The morphological basis of paradigm leveling*

Adam Albright
University of California, Los Angeles

1 Introduction

In older stages of Latin, many nouns exhibited paradigmatic *s* ∼ *r* alternations created by rhotacism in suffixed forms, as in (1a). In the period immediately before Classical Latin, these *s* ∼ *r* alternations were eliminated by extending the *r* to the nominative form. At approximately the same time, an independent change shortened long vowels before word-final sonorants, resulting in the paradigm in (1b):[1]

(1) Elimination of *s* ∼ *r* alternations in Pre-classical Latin

(a)	*Pre-leveling*	(b)	*Post-leveling*		
	[honoː**s**]	>	[hono**r**]	'*honor*'	(nom.sg.)
	[honoː**r**is]		[honoː**r**is]		(gen.sg.)
	[honoː**r**iː]		[honoː**r**iː]		(dat.sg.)
	[honoː**r**em]		[honoː**r**em]		(acc.sg.)
	etc. . .		*etc. . .*		

*I would like to thank the following people for their helpful comments and discussions: Bruce Hayes, Donca Steriade, Brent Vine, Carson Schütze, Angelo Mercado, and audiences at UCLA, MIT, and the ZAS Conference on Paradigm Uniformity, as well as two anonymous reviewers. All remaining errors and omissions are, of course, my own. This material is based upon work supported under a National Science Foundation Graduate Fellowship.

[1]There is some evidence for an intermediate stage in the late Pre-Classical stage in which [honoːr] with long [oː] was at least one possible variant of the nominative form. Early poets such as Plautus and Ennius still allowed final long [o:r], using it to satisfy metrical requirements for heavy syllables, and also sometimes used [-o(ː)r] instead of [-oːs], e.g., Ennius 545 (Skutsch 1985) *Clā́m[ō:]r ād cā́elūm vṓlvēndŭs pĕr ā́ethĕră vā́gĭt*; also Ennius 409, 428. As far as I know, there is no evidence concerning how "clean" this intermediate stage was – the shortened variant [honor] may also have been used from the very beginning of the [s] > [r] change. The analysis that I propose here is compatible with the existence of an intermediate [honoːr] stage, but does not rely on it. It would, however, make a crucial difference for some other analyses, such as the Uniform Exponence analysis sketched below.

As Hock (1991, pp. 179-190) points out, the change from [honoːs] to [honor] can be described in two different ways. The first is as a *four-part analogy*, in which [honoːs] was influenced by words that already had an [r] in the nominative, such as [soror] 'sister':

(2) Extension of *r* by four-part analogy:
[soroːris] : [soror] :: [honoːris] : X (X = [honor])

Although the four-part analogy notation expresses the change as the influence of one particular lexical item (in this case, [soror]), it is generally recognized that such changes are actually due to the collective influence of many words, such as [soror], [kruor] 'blood', and also the numerous agentive nouns ending in *-or* ([oːraːtor] 'speaker', [gladiaːtor] 'gladiator', etc.). However, even when we recognize that the four-part analogy notation is just a shorthand for the influence of a larger pattern, it is far from an explanation of the change. How many words does it take for speakers to construct such an analogy? How similar do they have to be? More importantly, it has often been noted that four-part analogy can not tell us why the influence was not from [honoːs] to [soror] ([honoːris]:[honoːs] :: [soroːris]:X, X=[soroːs]) (Hock 1991; Barr 1994; Kiparsky 1997), or why the change was not in the opposite direction, undoing rhotacism (something like [soror]:[soroːris] :: [honoːs]:X, X = [honoːsis]). This last question is especially puzzling, because the actual change in (2) violates the tendency for analogical change to extend from more "basic" or underived forms to less basic, or derived forms (Kuryłowicz 1947).

The four-part analogy notation can equally well capture the leveling of alternations ([honoːs] > [honor]) or the extension of alternations ([soror] > [soroːs]). An alternate account of the spread of [r] to the nominative in [honor] is as *paradigm leveling*, with the nominative form changing to match the remainder of the paradigm ([honoː**r**is], [honoː**r**iː], [honoː**r**em], etc.). The pressure to level paradigms has been formalized in Optimality Theory as constraints on paradigm uniformity or uniform exponence (Kenstowicz 1995; Steriade 2000); for example, Kenstowicz schematizes the [honoːs] > [honor] change as the promotion of a constraint demanding uniformity in noun paradigms (UE). In the first stage, shown in (3), the ban on intervocalic [s] (*VsV) outranks Faithfulness for /s/ (Faith-/s/), yielding rhotacism in suffixed forms. Faith-/s/ in turn outranks UE, meaning that rhotacism does not overapply in the nominative form. The result is a paradigm with $s \sim r$ alternations, as in (3a). Note that candidate (c) contains a vowel length alternation ([or] $\sim$ [oːris] – Kiparsky 1997; Hale et al. 1998; Baldi 1999, p. 323), for which Kenstowicz does not record a UE violation. It is entirely possible that there are separate UE constraints for different alternations, with UE for $s \sim r$ alternations ranked higher than UE for vowel length alternations. If that is the case, then UE($s \sim r$) is the relevant constraint here, and this what I will assume.

(3) Stage 1: *VsV ≫ Faith-/s/ ≫ UE(*s* ∼ *r*)

/hono:s/ (nom.), /hono:s-is/ (gen.), /hono:s-em/ (acc.)		*VsV	Faith-/s/	UE(*s* ∼ *r*)
a. ☞	[hono:s], [hono:r-is], [hono:r-em]	√	**	*
b.	[hono:s], [hono:s-is], [hono:s-em]	*(gen.)!*(acc.)	√	√
c.	[honor], [hono:r-is], [hono:r-em]	√	***!	√ (?)

