Exceptions

Class 1: Why exceptions?

Adam Albright

CreteLing 2024

creteling2024.phonology.party

Setting the stage

Why exceptions?

- Exceptions are noise?
- Exceptions have their own grammar?
- Exceptions as a window into “core” grammar?

The goal of grammatical analysis

A fitting problem

- Receive some data about distributions
- Find an analysis (lexicon, grammar) that best fits the observed distributions

Data about distributions

Contexts for sounds

- Phonological
  - Cantonese: no voiced obstruents
  - Dutch, Thai: no final voiced obstruents
- Lexical
  - Thai: voiced [b] *báàn* ‘house’, voiceless [p] *páà* ‘aunt’
  - Dutch: voiced [b] *bar* ‘bar’, voiceless [p] *paːr* ‘pair’
- Both ⇒ alternations
  - Dutch: voiceless [p] *rɪp* ‘rib’, voiced [b] *rɪbən* ‘rib-PL’
  - Dutch: voiceless [p] *lɪp*, *lɪpən* ‘lip/lip-PL’
- Dutch: no morphemes with [b] in some contexts, [m] elsewhere

Analyses that fit the observed distributions

- Analyses: lexicon, grammar
  - I use ‘grammar’ to mean system of rules/constraints
- Fit: “best” analysis (objective function)

Measuring goodness of fit

Phonology problem sets—e.g., Turkish possessives Halle & Clements (1983, p. 96)

‘rope’	ip	ipi	‘Ahmed’	ahmet	ahmedi
‘louse’	bit	biti	‘slipper’	pabuč	pabuǰu
‘reason’	sebep	sebebi	‘power’	güč	güǰü
‘wing’	kanat	kanadɨ	‘basket’	sepet	sepeti
‘honor’	šeref	šerefi	‘art’	sanat	sanatɨ
‘rump’	kɨč	kɨčɨ	‘cap’	kep	kepi
‘pilot’	pilot	pilotu	‘worm’	kurt	kurdu
‘bunch’	demet	demeti	‘hair’	sač	sačɨ
‘wine’	šarap	šarabɨ	‘color’	renk	rengi

Fitting phonology problem sets

- Small amount of data, chosen to be consistent
- Every single form can be accounted for with distinct UR for each morpheme, set of rules/constraint ranking
- No exceptions (enforced strictly)
- Simple rules (enforced fairly strictly)
- Don't predict unseen word types (little attention, esp. pre-OT)

Unseen word types

- Turkish possessives problem
  - Harmless pedagogically motivated simplification, avoiding complications: very little data about final /g/ (one example, in a cluster)
  - Pernicious simplification, suppressing facts that we have no way of explaining: no data on final k/g
- Target skill: crafting general statements
  - Values analyses that predict unseen forms
- Maybe human learners do, too?
- E.g., (Halle, 1978) on English plurals
- Further testimony: proper names, loanwords that frequently fill out problem set data

Actual Turkish

Kaisse (1986), Inkelas & Orgun (1995), and others

- Non-alternating voiceless

at	atɨ	‘horse’
devlet	devleti	‘state’
sap	sapɨ	‘stem’

- Alternating

kanat	kanadɨ	‘wing’
kalɨp	kalɨbɨ	‘mold’

- Non-alternating voiced

ad	adɨ	‘name’
öǰ	öǰü	‘revenge’
etüd	etüdü	‘etude’

Harmless or pernicious simplification?

- ‘A+’ solution to Halle and Clements problem doesn’t fit the Inkelas and Orgun data so well
  - Unexpected but attested: voiced *ad* ‘name’
  - Expected but un(der)attested: final /k/
- Analytical choice
  - Complicate analysis to fit the data better
    - Fancier lexical representations, fancier rules/constraints
  - Simpler analysis, plus exceptions

Fancier lexical representations

Inkelas & Orgun (1995); Inkelas et al. (1997)

- Non-alternating stems: specified [$\pm$voice]
- Alternating stems: underspecified, receive [voice] value from context
- Also some limited restrictions related to morpheme length
- Analysis provides an account for every form
  - Relatively simple grammar, but leaves lexical trends mostly unaccounted for

Fancier grammar

Becker et al. (2011)

