. While vowel height and vowel backness are statistically reliable predictors of voicing alternation rates that are right there in the lexicon for speakers to pick up on and potentially use in guiding their choices on unpredictable voicing alternations, speakers did not use this information at all, suggesting that they never actively incorporated these facts into their phonotactic knowledge. We view our results as compatible with models in which observed correlations in the lexicon are filtered through a set of principled biases that are available to the learner in advance of collecting statistical information. Before proceeding to a general discussion, we compare the results of Experiment 2 with the predictions made in a computer simulation by an "omnivorous" model of extracting statistical lexical generalizations.

## Simulation with *The Minimal Generalization Learner*

The Minimal Generalization Learner (MGL) of Albright and Hayes (2003) is an information-theoretic algorithm that generalizes patterns over classes of words that undergo similar alternations. MGL provides a reflection of robust trends in the lexicon and has the potential to generalize them to novel outputs. The MGL has been empirically shown to model humans' experimental results in novel word-formation tasks with the past tense in English, 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 Minimal Generalization Learner (MGL) works by processing paradigms of related forms, creating a rule for each paradigm, and then collapsing among those rules to yield a set of more general rules. These more general rules can be applied to novel forms, giving a set of possible outputs with a confidence score assigned to each.

The MGL's operation is exemplified in Table 14 below. Two alternating nouns, *kebap* 'kebab' and *šarap* 'wine' are read, and a rule is projected from each (14a,b). The MGL identifies the structural change in each paradigm ([p] 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 (14c). 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*) but its success rate is lower, since it will mistakenly apply to non-alternating nouns that end in *ap*. This tradeoff between scope and accuracy is balanced by calculating adjusted confidence scores for each postulated rule.

| | paradigm | rule | |
|---|---|---|---|
| a. | šarap$_2$ ∼ šarabɨ$_2$ | p → bɨ / | š a r a __ $_2$ |
| b. | kebap$_2$ ∼ kebabɨ$_2$ | p → bɨ / | k e b a __ $_2$ |
| c. | | p → bɨ / | X a __ $_2$ |

Table 14: A minimal generalization in the MGL. The subscript on the final consonant (e.g. '2') annotates the number of syllables in the word.

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.

### *Materials and method*

To simulate the experimental behavior of the human participants in Experiment 2, the MGL was provided with all the stop-final words in TELL as training data, and with the stimuli of Experiment 2 as test items. In addition, the MGL received a a list of the consonants and vowels of Turkish and the natural classes they form.

The MGL was given the stop-final nouns of TELL in paradigm form, such that each noun was paired with its possessive form. In addition, each bare noun and possessive form were annotated for the mono-/poly-syllabicity of the bare noun. This annotation was necessary to allow the MGL to find the size effect; the MGL is biased to find generalizations locally and would not have found the size effect on its own, since that would require counting over several segments far into the word.

For each test item, the MGL generated alternating and non-alternating responses, each response associated with a confidence score, which estimates the likelihood of obtaining that response from a human participant. To calculate the proportion of alternating responses that the MGL predicts, the confidence score of the alternating response was divided by the sum of the confidence scores of the alternating and non-alternating responses.

The results of the MGL simulation thus yielded a predicted alternation rate for each of the 72 test items of Experiment 2. To make the MGL results comparable with the lexicon and the experimental results using statistical methods, and to make the potential to find effects as balanced as possible, the MGL results were calculated 24 times (to match the number of subjects in Experiment 2), and each data point was multiplied by an independent random variable between 0 and 1 (to match the average variance in the experimental results). Each data point was rounded to zero or one, thus creating a data set of 1728 binary responses, in order to match entirely the number of responses of Experiment 2.

### *Results*

We report here the results obtained by running the MGL at the 75% confidence level, which is the level that generated results that are most similar to the human results. Figure 5 shows MGL's prediction for the nonce words of Experiment 2, grouped by size vs. place, plotted against the proportion of alternating words in TELL in the corresponding size and place. The MGL predictions match the lexicon well ($r(10)=.92$, $p<.001$).

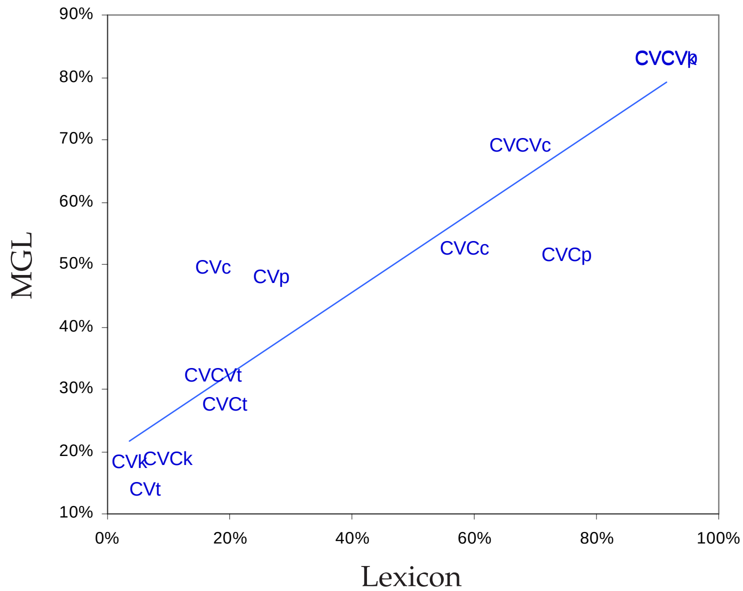


*Figure 5*. Rates of alternation in the lexicon, by place and size, plotted against the percentage of alternating responses predicted by the Minimal Generalization Learner.