Under this analysis, the change consists of promoting UE over Faith-/s/, so that rhotacism overapplies in the nominative, as in (4). This leaves two possible candidates for the nominative: [hono:r] (candidate (c)) and [honor] ((d) and (e)). The first of these violates a high-ranking phonotactic constraint against final [o:r] in Latin, favoring a paradigm with shortening in the nominative ((d) or (e)); of these, the paradigm that preserves long [o:] in the oblique forms is more faithful to the long [o:] of the input (Faith-/V:/), and thus candidate (d) is selected.

(4) Stage 2: *VsV, UE(*s* ∼ *r*) ≫ Faith-/s/, Faith-/V:/

/hono:s/ (nom.), /hono:s-is/ (gen.), /hono:s-em/ (acc.)		*o:r#	*VsV	UE(*s* ∼ *r*)	Faith-/s/	Faith-/V:/
a.	[hono:s], [hono:r-is], [hono:r-em]	√	√	*(nom.)!	**	√
b.	[hono:s], [hono:s-is], [hono:s-em]	√	*(gen.)! *(acc.)	√	√	√
c.	[hono:r], [hono:r-is], [hono:r-em]	*!	√	√	***	√
d. ☞	[honor], [hono:r-is], [hono:r-em]	√	√	√	***	*
e.	[honor], [honor-is], [honor-em]	√	√	√	***	**!*

This analysis avoids many of the problems pointed out by Hale et al. (1998) by limited UE to *s* ∼ *r* alternations. An alternative possibility, however, is that the Uniform Exponence analysis actually captures the creation of the intermediate variant [hono:r] (see fn. 1), with perfect UE satisfaction, and that the underlying form of this word had already been reanalyzed as /honoːr/ (as Hale et al. claim) by the time final vowel shortening occurred, so Uniform Exponence was irrelevant by that stage.

The uniform exponence account formalizes the intuition that paradigm leveling is due to a pressure for nonalternating paradigms, and that the resulting paradigm is one which (in this case) satisfies both paradigmatic constraints and also general phonotactic constraints of the language (such as *VsV, and *oːr#). However, it leaves many details unaccounted for. Why, for example, was the old *s* ∼ *r* alternation suddenly intolerable, at the same time that a new *o* ∼ *oː* alternation was being created? We might have expected the UE constraint to move above both the *oːr# and *VsV constraints, since there was no crucial ranking between them before the change, yielding a uniform paradigm with overapplication of both rhotacism and shortening: [honor], [honoris], [honorem], etc.). Furthermore, we might expect an increased drive for paradigm uniformity to level other alternations in noun paradigms, but in fact these remained by and large intact:

(5) Alternations preserved in Latin noun paradigms

gloss	'honor'	'city'	'art'
alternation	[o] ∼ [oː]	[p] ∼ [b]	∅ ∼ [t]
nom.	[honor]	[ur**p**s]	[ars]
gen.	[honoːris]	[ur**b**is] (*[ur**p**is])	[ar**t**is] (*[aris])
dat.	[honoːriː]	[ur**b**iː] (*[ur**p**iː])	[ar**t**iː] (*[ariː])

Another unexplained mystery is why was UE promoted over Faith-/s/ to extend rhotacism, rather than Faith-/s/ being promoted over *VsV to eliminate rhotacism? As mentioned above, the rhotacism constraint (*VsV) plays a crucial role in Kenstowicz' analysis, ensuring that the resulting paradigm will have uniform [r] and not uniform [s]. However, we could just as easily have used the same analytical device (historical reranking of two constraints) to predict the opposite change, promoting the IO Faith-/s/ constraint to yield uniform [s]. Thus, the mere existence of an active rhotacism constraint at one stage in the grammar is not sufficient to explain why it should continue to be true at the next stage in the grammar. A proposal by McCarthy (1998) that output-to-output constraints start at the top of the grammar in the initial state might help to explain the tendency for uniform exponence to move up in grammars over time. If this is correct, then we need only assume that learners sometimes fail to demote UE below the relevant markedness constraints – but why would one generation of Latin learners suddenly fail to apprehend the correct ranking of Faith-/s/ $\gg$ UE? If uniform exponence is to have any explanatory force in accounting for paradigm leveling, we would ideally like to be able to predict when such UE $\ll\!=\!\gg$ Faithfulness flips are likely to occur, and which faithfulness constraints will be demoted.

A uniform exponence analysis also fails to capture various other details about the [hono:s] > [honor] change. Most notably, the spread of [r] to the nominative form was complicated by the fact that it was restricted primarily to polysyllabic, non-neuter nouns such as [hono:s] 'honor' (masc.) and [arbo:s] 'tree' (fem.), shown in (6a) (Hock 1991; Barr 1994; Kiparsky 1997). Monosyllabic nouns, such as [flo:s] 'flower' (masc.) were not affected (6b), nor were polysyllabic neuter nouns, such as [korpus] 'body' (neut.) (6c).