- Lexicon: voicing correlated (imperfectly) with root length, place of articulation, C vs. CC, vowel quality
    - NB: Becker et al. focus on voiceless vs. alternating, but it is also well-known that non-alternating voiced are concentrated among monosyllabic roots
- Wug test: speakers productively generalize trends to project voicing for novel roots
    - Root length, place of articulation, *not* vowel quality(?)
- Suggestion: rather than stipulating voicing in lexicon, learn a more complex grammar that determines voicing values
- Analysis doesn't account for every form, but captures lexical trends

Some open questions

- What kinds of complications do human learners have in their phonological repertoire?
  - Contextual refinements to rules/processes
  - Representational tricks to distinguish morphemes
- When do learners employ these complications?
  - “Best” fit: simplicity/accuracy trade-off

Goals of this class

Explore trade-off between analytical complexity and simplicity, and consequences for exceptions

- How poorly do our analyses actually do, if we allow some underfit for the sake of simplicity?
- How much better could our analyses do, if we allow greater grammatical complexity?
- Theory of URs modulates fit
    - Abstraction: "cobbling", underspecification, diacritics, gradient representations, etc.
    - More abstraction, fewer exceptions
- Theory of grammar modulates fit
    - Range of processes that can be expressed, contextual restrictions, process interactions
    - More parameters, fewer exceptions

Goals of this class

Consider sources of evidence for which forms as exceptions

- Psycholinguistic evidence
- Acquisition evidence
- Historical change
- Distributional consequences

What do these tell us about the human objective function (preferred fit)?

Logistical notes

- Assumed background
- Class discussion
- Website

Evaluating analyses

Fit to data

- What makes one analysis “better” than another?
- Starting small: what does it mean for one rule/constraint to be better than another?

The *Sound Pattern of English*

Chomsky & Halle (1968), chap. 8: *On the evaluation procedure and the form of phonological rules*

- Assumption: our job as linguists is to build a theoretical framework that settles on the same grammar that humans do, given the same data
    - They call the theoretical framework a “system of formal devices”
- Theoretical frameworks define hypothesis spaces: sets of analyses that can be expressed
- Goal: find theoretical framework + objective function, such that for any set of data, the highest scored hypothesis expressible in framework is also the ones that humans prefer

SPE evaluation

- Theoretical framework
  - Chomsky and Halle assume basic principles of Generative Phonology
  - Compare frameworks that differ in key formal devices
    - Phonological features, variables, rule ordering, etc.
- Evaluation
  - Goal: find the simplest hypothesis compatible with the data

The simplest hypothesis compatible with the data

- Compatible with the data, in SPE = UR+rules generate every form
  - p. 331: "We will not concern ourselves here with the nontrivial problem of what it means to say that a hypothesis—a proposed grammar—is compatible with the data"
  - p. 331: "In other words, we make the simplifying and counter-to-fact assumption that all of the primary linguistic data must be accounted for by the grammar and that all must be accepted as "correct"; we do not here consider the question of deviation from grammaticalness, in its many diverse aspects."
  - I.e., no errors, no exceptions
- Simplest = shortest (fewest symbols)

The simplest hypothesis, compatible with the data

One way of stating the SPE objective function

$$\text{argmax}\left(\frac{1}{\text{length(grammar)}} - \infty\big(\text{num of exceptions}\big) \right)$$

- First term: score is inverse to length (i.e., shorter=better)
- Second term: any exceptions immediately disqualify the hypothesis
- SPE states as two separate conditions, but combining them allows us to express something that can be evaluated in learning, and which explicitly states the balance

SPE evaluation

- Features: allow more compact expression of rules
  - [$-$sonorant] $\rightarrow$ [$-$voice] / [$-$voice]____ #
  - [$-$sonorant] $\rightarrow$ [$-$voice] / {p,t,k,f,θ,s,ʃ}____ #
  - {d,z} $\rightarrow$ [$-$voice] / {p,t,k,f,θ,s,ʃ}____ #
  - d $\rightarrow$ t / {p,k,f,θ,s,ʃ}____ #, z $\rightarrow$ s / {p,t,k,f,θ,}____ #,
  - Halle (1978): first rule correctly predicts generalization to [xd], [xz]
- Variables: can unify rules
  - [$-$sonorant] $\rightarrow$ [αvoice] / ____ [αvoice]
- Etc.

