When and Why to Ignore Lexical Patterns in Turkish Obstruent Alternations

Michael Becker, UMass Amherst (michael@linguist.umass.edu)
Nihan Ketrez, Yale (nihan.ketrez@yale.edu)
Andrew Nevins, Harvard (nevins@fas.harvard.edu)

Highlights:

- In the lexicon, alternation rates can be predicted by the place of articulation of the stem-final stop, the size of the word, and the height and backness of the stem-final vowel
- Speakers are attuned to the place of articulation and size effects; they ignore vowels – they distinguish between phonologically-motivated generalizations and accidental generalizations
- We propose that UG acts as a filter on the generalizations the humans can make; UG contains information about possible and impossible interactions between phonological elements
- We show that computational models that do not have information about possible interactions reproduce accidental generalizations
- We propose an OT model that can capture the categorical behavior of known nouns and the stochastic behavior of novel nouns

1 The Turkish lexicon

Turkish regulates the voicing of stem-final stops, productively enforcing final devoicing and intervocalic voicing:

(1)	rop ~ rob-u	< French [rɔb]	‘dress’
	tüp ~ tüb-ü	< French [tüb]	‘tube’
	gurup ~ gurub-u	< French [grup]	‘group’
	ešarp ~ ešarb-ɨ	< French [ešarp]	‘scarf’

But lexical exceptions abound:

(2) Failure of final devoicing: ad ~ ad-ɨ 'name'
Failure of intervocalic voicing: top ~ top-u 'ball'

Exceptions to final devoicing are fairly rare (~2% of the lexicon at best)
We focused on exceptions to intervocalic voicing (~46% of the lexicon)

Application of intervocalic voicing is unpredictable for any given existing lexical item, but certain factors are known to affect it. We mined the data in TELL (Inkelas et al. 2000) for such factors:

(3) **Size**

	n	% alternating
Monosyllabic, simplex coda (CVC)	137	12%
Monosyllabic, complex coda (CVCC)	164	26%
Poly-syllabic (CV.CVC and longer)	2701	59%

(4) **Place of articulation of the stem-final stop**

	n	% alternating
p	294	84%
t	1255	17%
č	191	61%
k	1262	85%

(5) **Height of the stem's final vowel:**

	n	% alternating
–high (a, e, o ,ö)	1690	42%
+high (ɨ, i, u, ü)	1312	72%

(6) **Backness of the stem's final vowel:**

	n	–back (e, i, ö, ü)	+back (a, ɨ, o, u)	difference
p	294	90%	79%	–11%
t	1255	16%	18%	+1%
č	191	44%	74%	+30%
k	1262	84%	86%	+1%

2 Speakers’ knowledge about voicing alternations

A hint comes from Inkels & Orgun (1995) and Inkelas et al. (1997), who mention **size** and **place**, but not the vowel effects.

Our bottom line: speakers not only replicate the lexical **size** effect and **place** effect, but also their **intersection**. Speakers do not replicate any vowel effects – neither **height** not **backness**.