(6) Leveling restricted to masc. and fem. polysyllables

a.	hono:s	>	honor	'honor' (*masc.*)	
	arbo:s	>	arbor	'tree' (*fem.*)	
	odo:s	>	odor	'odor' (*masc.*)	
	(augus)	>	augur	'omen' (*masc.*)	
b.	flo:s	>	flo:s	'flower' (*masc.*)	(≯flor)
	o:s	>	o:s	'mouth' (*neut.*)	(≯or)
	mo:s	>	mo:s	'custom' (*masc.*)	(≯mor)
c.	korpus	>	korpus	'body' (*neut.*)	(≯korpor, korpur)
	tempus	>	tempus	'time' (*neut.*)	(≯tempor, tempur)
	onus	>	onus	'burden' (*neut.*)	(≯oner, onur)

Furthermore, the new [r] forms appear to have replaced the older [s] forms relatively slowly; in the historical period we find both [hono:s] and [honor], [odo:s] and [odor], [arbo:s] and [arbor], [lepo:s] and [lepor] 'charm', [labo:s] and [labor] 'labor' (Leumann 1977, p.179).[2] Both of these problems could be handled by various means – UE could be restricted to non-neuter polysyllables in some way, for example, and the free variation could be accomplished by the gradual promotion of stochastically ranked constraints (Boersma and Hayes 2001). As above, the real problem is not in finding theoretical machinery that can describe the change; it is in understanding why the change should have occurred in this direction, to these particular words.

In sum, there are a variety of questions that must be answered if our understanding of the Latin [honor] analogy is to move beyond description to actual explanation:

[2]It is difficult to know in many cases whether the occurrence of forms like *hono:s* alongside *honor* reflects free variation, or simply literary archaism. Cicero, for example, systematically used the form *hono:s* instead of *honor*, in both philosophy and oratory texts (elevated styles), as well as in letters (potentially less elevated/archaic); at the same time, he used *labor* instead of *labo:s* in all three contexts.

(7) a. Why did was a basic, “unmarked” isolation form (the nominative) rebuilt on the basis of more marked suffixed forms, contrary to the usual direction of analogical change? (Lahiri and Dresher 1984; Bybee 1985, chap. 3)

b. What role (if any) did similar words, like [soror], play in the change? is there a minimum number of such words that are necessary to effect such a change?

c. Why did [honoːs] change to [honor], and not [soror] to [soroːs]?[3]

d. Why were monosyllables and neuters generally not affected?

e. Why did both [oːs] and [or] variants persist for so long, and why was no variation induced in etymological [or] words such as agentives?

In this paper, I propose that the change of [honoːs] to [honor] was driven by more than just a phonological change involving paradigm uniformity constraints. I suggest that it was actually a morphological effect, resulting from the way that paradigms are learned. I present a computationally implemented model of paradigm acquisition that proceeds in two stages: first, it compares all of the available paradigms and selects the base form that allows the remainder of the paradigm to be projected as reliably and effectively and efficiently as possible, in a way which will be quantified in section 2. It then develops a gradient morphological grammar to project the rest of the paradigm, using the *minimal generalization* method of morphological rule induction, developed by Albright and Hayes (1999, 2000). When this process is applied to Latin, it emerges that the preferred base is an oblique form, not the nominative form, accounting for the unusual direction of the analogical change (7a). Furthermore, when an oblique form is used to project nominative forms, the system makes essentially the right predictions for the *honor* analogy: [-oːs] nominatives are strongly favored for monosyllabic and neuter nouns, [-or] nominatives are preferred for polysyllabic masculine and feminine nouns with [-oːs] remaining a strong second choice, and [-or] is strongly favored for agentive nouns.

This analysis is similar in spirit to the four-part analogy explanation of the change, but the gradient nature of the rules in this system gives us a quantitative

[3] A reviewer points out that the paradigm of [honoːs] already had [r] forms in it, while the paradigm of [soror] never had [s] forms. Following Steriade (1994) we might call this the “lexical conservatism” analysis, in which speakers may only use or extend allomorphs that are already attested. It is possible that a lexical conservatism analysis could explain this part of the asymmetry ([honor] but not *[soroːs]), but there are hints that Latin speakers were not always bound by lexical conservatism – for example, see fn. 12 regarding the use of [femus] for [femur], which otherwise never had an [s] anywhere in the paradigm. The lexical conservatism analysis would also tell us nothing about (7d), since monosyllables and neuters also had an available [r] allomorph which could have been extended to the nominative singular.

expression of the influence of other lexical items (7b), and also helps predict which pattern should win out for each class of words (7cd). Finally, the close competition between [-oːs] and [-or] for many forms suggests an intriguing interpretation for the persistence of [-oːs] nominatives into the attested period: I conjecture that perhaps they were not merely a conservative retention of memorized archaic forms, but that they may also have supported to a certain extent by the synchronic grammar of Classical Latin. In other words, when Latin speakers heard an archaism like [honoːs], even if it was not the synchronically preferred form, it may have struck them as moderately grammatical, allowing [-oːs] forms to persist as an archaism much longer than some other archaic features. Clearly more philological work is needed to support this hypothesis, but it is a good example of how the current model makes predictions about not only how forms can be innovated, but how they may also be retained or lost.

2 A model of base discovery

The premise of the model proposed here is that language learners seek the most accurate rules possible to describe their language. In order to do this, they must select certain forms to memorize, and then they must construct rules to derive the remaining forms. In this paper, I will pursue the hypothesis that learners initially explore many different candidates for base status, comparing the ease with which other parts of the paradigm can be derived from each potential base. Once a sufficient number of words has been learned, a base is chosen, and the learner focuses on constructing a grammar to derive the rest of the paradigm.