Applied to the Turkish possessives problem set

‘rope’	ip	ipi	‘Ahmed’	ahmet	ahmedi
‘louse’	bit	biti	‘slipper’	pabuč	pabuǰu
‘reason’	sebep	sebebi	‘power’	güč	güǰü
‘wing’	kanat	kanadɨ	‘basket’	sepet	sepeti
‘honor’	šeref	šerefi	‘art’	sanat	sanatɨ
‘rump’	kɨč	kɨčɨ	‘cap’	kep	kepi
‘pilot’	pilot	pilotu	‘worm’	kurt	kurdu
‘bunch’	demet	demeti	‘hair’	sač	sačɨ
‘wine’	šarap	šarabɨ	‘color’	renk	rengi

Hypotheses

- Medial voicing: $[-\text{son}] \rightarrow [+\text{voi}]$ / $[+\text{son}]$ ___ $[+\text{son}]$
- Final obstruent devoicing: $[-\text{son}] \rightarrow [-\text{voi}]$ / ___ #
- Final stop devoicing: $[-\text{son}, -\text{contin}] \rightarrow [-\text{voi}]$ / ___ #
- Final b,d,ǰ,g devoicing: {b,d,ǰ,g} $\rightarrow [-\text{voi}]$ / ___ #

Evaluating Turkish hypotheses

Schematically:

Medial voicing	$[-\text{son}] \rightarrow [+\text{voi}] / [+\text{son}]\_\_\_\ [+\text{son}]$	$-\infty$(exceptions)
Final obstruent devoi	$[-\text{son}] \rightarrow [-\text{voi}] / \_\_\_\ \#$	1/5
Final stop devoi	$[-\text{son}, -\text{voi}] \rightarrow [-\text{voi}] / \_\_\_\ \#$	1/7
Final b,d,ǰ,g devoi	$\{\text{b,d,ǰ,g}\} \rightarrow [-\text{voi}] / \_\_\_\ \#$	1/7

- Also: no incentive to include separate rules that incorporate additional factors, such as root length, place of articulation, etc.

An exacting objective

- Recall that actual Turkish has words like *ad* $\sim$ *adɨ* ‘name’, *etüd* $\sim$ *etüdü* ‘etude’
- Final devoicing hypotheses are now excluded (exceptions)
- Limiting devoicing to polysyllables correctly avoids exceptions like *ad*, but introduces others, and still doesn’t cover polysyllabic *etüd*
- A costly complication: add list of morphemes to devoicing rule
- Helps illustrate why adding a distinct UR, such as [$\varnothing$ voice], is preferable under these assumptions

The augmented hypothesis

- $\begin{bmatrix} -\text{sonorant} \\ \varnothing\ \text{voice} \end{bmatrix} \rightarrow [-\text{voi}]$ / ___ #
- $\begin{bmatrix} -\text{sonorant} \\ \varnothing\ \text{voice} \end{bmatrix} \rightarrow [+\text{voi}]$ / $+$sonorant___ $+$sonorant
- Somewhere, we should also include cost of allowing three-way lexicon contrast between [$+$voice], [$-$voice], [$\varnothing$ voice] (more on this soon)

A “baseline” analysis

- Since final voicing status is unpredictable, put it in the lexicon
- Analysis has no exceptions, in the sense of surface forms that can’t be generated with UR + rules
- “Standard” approach, given priority for exceptionless hypotheses
- A natural question: should the analysis cover less, using simpler UR’s and admit some exceptions?
- I’ll push here an alternative: the analysis should cover *more*, using even simpler UR’s, more exceptions, but greater predicitivity

Admitting some exceptions

- An obvious intermediate hypothesis
  - Turkish final obstruents are either [+voi] or [−voi] underlyingly
  - Final devoicing rule
  - Some exceptions to final devoicing
- But there are many other hypotheses to consider
  - No final devoicing, but some morphemes exceptionally undergo
  - Restricted final devoicing (polysyllables? some places of articulation?)
  - Final devoicing + intersonorant voicing
- These represent different balance of grammatical complexity vs. number of exceptions
  - How to navigate this space, without the simplifying assumption of 'no exceptions'?