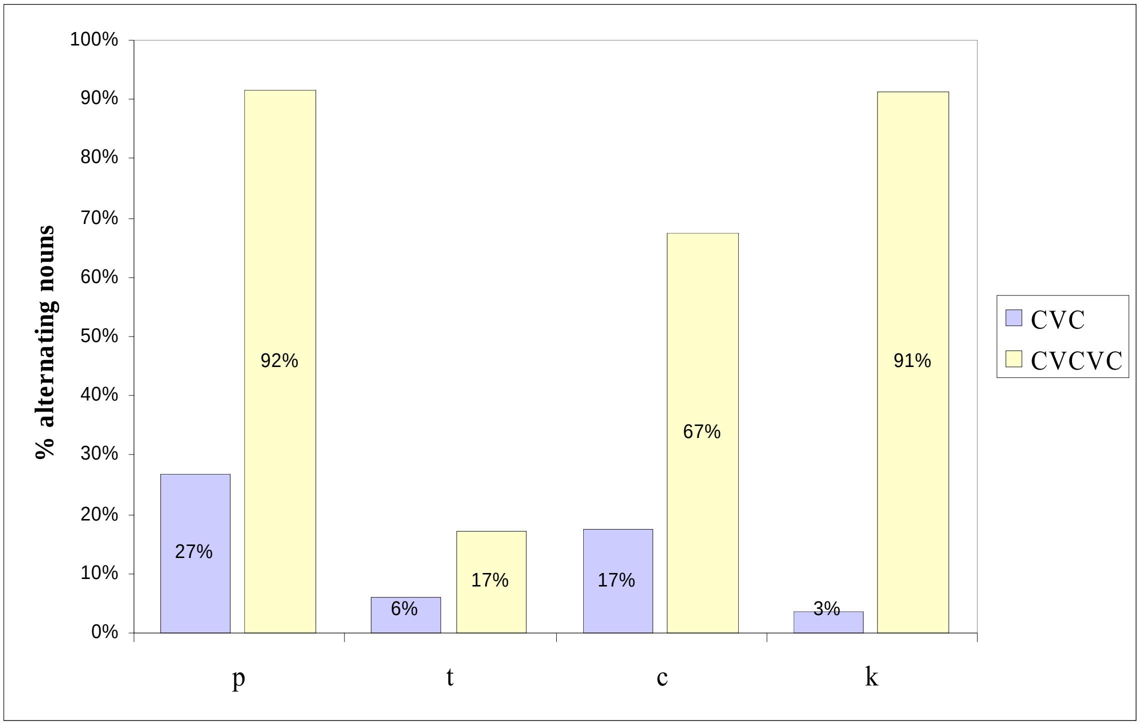
2.1 Experiment 1: Size × place

We showed 10 kids (average age 4;0) various toys, presenting each with a name, e.g. yüč. We asked them to name Kermit’s yüč, hoping to get either “Kermit’in yüč-ü” or “Kermit’in yüǰ-ü”. We also ran the same experiment with 11 adults.

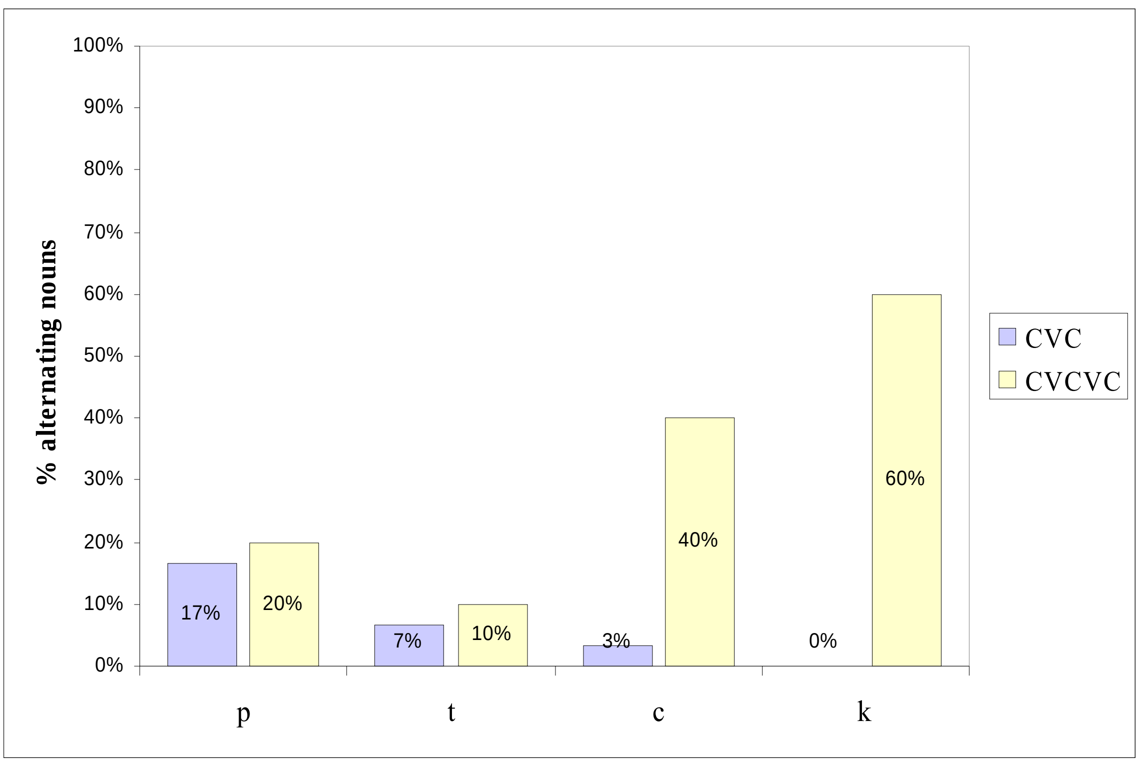
Results: Both kids and adults not only replicated the **size** and **place** effects from the lexicon, but also their **intersection** (see also Nevins & Kamali 2005, Nakipoğlu & Üntak 2006)