The MGL results were analyzed with a logistic regression model that had the same variables as the ones described in the analysis of the lexicon and of Experiment 2: *size*, *place*, *height*, *backness* and *roundness*. Using the forward stepwise (likelihood ratio) method of SPSS, *size*, *place*, *backness* and *height* were entered, as shown in Table 15.

| | $\beta$ | $\chi^2$ | $R^2$ | $p$ |
|---|---|---|---|---|
| size | 1.057 | 241.1 | .185 | $<.001$ |
| place | .933 | 212.9 | .327 | $<.001$ |
| backness | $-.419$ | 11.5 | .355 | $<.005$ |
| height | .409 | 10.1 | .341 | $<.005$ |

Table 15: Logistic regression results for TELL.

The MGL results mirror the results of the lexical analysis, with *backness* and *height* reaching significance. These results contrast with the results of Experiment 2 that show no effect of *backness* and *height*. In other words, the MGL overfit the data by projecting generalizations from the lexicon that were not projected by humans.

The correlation between the lexical statistics and the MGL results for the height effect, as shown in Figure 6, is quite good ($r(10)=.52$, $p<.05$). This contrasts sharply with the lack of a reliable correlation between the lexical statistics for vowel quality and the results from Experiment 2 (see Figure 4 above).

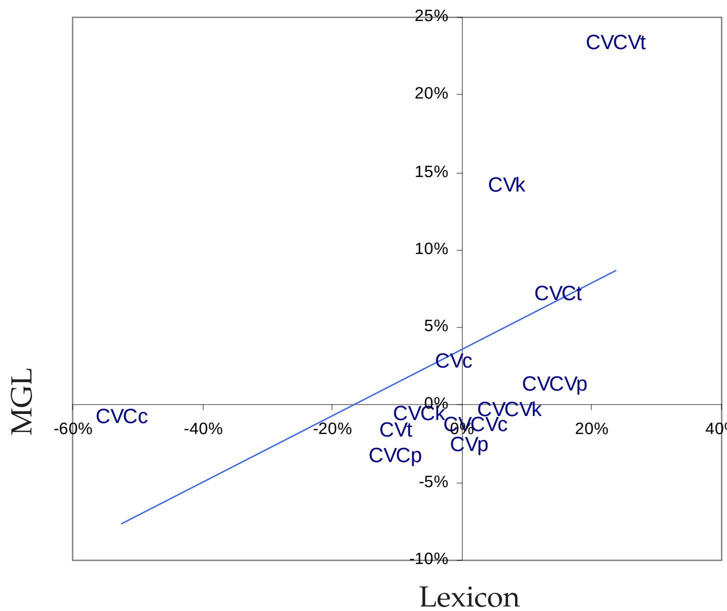


*Figure 6*. The difference in rates of alternation between high and non-high vowels, by size and place, in the lexicon and in the MGL results. Positive values indicate more alternations with [+high] vowels, negative values indicate more alternations with [−high] vowels.

## *Discussion*

The MGL's impressive performance in matching the lexical trends of Turkish voicing alternations is not an accurate model of the participants' behavior. The MGL outperformed the participants of Experiment 2, tracking and extending statistical generalizations that humans did not.

The MGL is a powerful learner of phonotactic regularities: given nothing but a list of paradigms and the natural classes that the segments therein belong to, the MGL learned that Turkish has voicing alternations, and that there are factors that are correlated with its 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.

The MGL results are representative of a wider range of learning algorithms (e.g. Quinlan (1993)) that can use purely distributional properties of a lexicon to predict human behavior. The speakers of Turkish in Experiment 2 showed a bias to ignore any effect that vowel quality might have on the voicing of a neighboring consonant. The current simulation demonstrated that a model that is not equipped with a set of biases of what to pay attention to and what to ignore is not able to successfully reproduce humans' experimental responses.

## General Discussion

The findings that Turkish speakers extend some but not all lexical statistics in nonce word tasks raise a number of points of discussion. The positive findings of Experiments 1 and 2, that speakers do employ some statistical information in generalizing phonotactic patterns, may be considered as an instance of the more general cognitive strategy of use of base rate information in the context of deneutralization tasks. The findings of Experiment 2, 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.

### *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* does not alternate when a vowel-initial suffix is added, as in the possessed form *sop-u*, but the noun stem *jop* does: its possessed form is *jo**b**-u*. Given a nonce word like *zop*, 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, Tenenbaum, and Gopnik (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] 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]-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* and asked whether to judge whether the deneutralized form should be *zo**p**-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 Experiments 1 and 2 reported here show that Turkish speakers can and do use this information in reasoning whether a word like *zop* 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 and 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.