Numbers matter

- Recall restatement of SPE objective as a unified function

$$\text{argmax}\left(\frac{1}{\text{length(grammar)}} - \infty\big(\text{num of exceptions}\big) \right)$$

- By allowing exceptions, we're considering decreasing the cost of exceptions (non-infinite)
- Weights between 0 and $-\infty$ allow some exceptions, in trade-off with complexity
- Hypotheses on previous slide must be compared for how many exceptions they create

Reducing exceptions, increasing predictiveness

- Number of exceptions is the most obvious metric of what the analysis fails to predict
- Information in the lexicon is also not predicted
- Evaluating analyses based on exceptions, but not on how much is specified in the lexicon, continues to privilege overloading the lexicon to avoid exceptions
- Makes sense to shift from ‘how can we reduce exceptions?’ to ‘how much can a good grammar predict?’

Predictability

How (un)predictable are contrasts?

- Listing a binary distinction in the lexicon as $+/-$ represents complete uncertainty
  - 50% probability of each, information content $=$ 1 bit
- In actuality, few contrasts are truly 50/50
  - Voicing: [$-$voi] generally favored for obstruents, [$+$voi] strongly favored or required for sonorants
  - Place: asymmetries, especially by position
- In principle, grammar of preferences can reduce burden on lexicon
  - E.g., radical underspecification of ‘unmarked’ feature values

Not markedness of a single value

- Voicing: preferred value often depends on context
  - Obstruents: [$-$voi]
  - Sonorants: [$+$voi]
- Turkish: contexts are more complex, somewhat language-specific
  - Hayes (1995) cites observation by Lewis (1967) that *all* polysyllabic stems are voiced when suffixed
  - Becker et al. (2011) show that this is not absolute (loanwords, etc. included), but it is a strong trend, especially for labials and dorsals
  - Conversely: contexts where voiceless dominates (e.g., coronals in monosyllables)
- No single ‘unmarked’ value, but predictable nonetheless

A couple possibilities for Turkish

- Leave final voicing unspecified in lexicon, assign by rule in suffixed forms
- Use isolation/unsuffixed form as UR, voice by rule in suffixed forms (Hayes, 1995)
- In both cases, the grammar is complex, but the lexicon is simpler

Not an isolated quirk of Turkish

- Final vowel deletion in Yidiɲ (Hayes, 1997)
  - $XV_1C_1C_2V_2 \rightarrow XV_1C_1$, if:
    - Word is odd-syllabled
    - C_1 is permissible word-finally
- Examples:

gidaːn	gidanu-ŋgu	‘moon(-ERG)’
buɲa	buɲaː-ŋ	‘woman(-ERG)’

- In principle, learners need to hunt around to find quality of deleted vowels in even-syllabled words

Not an isolated quirk of Turkish (*cont.*)

- In actuality, quality is usually predictable
  - /$CV_1C_0VNV_2$/: vowel is *u*
  - Other /$CV_1C_0VCV_2$/: $V_2 = {}_1$
- Hayes shows that these statements are not true of morphemes in general, but only morphemes subject to final V deletion
- Interpretation: lexicon contains form with deletion, grammar ‘restores’ the vowel (with a few exceptions)

Predicting the unpredictable, cont.: Dutch

- Like Turkish, Dutch has voicing alternations for stem-final obstruents

Sg.	Pl.	Gloss
vut	vutən	‘foot’
bɛt	bɛdən	‘bed’
hɔnt	hɔndən	‘dog’
dif	divən	‘thief’
zak	zakən	‘pocket’
mœys	mœyzən	‘mouse’

- Problem set ‘A+’: voicing is underlying, devoicing process in final position

Gradient restrictions

OBSTRUENT	VOICED		VOICELESS		TOTAL	
	#	%	#	%	#	%
P	20	9	210	91	230	100
T	177	25	542	75	719	100
S	151	33	300	66	451	100
F	116	70	50	30	166	100
X	127	97	4	3	131	100

TABLE 1. Morphemes ending in underlyingly voiced or voiceless final obstruents in the CELEX data set, by type of obstruent.