(7)	CVC	kids	adults	CVCVC	kids	adults
p	tup	0%	17%	yɨyap	20%	42%
	dap	20%	42%	ǰisip	20%	50%
	zip	30%	42%	kunup	20%	67%
t	nut	10%	17%	gevit	20%	25%
	hit	0%	0%	muyut	10%	42%
	fet	10%	0%	niket	0%	25%
č	nɨč	0%	25%	bölüč	50%	50%
	yüč	10%	42%	heveč	40%	58%
	peč	0%	17%	maɨč	30%	50%
k	sük	0%	0%	banɨk	40%	100%
	pɨk	0%	0%	nönük	80%	92%
	vek	0%	8%	mesek	60%	92%

(8) The lexicon (according to TELL)



(9) Kids' choices (Experiment 1)



(10)	Kids vs. lexicon:	$r(8) = .82$	$p < .01$
	Adults vs. lexicon:	$r(8) = .90$	$p < .01$
	Adults vs. kids	$r(24) = .84$	$p < .001$

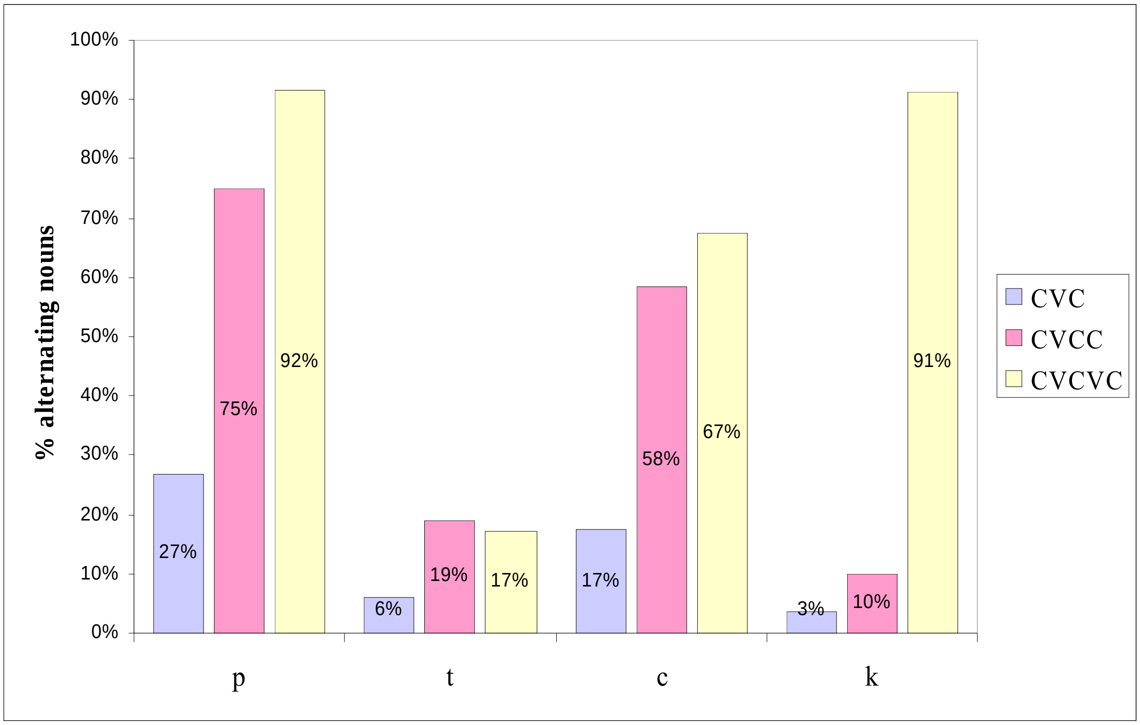
2.2 Experiment 2: Size × place × vowel quality

We showed 24 adults nouns, e.g. köč. They were shown a possessor, and asked to choose between two vocal renditions of the possessed noun, i.e. "köč-ü" or "köǰ-ü".