The choice of which forms are memorized and which are derived often has a substantial effect on the ability to construct accurate rules. Consider, for example, the hypothetical language in (8), which has a single nominative ending ([-us]), but two possible genitive endings ([-iː] and [-oris]).

(8) Neutralization in the nominative

nom.		*gen.*
[gluptus]	∼	[gluptiː]
[nokus]	∼	[nokiː]
[reptus]	∼	[reptoris]
[kortus]	∼	[kortoris]

In this case, there are two classes of words: the [-us] ∼ [-iː] words and the [-us] ∼ [-oris] words. The two classes are neutralized in the nominative, so the mapping from the nominative to the genitive is unpredictable. If a learner were to memorize just the nominative form, then two rules would be necessary to project genitives ([-us]→[-iː] and [-us]→[-oris]). Each of these rules would only have 50% accuracy in the existing lexicon, each covering 2 out of 4 words. Furthermore, for a hypothetical new word [tulpus], there would be two possible genitives ([tulpiː] and [tulporis]), each with a 50% chance of being right, leaving a 0% margin to decide by.

In such a situation, the learner should recognize that the genitive form simply contains unpredictable information that cannot be derived by rule, but must rather be memorized. If we were to derive nominative forms from genitive forms, we still need two rules ([-oris]→[-us] and [-iː]→[-us]), but each would have 100% accuracy in the existing lexicon.[4] The hypothetical new genitive forms [tulpiː] and [pulkoris] could only have nominative forms [tulpus] and [pulkus], respectively, with 100% certainty.

Thus, it appears that we can identify the "most informative" base algorithmically, given two abilities: a capacity for discovering the rules needed to derive one part of the paradigm from another part, and a set of metrics to estimate how effective or reliable the rules are in each direction.

2.1 The minimal generalization model of rule induction

One system for hypothesizing morphological rules is the *minimal generalization* algorithm, outlined by Pinker and Prince (1988) and developed and implemented by Albright and Hayes (1999). The minimal generalization procedure starts with pairs of related forms, such as the ones given above in (8). It compares the left and right member of each pair, in order to determine what material is constant across both forms of the word, and what material changes. The result is a word-specific rule for each input pair, as in (9):

(9) Word-specific morphological rules for the words in (8)

a. [us] → [iː] / glupt___#
b. [us] → [iː] / nok___#
c. [us] → [oris] / rept___#
d. [us] → [oris] / kort___#

If the learner's only task was to reproduce inflected forms that had already been heard, then these rules would be sufficient. However, in real life, speakers

[4] I will leave aside for the moment the possibility of memorizing just abstract roots (/glupt-/, /nok-/, etc.), because such an ability would not help significantly in this hypothetical language; we could have a single nominative rule adding [-us] to roots, but the words with [-oris] genitives would either have roots [rept-, [kort-, requiring a separate [-oris] rule for the genitive, or they would have roots [reptor- and [kortor-, requiring readjustment rules to delete the [or] in the nominative.

	residue	shared features	shared segments	change
comparing A:	re	p	t	us → oris
with B:	ko	r	t	us → oris
yields C:	X	$\begin{bmatrix} \text{-syl} \\ \text{+cons} \\ \text{-nas} \\ \text{-lat} \\ \text{-dors} \\ \textit{etc.} \end{bmatrix}$	t	us → oris

Figure 1: Minimal generalization to discover environments

must frequently produce new forms of words – particularly in richly inflected languages. Therefore, it is necessary to be able to generalize. Under the minimal generalization approach, this is done by comparing word-specific rules that share the same change. For example, in this hypothetical language, [reptus] and [kortus] both exhibit the change [us] → [oris]; these two words are thus compared further to see whether there is some shared property that may be conditioning this particular change. Figure 1 shows how two words are compared to construct the environment for a new, generalized rule.

Such pairwise comparisons are iterated across the entire lexicon, yielding a sizable set of (often redundant) generalized rules for each morphological change. When a change occurs in a diverse set of phonological environments, then the generalization procedure in Figure 1 can yield very general rules; for example, comparing the words [reptus] ∼ [reptoris] and [kraus] ∼ [kraoris] would yield the very general rule [us] → [oris] / ___# (assuming that [a] and [t] have no feature values in common).

In addition to hypothesizing rules, we also need a way to keep track of how accurate, or trustworthy the rules are. This is done by calculating the *reliability* of each rule, defined in (10):

(10) Definition of a rule's reliability :

$$\frac{\text{\# of forms included in the rule's structural change (= \textit{hits})}}{\text{\# of forms included in the rule's structural description (= \textit{scope})}}$$

For example, the generalized rule created in Figure 1 has a structural description that covers three of the words in this hypothetical language ([gluptus], [reptus], and [kortus]), but only two of those words actually take [-oris] in the genitive ([reptoris] and [kortoris]). Thus, the reliability of the rule is 2/3 = .67. These reliability ratios are then adjusted using lower confidence limit statistics to yield a *confidence* value, following a suggestion by Mikheev (1997), so rules that attempt to cover just a few forms are penalized for being too unambitious. For example, when using a confidence level of α = .95, a rule with reliability of 5/5 is assigned a confidence of .825, while a rule with reliability of 1000/1000 is assigned a confidence of .999. (For details of calculating confidence limits, see Mikheev 1997.)