- Ernestus & Baayen (2003): Stem-final voiced and voiceless obstruents are not equally common in lexicon
- Distribution also depends on place of articulation, continuancy, and other factors

Gradient restrictions, cont.

OBSTRUENT TYPE	NO PRECEDING CONSONANT				PRECEDING SONORANT				PRECEDING OBSTRUENT			
	VOICED		VOICELESS		VOICED		VOICELESS		VOICED		VOICELESS	
	#	%	#	%	#	%	#	%	#	%	#	%
P	0	0	52	100	–		–		–		–	
T	55	32	116	68	22	71	9	29	4	17	20	83
S	78	82	17	18	14	67	7	33	0	0	11	100
F	62	100	0	0	1	100	0	0	–		–	
X	65	98	1	2	–		–		–		–	
	$\chi^2(4) = 231.1, p < 0.001$				$\chi^2(2) = 0.551, p = 0.759$				$\chi^2(1) = 0.751, p = 0.386$			

TABLE 3. Morphemes ending in underlyingly voiced or voiceless obstruent in the CELEX data set, by type of obstruent and type of preceding segment (all morphemes end in syllables with long vowels).

- Preceding consonant matters, in consonant clusters
- Effect differs depending on preceding vowel quality and length
- Net result: high degree of predictability

Gradient restrictions, cont.

CODA	[i, u, y]				LONG VOWELS				SHORT VOWELS			
	VOICED		VOICELESS		VOICED		VOICELESS		VOICED		VOICELESS	
	#	%	#	%	#	%	#	%	#	%	#	%
- P	0	0	20	100	0	0	52	100	20	22	72	78
- T	13	24	41	76	55	32	116	68	21	15	113	85
Son T	5	50	5	50	22	71	9	29	56	35	103	65
Obstr T	0	0	11	100	4	17	20	83	1	1	124	99
- S	24	71	10	29	78	82	17	18	3	3	115	97
Son S	1	25	3	75	14	67	7	33	31	41	45	59
Obstr S	0	0	7	100	0	0	11	100	0	0	85	100
- F	21	78	6	22	62	100	0	0	4	10	38	90
- X	11	100	0	0	65	98	1	2	32	91	3	9

Son = sonorant consonant   Obstr = obstruent

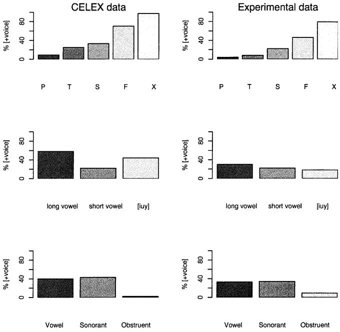
TABLE 5. Underlyingly voiced and voiceless obstruents for morphemes containing [i, u, y] and for corresponding morphemes with long and short vowels, by type of final coda.

- Preceding vowel quality and length matters
- Distribution also depends on place of articulation, continuancy, and other factors
- Net result: high degree of predictability

Synchronically 'active' prediction

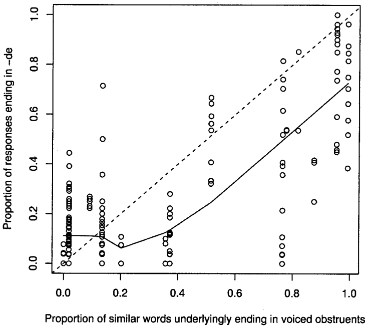
- Ernestus & Baayen (2003): wug test of novel verbs (192 types, most possible rhymes)
- Presented neutralized 1SG (*ik* [tif])
- Prompted to produce past ([tifdə]/[tivdə])

Synchronically ‘active’ prediction



- Result: participants largely reproduce lexical trends
- Ernestus & Baayen (2003) suggest that participants are using their knowledge of the lexicon
- Also compatible with Hayes’s approach: grammar predicts suffixed forms (with some exceptions)

Synchronically 'active' prediction



- Result: participants largely reproduce lexical trends
- Ernestus & Baayen (2003) suggest that participants are using their knowledge of the lexicon
- Also compatible with Hayes's approach: grammar predicts suffixed forms (with some exceptions)

How does grammar predict the ‘unpredictable’?