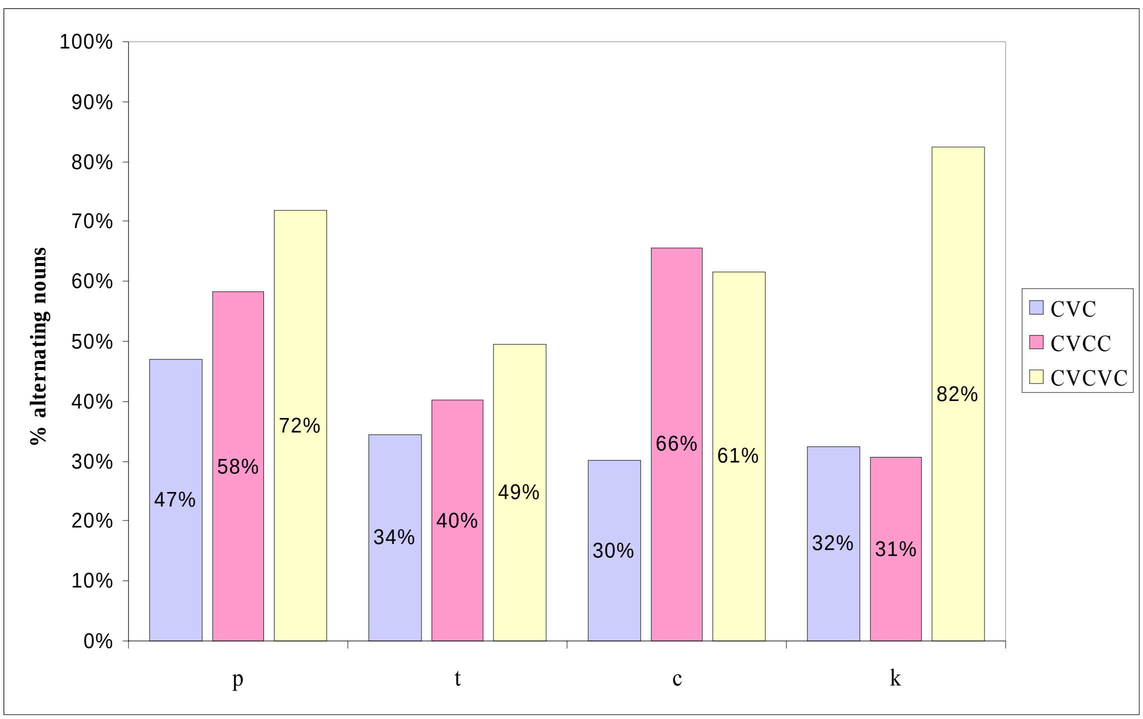
(11) Sizes: CVC, CVCC and CVCVC
Places: p, t, č, k
Vowels: a, ɨ, e, i, o, u, ö, ü

			CVC		CVCC		CVCVC	
			–high	+high	–high	+high	–high	+high
p/č	–r	–b	gep	yič	telp	ginč	heveč	ǰisip
		+b	dap	nɨč	panč	dɨrp	yɨyap	maɨč
	+r	–b	köč	züp	yönč	kürp		bölüč türüč
		+b	poč	tup	solp	munč		kunup goyup
t	–r	–b	pet	hit	zelt	čint	niket	gevit
		+b	fat	mɨt	hant	šɨrt	yaat	pɨsɨt
	+r	–b	söt	ǰüt	gönt	nürt		sölüt bünüt
		+b	yot	nut	ǰolt	bunt		čorut muyut
k	–r	–b	vek	zik	helk	tink	mesek	perik
		+b	ǰak	pɨk	vank	nɨrk	tatak	banɨk
	+r	–b	hök	šük	sönk	pürk		nönük düyük
		+b	mok	nuk	bolk	dunk		zoruk yuluk

(12) The lexicon (according to TELL)



(13) Adults' choices



(14) Adults vs. lexicon, size × place: $r(12) = .94$ $p < .001$
Adults vs. lexicon, vowel height: $r(12) = –.26$ $p > .1$
Adults vs. lexicon, vowel backness: $r(12) = –.13$ $p > .1$