This generalization algorithm constructs rules that are quite traditional in their format, but the resulting grammar is unconventional in allowing a large number of specific and often overlapping or redundant rules. When the grammar is used to derive new forms, it is therefore possible that there is more than one applicable rule. In this case, the rule with the highest reliability value is the one which gets to apply. Continuing with the example from the previous paragraph, we can consider the predictions of the grammar for a novel word [loptus]. The comparison of [gluptus] and [nokus] would yield a rule of [us] → [iː] after voiceless stops, yielding an output of [loptiː] for the novel word. The reliability of the rule used to derive [loptiː] is 2/4 = .5, adjusted downward to .31, so we can say that [loptiː] is projected with a confidence of .31. The comparison of [reptus] and [kortus], on the other hand, yields a rule of [us] → [oris] specifically after [t], which works for 2/3 of the existing words, for a reliability of .67, adjusted downward to .40. This rule generates the output [loptoris] for the novel word, with a confidence of .40. Thus, the [us] → [oris] rule is more reliable in this environment, and [loptoris] would be the winning output. The winning margin between the winning output and the next best competitor in this case would be .40 - .31, or .09.

The minimal generalization approach is designed to discover which morphological processes have the highest reliability, and in which phonological environments. In many cases, however, the true reliability of a morphological process may be obscured by phonological changes; consider, for example, the language in (11):

(11)

nom.		*gen.*
[neks]	∼	[negis]
[arks]	∼	[argis]
[teks]	∼	[tekis]
[flurs]	∼	[fluris]

In this case, the standard analysis would be to say that all of the words behave the same morphologically (nominative [-s], genitive [-is]), but there is a phonological process enforcing voicing agreement in final obstruent clusters, yielding /negs/ → [neks], /args/ → [arks]. The procedure outlined above, however, would analyze [neks] ∼ [negis] with the morphological rule [ks] → [gis] / ne___#. What we need is a way to discover that this word could also be analyzed as [s] → [is] / [neg]___#, with a further phonological process fixing the illegal word-final *[gs]. Note that it would not be reasonable to equip the morphological learner with the necessary phonological rule or constraint ranking ahead of time, because the evidence for such a rule would come precisely from alternations in morphologically related words.

The Albright and Hayes implementation of minimal generalization learning uses the following approach to solve this problem: the learner is assumed to have prior knowledge of what sequences are phonotactically legal in her language, but not the specific rules governing alternations.[5] When assessing the reliability of a rule, the learner then checks whether the output of the rule would yield an illegal sequence. For example, in learning rules for the forms in (11) in the genitive $\rightarrow$ nominative direction, the forms [tekis] $\sim$ [teks] and [fluris] $\sim$ [flurs] would generalize to yield a morphological rule [is] $\rightarrow$ [s] / ___#. In order to calculate the reliability of the rule, we can try applying it to all of the input forms ([fluris], [tekis], [argis], [negis]). This yields the correct result for [flurs] and [teks], but the wrong result for *[args] and *[negs]. However, given the knowledge that [gs] is phonotactically ill-formed, the learner posits a phonological rule changing /gs/ to [ks]. This rule, when applied after the morphological operation, allows the [is] $\rightarrow$ [s] rule to yield the correct outcomes for [arks] and [neks] as well, for a reliability of 4/4.

The system described in this section provides an automated method for hypothesizing a set of morphological rules and estimating the effectiveness of individual rules in explaining the input data. The examples so far have involved the mapping between just two slots in the paradigm (e.g., *nom.* $\rightarrow$ *gen.*); I will refer to the set of rules between just two forms as a *subgrammar*. The complete grammar of a language would require many such subgrammars – e.g., *nom.* $\rightarrow$ *gen.*, *nom.* $\rightarrow$ *dat.*, and so on. What remains, then, is to define a way to estimate the reliability of an entire subgrammar, in order to decide which parts of the paradigm can be derived relatively effectively by rule, and which forms should be memorized instead.

2.2 Criteria for base selection

How can we decide that it is easier to project form X from form Y, rather than the other way around? In other words, what is the criterion for deciding that the subgrammar of rules from X $\rightarrow$ Y is "better" than the subgrammar from Y $\rightarrow$ X? Given a subgrammar of rules with reliability values attached to them, there are various possibilities for how to evaluate the subgrammar as a whole.

[5]Recent research on infants has shown that children are able to distinguish phonotactically illegal sequences from legal ones at a remarkably early age (Friederici and Wessels 1993; Jusczyk, Friderici, Wessels, Svenkerud, and Jusczyk 1993; Jusczyk, Luce, and Charles-Luce 1994); see Hayes (to appear) and Smolensky et al. (to appear) for overviews, and specific proposals for how phonotactic distributions could be encoded grammatically at the earliest stages.

One approach would be to examine inherent properties of the subgrammars involved. We could, for instance, compare the number of rules that are hypothesized in mapping X → Y and Y → X. Another possibility would be to measure the average reliability of the rules in the subgrammar.

A second approach is to examine the performance of the grammar in deriving forms. The word-specific rules in (9) ensure that the system will always be able to derive the correct answer for words that were part of the input data. However, if we deprive the system of these word-specific rules and then ask it to derive outputs for the training data, we can measure how well the system is able to capture the data using higher-level generalizations. One intuitive measure of the success of a subgrammar is the proportion of times that the "correct" (attested) form is chosen as the winning form. Another possible measure is the average reliability of the rules that derive the winning outputs, on the theory that more ambiguous mappings will yield rules with lower reliabilities. A third way to measure the performance of a mapping would be to count the average number of competitors for each output, since there will always be at least two possible outcomes for an ambiguous input. Finally, we could measure the margin between the winning output and the next highest competitor, since truly ambiguous mappings will lead to "coin-toss" situations where two possible outcomes are equally likely.