- Hayes’s approach: store isolation form, use rules to project suffixed forms
  - Dutch, Turkish: lexical representation is (generally) voiceless, rules voice final obstruents in certain contexts
  - Grammar is imperfect $\Rightarrow$ exceptions (item-specific ‘rules’?)
- There are alternatives, e.g.: lexicon leaves unspecified unless exceptional (Kiparsky, Inkelas, etc.)
- What these approaches have in common: final obstruent voicing not ‘fully contrastive’

From ‘exceptionality’ to ‘unpredictability’

- If we penalize analyses for exceptions to the grammar, this line of analysis seems to lead to an uncomfortably large number of exceptions
- However, this ignores the fact that there are two types of unpredictability in a generative analysis
  - Lexical specifications
  - Exceptions to rules
- Rather than eliminating exceptions, perhaps a more sensible learning objective: minimize the information needed to describe the data
  - Concentrate probability mass, maximize likelihood
  - Side note: this makes contact with MDL approaches to phonology (Rasin & Katzir, 2015, 2016)
- Payoff to predicting features in grammar (even imperfectly) is reduced information in lexicon

Why exceptions?

- The more ambitious our analysis is about capturing trends in the data, the more exceptions there will be
- Trade-off not only between grammatical simplicity and exceptions, but also lexical information content and exceptions
- Limits on grammar also lead to exceptions, so by identifying exceptions, we learn about what grammars can/can't capture

Where next

- So far: introduction to the idea that reasoning about what's rule-governed vs. exceptional makes most sense in a framework where grammars try to predict as much as possible
- Discussion of how the grammar does this has been high level, conceptual so far
- Next time: introduce one concrete way of doing this (Minimal Generalization Learner), examine predictability (and exceptionality) in cases like Dutch, Turkish

References

Becker, Michael; Nihan Ketrez; and Andrew Nevins. 2011. The surfeit of the stimulus: Analytic biases filter lexical statistics in Turkish laryngeal alternations. *Language* 87.84–125.

Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.

Ernestus, Mirjam, and R. Harald Baayen. 2003. Predicting the unpredictable: Interpreting neutralized segments in Dutch. *Language* 79.5–38.

Halle, Morris. 1978. Knowledge unlearned and untaught: What speakers know about the sounds of their language. *Linguistic theory and psychological reality.*, ed. by Morris Halle, Joan Bresnan, and George Miller, 294–303. MIT Press.

References (*cont.*)

HALLE, MORRIS, and GEORGE N. CLEMENTS. 1983. *Problem book in phonology: A workbook for introductory courses in linguistics and in modern phonology*. Cambridge, MA: The MIT Press.

HAYES, BRUCE. 1995. On what to teach the undergraduates: Some changing orthodoxies in phonological theory. *Linguistics in the morning calm*, ed. by Ik-Hwan Lee, vol. 3, 59–77. Seoul: Hanshin.

HAYES, BRUCE. 1997. Anticorrespondence in Yidiɲ. UCLA manuscript.

INKELAS, SHARON, and CEMIL ORHAN ORGUN. 1995. Level ordering and economy in the lexical phonology of Turkish. *Language* 71.763–793.

INKELAS, SHARON; ORHAN ORGUN; and CHERY ZOLL. 1997. The implications of lexical exceptions for the nature of grammar. *Derivations and constraints in phonology*, ed. by Iggy Roca, 393–418. Oxford University Press.

References (*cont.*)

KAISSE, ELLEN. 1986. Locating Turkish devoicing. *Proceedings of the west coast conference on formal linguistics 5*, 119–128. Stanford, CA: Stanford Linguistics Association.

LEWIS, G. L. 1967. *Turkish grammar*. Oxford: Clarendon Press.

RASIN, EZER, and RONI KATZIR. 2015. Compression-based learning for OT is incompatible with Richness of the Base. *Proceedings of nels 45*, ed. by Thuy Bui and Deniz Özyıldız, vol. 2, 267–274.

RASIN, EZER, and RONI KATZIR. 2016. On evaluation metrics in Optimality Theory. *Linguistic Inquiry* 47.235–282.