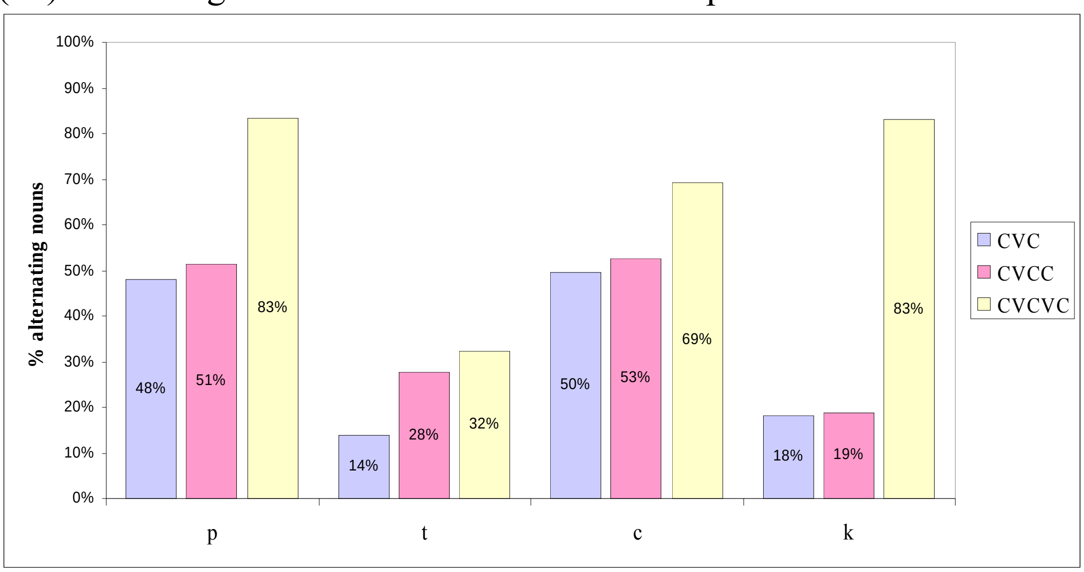
3 UG-less modeling of speakers' knowledge

Speakers' knowledge is tightly correlated with the lexicon; we assume that they project their knowledge from the lexicon. We used the Minimal Generalization Learner (henceforth MGL, Albright & Hayes 2002), which can learn generalizations from a corpus and project them onto novel words.

We gave the MGL the stop-final words of TELL and their possessive forms, annotated for mono-/poly-syllabicity, and a table of segments and their features. The MGL starts with creating a rule for each word of the lexicon, and collapses phonologically similar rules to form generalizations. Ex.:

(15)	šarap$_2$ ~ šarabɨ$_2$	rule:	p → bɨ /	š a r a __ $_2$
	kebap$_2$ ~ kebabɨ$_2$	rule:	p → bɨ /	k e b a __ $_2$
	Generalization:		p → bɨ /	X a __ $_2$

(16) Then we gave the MGL the words from experiment 2:



The MGL modeled the lexicon *too well*, generating data that fit the lexicon more closely than humans. Most importantly, it failed to ignore the vowel effects that are in the lexicon.

(17)	MGL vs. lexicon, size × place:	$r(12) = .92$	$p < .001$
	MGL vs. humans, size × place:	$r(12) = .87$	$p < .001$

The MGL matched the size and place effects of the lexicon and of human behavior very well

(18)	MGL vs. lexicon, vowel height:	$r(12) = .52$	$p = .079$
	MGL vs. humans, vowel height:	$r(12) = .17$	$p > .1$

The MGL matched the lexicon's height effect fairly well (at 7.9% confidence level), but did not match human behavior.

(19)	MGL vs. lexicon, vowel backness:	$r(12) = .63$	$p = .027$
	MGL vs. humans, vowel backness:	$r(12) = -.9$	$p > .1$

The MGL matched the lexicon's backness effect well, but did not match human behavior.

We conclude that MGL reproduced both the phonologically-motivated generalizations (size and place effects) and the accidental generalizations (vowel height and backness) that were found in the lexicon.

What's missing from the MGL is a theory of possible and impossible interactions between phonological elements.

4 The ingredients of a UG-based analysis

We assume that UG acts as a filter on learning the lexicon. UG constrains the learning process, letting speakers find phonologically-motivated generalizations and ignoring others.

When speakers derive novel forms, they **do not access their lexicon**. They only use their grammar, which has the phonologically-motivated aspects of the lexicon built into it.