A reliable or effective subgrammar, then, should have the following properties: high reliability rules, high accuracy in deriving the correct form for the training data, high confidence in the winning output, few competing outputs for each input form, and a large margin between the winning output and the next highest competitor. Although these metrics reflect logically distinct features of grammars, and we could imagine cases in which they did not agree with one another (for example, a grammar with high accuracy but low winning margins), in practice they turn out to be highly correlated with one another; therefore, rather than selecting a single metric or coming up with some formula to combine them, I will simply present the results of all of these metrics side-by-side, for the reader to compare.

For the purposes of this paper, I will assume that a learner faced with the task of learning Latin noun paradigms is attempting to identify one single form that will serve as the base for the entire paradigm. This is not the only possible approach. We could imagine, for example, a system under which the learner is allowed to choose multiple bases, upon discovering that certain mappings (such as between the dative and the ablative) are nearly 100% predictable, while other mappings (such as between the nominative and the genitive) are not nearly so predictable. Indeed, a multiple-base approach of this type would be required to learn something like the standard dictionary entries for Latin nouns, which list two forms (the nominative and genitive singular). Confining the learner to a single base form, at least for the purposes of learning a "smallish" paradigm of 5 cases and 2 numbers, provides a more restrictive initial hypothesis that makes stronger predictions about possible paradigmatic changes.

2.3 Results for Latin noun paradigms

In the small, hypothetical examples discussed above, the neutralizations in the nominative made the genitive → nominative mapping clearly easier than the nominative → genitive mapping. In real languages, however, the situation is rarely so clear. Neutralizations typically affect only a subset of the segments in the language, so the uncertainty that they cause may only affect a small number of words. Furthermore, neutralizations often affect different parts of the paradigm for different words. Thus, it is not always easy to intuit whether a mapping is easier in one direction than the other, or the magnitude of the asymmetry.

The question of interest for this paper is whether Latin nouns were easier to project in the oblique → nominative direction than vice versa. Latin nouns are traditionally divided into five classes, or "declensions," each of which was inflected for five major cases: the nominative, genitive, dative, accusative, and ablative. (Two additional cases, the vocative and locative, were almost always identical to other cases.) A full description of all of the declensions and their subclasses is clearly beyond the scope of this paper – see Leumann (1977), Kühner (1912), or Allen (1903) for in-depth discussions. What is important here is that the distinctions between many of these classes were neutralized or nearly-neutralized in various parts of the paradigm.

(12) Neutralizations in the nominative

a.		*nom.*	*gen.*	*gloss*
		[popul**us**]	[populiː]	'people'
	vs.	[man**us**]	[manuːs]	'hand'
	vs.	[korp**us**]	[korporis]	'body'
	vs.	[gen**us**]	[generis]	'kind'
b.		[ag**er**]	[agriː]	'field'
	vs.	[gen**er**]	[generiː]	'son-in-law'
	vs.	[fraːt**er**]	[fraːtris]	'brother'
	vs.	[kark**er**]	[karkeris]	'prison'
	vs.	[it**er**]	[itineris]	'journey'
c.		[da**ps**]	[dapis]	'feast, banquet'
	vs.	[ur**p**s]	[urbis]	'city'

(13) Neutralizations in the genitive

a.		[populus]	[popul**iː**]	'people'
	vs.	[bonum]	[bon**iː**]	'good(ness)'

b.		[soror]	[sor**oːris**]	‘sister’
	vs.	[honoːs]	[hon**oːris**]	‘honor’
	vs.	[korpus]	[korp**oris**]	‘body’ (*distinguished only by short [o]*)
c.		[miːles]	[miːl**itis**]	‘soldier’
	vs.	[kaput]	[kap**itis**]	‘head’
	vs.	[ankeps]	[ankip**itis**]	‘danger’
d.		[lak]	[la**ktis**]	‘milk’
	vs.	[noks]	[no**ktis**]	‘night’

Some of these neutralizations affect relatively large numbers of words. The neutralization between masculine ([-us]) and neuter ([-um]) second declension nouns in (13a) involves two very large (and productive) classes of nouns. The neutralization caused by voicing agreement in final obstruent clusters in (12c), on the other hand, affected relatively fewer words. An additional complication is that other factors, such as grammatical gender, could help the speaker know which suffix to use in a potentially ambiguous situation – so, for example, the use of [-us] or [-um] in the nominative of a word with [-iː] in the genitive is almost completely predictable given the gender of the word. Thus, if there is an asymmetry in predictability between nominative and other forms, it would be because of differences in the “severity” of the neutralizations involved, and the ability to predict the correct form using gender. Were the neutralizations in the nominative in fact more severe than in other cases?

In order to answer this question, I started with a database of fully inflected classical Latin nouns, prepared in 1997-1998 by a group working under the supervision of Bruce Hayes at UCLA. This database contained all of the nouns with five or more tokens in a lemmatized frequency count from classical texts (Delatte, Evrard, Govaerts, and Denooz 1981), based on a corpus of approximately 800,000 words (582,000 from prose, 212,000 from poetry). Nouns beginning with the letters R through Z were omitted from the current study because the database was found to have incomplete information for many paradigms in this section of the alphabet. Nominative forms were listed in their forms prior to the [honoːs] > [honor] change; in cases of uncertainty, words were listed with a final [s]. The rationale for this was that we are interested in seeing if the model will favor [r] forms in spite of numerous [s] forms in the training data, and we do not want this to be the result of the influence of spurious [r] forms.