### *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. Nakipoğlu and 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. *kitap* → *kipat* 'book') do not occur with monosyllables (e.g. *yap*) 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 arguably 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 and Abramson (1964); Ohala (1983); Volatis and 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 and 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 (e.g. Canadian Raising (Chambers (1973), Moreton and Thomas (2005)), 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 (2006), 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 unable to generalize 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 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 *či* or *ši*; 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, Perfors, and Tenenbaum (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.

### *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 (1999); 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 and Kaye (1990); Archangeli and 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 never used and in fact excluded or disfavored by learning biases, why do these phonetically-unmotivated patterns 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 [ǰ] than [ʒ], are ultimately traceable to lexical borrowing from French centuries ago, when Old French had [ǰ] 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 noun to become alternating (e.g. Arabic *burǰ* 'sign' > Turkish *burč* ∼ *burǰ-u*), whereas source words that end in a voiceless stop are non-alternating across the paradigm. Arabic lacks the consonants [p] and [č] and has many nouns that end in [b] and [ǰ], 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 (such as, in an extreme case, the children tested in Experiment 1), 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. 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 (2006) that English speakers could not learn an artificial language pattern with height-voicing interactions, while they were able to learn non-adjacent V-V interactions, in which high vowels were followed by high vowels in the adjacent syllable. In Turkish, the case is arguably 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 the results of our Experiment 1 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.

## References


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

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

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

Beckman, J. (1998). *Positional Faithfulness*. Unpublished doctoral dissertation, UMass, Amherst.

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

Blust, R. (2000). Low-vowel fronting in northern sarawak. *Oceanic Linguistics*, *39*(2), 285–319.

Bonatti, L., Peña, M., Nespor, M., & Mehler, J. (2005). Linguistic constraints on Statistical Computations. *Psychological Science*, *16.6*, 451–459.

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

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

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

Cho, Y.-m. Y. (1999). *Parameters of Consonantal Assimilation*. Munich: Lincom Europa.

Cohn, A. (1993). *Phonetic and Phonological Rules of Nasalization*. Unpublished doctoral dissertation, UCLA.

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

Ernestus, M., & Baayen, H. (2003). Predicting the Unpredictable: Interpreting Neutralized Segments in Dutch. *Language*, *79.1*, 5–38.

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

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

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

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

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

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

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

Hayes, B. (1989). Compensatory Lengthening in Moraic Phonology. *Linguistic Inquiry*, *20.2*, 253–306.

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

Inkelas, S. (1999). Exceptional stress-attracting suffixes in Turkish: representations versus the grammar. In *The Prosody-Morphology Interface* (pp. 134–187). Cambridge University Press.

Inkelas, S., Küntay, A., Lowe, J., Orgun, O., & Sprouse, R. (2000). *Turkish Electronic Living Lexicon (TELL).* Website, http://socrates.berkeley.edu:7037/.

Inkelas, S., & Orgun, C. O. (1995). Level Ordering and Economy in the Lexical Phonology of Turkish. *Language*, *71.4*, 763–793.

Inkelas, S., Orgun, C. O., & Zoll, C. (1997). The implications of lexical exceptions for the nature of grammar. In *Derivations and Constraints in Phonology* (pp. 393–418). Oxford: Clarendon.

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

Kemp, C., Perfors, A., & Tenenbaum, J. (2007). Learning overhypotheses with hierarchical Bayesian models. *Developmental Science*, *10.3*, 307–321.

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

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

Kopkallı, H. (1993). *A phonetic and phonological analysis of final devoicing in Turkish*. Unpublished doctoral dissertation, University of Michigan.

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

Lewis, G. L. (1967). *Turkish Grammar*. Oxford: Clarendon.

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

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

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

Moreton, E. (2006). Phonotactic learning and phonological typology. *submitted*.

Moreton, E., & Thomas, E. (2005). Origins of Canadian Raising in voiceless-coda effects: a case study in phonologization. In *Papers in Laboratory Phonology 9*. .

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

Nakipoğlu, M., & Üntak, A. (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.

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

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

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

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

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

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

Sezer, E. (1981b). On non-final stress in Turkish. *Journal of Turkish Studies*, *5*, 61–69.

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

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

Vaux, B. (2005). *Formal and empirical arguments for morpheme structure constraints.* Paper presented at the Linguistic Society of America meeting, San Francisco, CA.

Volatis, L., & Miller, J. (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, J., & Dell, G. (2006). Speech errors reflect newly learned phonotactic constraints. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *32.2*, 387Ð398.

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

Wilson, S. M. (2003). A phonetic study of voiced, voiceless, and alternating stops in Turkish. *Newsletter of the Center for Research in Language*, *15.1*, 3–13.

Zimmer, K., & Abbott, B. (1978). The k/$\emptyset$ alternation in Turkish: Some experimental evidence for its productivity. *Journal of Psycholinguistic Research*, *7*, 35–46.