4.1 Principled vs. accidental generalizations

Phonetically-grounded generalizations:

(20) Voiced stops require maintaining higher sub-glottal pressure the closer the obstruction is to the glottis (g > ǰ > d > b)
(21) The release of voiceless stops is more audible the farther they are from the lips (p > t > č > k)
(22) Voiced stops and high vowels share a lowered F_1
(23) Vowel backness and rounding (F_2) share no features with voicing

Formally-/phonologically-based generalizations:

(24) Initial syllables are protected by positional faithfulness

Information about whether or how to repair marked structures is phonological, not phonetic

4.2 Lack of effect from lack of constraints

We use Optimality Theory (OT, Prince & Smolensky, 2003), which encodes universally possible and impossible interactions in the structure of the constraint inventory (CON).

(25) No language is known to change obstruent voicing based on the quality of a neighboring vowel

Therefore, OT has no constraints like

(26) *[+back][–voice]

Any relationship between vowel quality and obstruent voicing is necessarily accidental; speakers cannot be attuned to it.

Speakers do have constraints that relate obstruent voicing to size and place. Ex.:

(27) *VT]V – No voiceless stops between vowels in a derived environment (enforces intervocalic voicing)

(28) IDENT(voice)$_{\sigma 1}$&*LAB – Labial stops in the initial syllable must surface with their underlying voicing specification

The constraint in (28) is one member in a family of 12 IDENT(voice) constraints relativized to three sizes × four places.

4.3 A stochastic grammar without stochastic ranking

Known words, like top and ǰep, always show up the same way, i.e. top-u and ǰeb-i. A novel word like züp will likely show up as züp-ü, but züb-ü is also possible.

We propose a learning model in which speakers detect inconsistency in the grammar (Pater 2004) and then start keeping track of the behavior of individual items:

(29)

/top+V/	IDENT(voice)	*VT]V
→ topu		1
tubo	W_1	L_0

(30)

/ǰep+V/	IDENT(voice)	*VT]V
→ ǰebi	1	
ǰepi	L_0	W_1

(31) IDENT(voice)$_{\text{top}}$ » *VT]V » IDENT(voice)$_{\text{ǰep}}$

As more words are learned, each instance of Ident will accumulate "weight", and this "weight" is projected onto novel words:

(32) IDENT(voice)$_{\text{top,ip,sop,bap,čap,hep,kip…}}$ » *VT]V » IDENT(voice)$_{\text{ǰep,tüp,kap…}}$

Thus, the ratio of alternating and non-alternating nouns is built into the constraint ranking. A novel word like züp will be attracted by the heavier top-ranking IDENT, so züp-ü is more likely than züb-ü.

5 Conclusion

- Turkish speakers extract information about voicing alternations from their lexicon, relativized to size and place; they ignore the effect of the stem-final vowel
- When speakers derive novel words, they do not access the lexicon. They access their grammar, which is constrained by UG.
- UG contains information about possible and impossible interactions between phonological elements
- We propose an OT model that can learn the phonologically-motivated aspects of the lexicon. The resulting grammar has a categorical effect on known words and a stochastic effect on novel words.

Acknowledgments

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, Michael Key, Joe Pater, John McCarthy, Anne-Michelle Tessier, and Matt Wolf. We would also like to thank the audience at the 2004 HUMDRUM, especially Luigi Burzio, Jane Grimshaw, Alan Prince and Ellen Woolford. Thanks to the many speakers of Turkish who helped with various aspects of this project, including Mine Doğucu, Anıl and Alper Duman, Beste Kamali, Leyla Zidani-Eroğlu, and all the kind people who filled our questionnaire. Any remaining errors, if any, would be ours.

Selected References

Albright & Hayes (2002). Modeling English Past Tense Intuitions with Minimal Generalization. In Maxwell, Michael (ed) *Proceedings of the Sixth Meeting of the ACL Special Interest Group in Computational Phonology*. Philadelphia, July 2002. ACL.

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.

Nakipoğlu, Mine and Aslı Üntak (2006) What does the acquisition of stems that undergo phonological alternation reveal about rule application? Talk given at ICTL 2006, Uppsala

Nevins, Andrew & Beste Yolcu-Kamali (2005). Constructing underlying representations based on informed guesses: Turkish evidence. Talk given at the conference of the Central Eurasian Studies Society, Boston University.

Pater, Joe (2004). Exceptions and optimality theory: Typology and learnability. Talk given in the Conference on Redefining Elicitation: Novel Data in Phonological Theory. New York University.