The model of base selection being tested here is that learners evaluate the usefulness of prospective bases early in the learning process. Therefore, the only input data which would be available to the learner for comparisons would be the most common words. As an idealization, words with 50 or more tokens in Delatte, Evrard, Govaerts, and Denooz (1981) were selected, for a total of 494 input nouns. Six forms were considered as possible bases: the nominative, genitive, dative, accusative, and ablative singular, and the nominative plural. For each possible base form, training data files were then constructed to project each of the remaining forms (*nom.*→*gen.*, *nom.*→*dat.*, *nom.*→*acc.*, etc.), yielding 30 (=6×5) training sets in total.[6]

Nouns in the input files were listed in phonemic transcription. In order to take into account the effect of phonological processes, a list of illegal sequences was also prepared, including final clusters disagreeing in voicing (*bs#, *gs#, *ds#), final geminates (*ll#, *dd#, *ss#), the clusters *rts, *lts, and *nts, and a few other illegal sequences whose repair caused alternations in the nominative (*oːr#, *kt, *iiː). Because some of these illegal sequences refer crucially to word boundaries, word boundaries were also marked explicitly in the input files with brackets. Since rhotacism is not surface-true in this stage of Latin (cf: [kaːsus] 'fall', [rosa] 'rose'), intervocalic [s] was not included as an illegal sequence.[7] Finally, each noun was provided with a numeric code indicating the grammatical gender and the number of syllables (monosyllabic vs. polysyllabic), since the current implementation of the minimal generalization learner does not have an independent capacity for considering general prosodic properties of words. Token frequencies were also included in the input files, but they were not employed in the simulations reported here.

The input files were submitted to the minimal generalization learner, yielding subgrammars of rules with confidence values. The word-specific rules were then

[6] All of the training input sets and results files for the simulations discussed here, as well as the original database of nouns, can be downloaded from http://www.linguistics.ucla.edu/people/grads/aalbrigh/papers/latin.html.

[7] A reviewer points out that the presence of intervocalic [s] does not necessarily preclude the possibility that rhotacism continued to be a synchronically active process in Latin, possibly restricted to a particular morphological environment, such as /V___+V, where '+' indicates a morpheme boundary. For the purposes of the current model, the synchronic status of rhotacism actually makes very little difference. Including a *Vs+V constraint would improve the reliability of nominative → oblique grammars slightly, because the model could learn to apply rhotacism in mappings like [honoːs] → [honoːris] (instead of [honoːsis]); however, *-oːs* nominatives make up only a smaller fraction of the language as a whole, so improving the model's predictions for this subset of the vocabulary does not make a substantial difference in the calculations reported below. Note also that including a rhotacism constraint does not help the model at all in the oblique → nominative direction, since an oblique form with *-oːris* could come from either underlying /s/ or underlying /r/. More generally, assuming that rhotacism was synchronically active in Latin can help to explain why the paradigm of 'honor' was not leveled to [honoːs], [honoːsis], etc., but it can not explain other facts, like why speakers did not assume that forms like [soroːris] were also the result of rhotacism (predicting the incorrect nominative [soroːs]), or why speakers did not simply tolerate the rhotacism alternation to remain. It could also be added that many authors have tried to make use of the exact opposite intuition: if we assume that rhotacism was no longer synchronically active, then we can understand why [honoːs] and [honoːris] could no longer be related to one another by an automatic phonological process, and why the alternation was then open to leveling (Klausenburger 1979; Wetzels 1984, and others).

eliminated, and the resulting subgrammars were tested on the input forms. The metrics proposed in section 2.2 were calculated for each subgrammar, to obtain an estimate of the usefulness of each slot in the paradigm for predicting the remainder of the paradigm. The results, given in Appendix A, show that the predictability between *all* forms is quite high (over 80%); it is not the case that any form suffers from neutralizations that affect the majority of forms. Nevertheless, the oblique forms tend to be substantially better than the nominative form on almost all of the metrics considered. In Figure 2, the candidates for base status are compared in terms of their mean effectiveness in projecting the five other forms in the paradigm.

As can be seen, the criteria proposed in section 2.2 generally agree on the relative effectiveness of the various forms as possible bases. The only exception is the number of rules in the grammar, which yields an uninterpretably different ranking. However, we saw above that the number of rules failed to distinguish even the simple four-word language in (8), so it is not surprising that it performs poorly here. The combined results from all five remaining criteria (excluding number of rules) are shown in Figure 3.

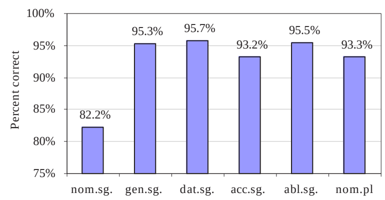
The nominative is the worst choice of base under all criteria, and thus receives the lowest rank for all metrics. This reflects the fact that the nominative suffers from more neutralizations, affecting both more words and more segments, than the oblique forms. Interestingly, the accusative also fares relatively poorly, because it is the same as the nominative for all neuter nouns, and thus shares many of the same neutralizations. Among the remaining forms, the ablative comes out slightly ahead of the genitive and the dative, because there are a few nouns (the so-called "i-stems") that have an unpredictable [i] in the ablative singular and genitive plural, and often in the nominative and accusative plural forms as well:

(14)

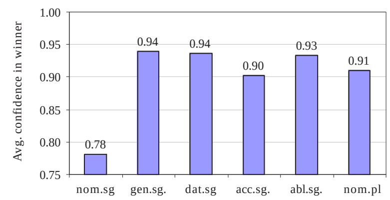
	gen.	abl.	gloss
	[ignis]	[igniː]	'fire' (masc.)
	[imbris]	[imbriː]	'rain' (masc.)
vs.	[patris]	[patre]	'father' (masc.)
	[fiːnis]	[fiːne]	'end' (masc.)

Although Allen (1903) lists a large number of i-stem nouns, the database employed here contained only 6 of them; thus, there is only a very slim advantage to choosing a base form in which the i-stems are distinct.[8]

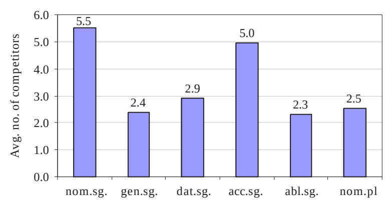
[8]In fact, the synchronic status of many i-stems seems doubtful; Allen writes that "The i-declension was confused even to the Romans themselves, nor was it stable at all periods of the language, early Latin having i-forms which afterwards disappeared." (§73). For this reason, some of the nouns listed by Allen as i-stems were not listed as such in other sources, and were included in the database without the [i].



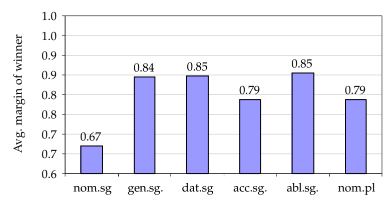
(a) Percent correct



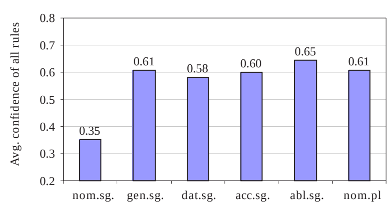
(b) Avg. confidence in winner



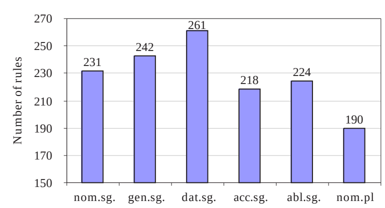
(c) Avg. no. of competitors



(d) Avg. margin between winner and nearest competitor



(e) Avg. confidence of rules



(f) No. of rules in the grammar

Figure 2: Comparison of potential bases according to the criteria in 2.2

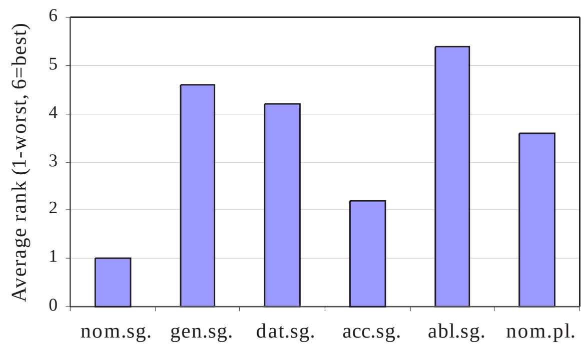


Figure 3: Average rank of forms as potential bases

An additional factor that has not been discussed here is the relative frequency of the different forms in the paradigm. As an idealization, I have assumed that learners have access to all forms of all nouns. Clearly this is not true in real life however; some cases are more frequent than others, and the frequency of cases may differ from word to word. As mentioned in section 2.1, the Albright and Hayes implementation of minimal generalization uses confidence statistics to estimate the effectiveness of rules, so that rules covering a few forms are penalized more than rules covering many forms. Therefore, with more realistic input data, including different amounts of data about different cases, subgrammars involving less frequent cases would be penalized because their rules would be based on fewer forms.

A simulation taking this into account would require more detailed frequency information about Latin noun paradigms than is currently available to me. Nevertheless, intuitively, it seems that there are substantial differences in the frequency of the oblique cases, and this would probably be the decisive factor in choosing a base from among the oblique forms that are more or less equivalent by all other criteria. For the purposes of the [honoːs] > [honor] change, it is sufficient that the model proposed here selects something other than the nominative as the base form; in the discussion that follows, I will use the genitive singular as the base for deriving the nominative, but the same result could be achieved using the ablative or dative singular.[9]

[9] A common intuition is that the [honoːs] > [honor] change may be due to the collective influence of *all* of the oblique forms combined, and not the effect of a single oblique form on the nominative singular. I will discuss this possibility further in section 4.1.

3 Projecting nominatives from the genitive

Choosing an oblique form as the base in Latin noun paradigms gives us only half of the explanation for the [honoːs] > [honor] change. In particular, it explains the "backwards" direction of the change (oblique forms affecting nominatives). This answers the question in (7a), of why it was the nominative that changed in Latin. What remains to be shown, then is that once an oblique form has been chosen as the base, the model makes the right predictions for nominative forms: namely, that polysyllabic non-neuter *-oːs* nouns changed to *-or*.

Recall that an assumption of the current model is that bases are selected early in the learning process, but learners continue to fine-tune their grammars to derive the remainder of the paradigm.[10] Therefore, in order to test the predictions of the model for nominatives using an oblique form as the base, the model was trained on the full set of 1,687 nouns in the gen.→nom. direction. The resulting grammar was then used to generate possible nominatives for all genitive forms ending with sequences that could potentially arise from rhotacism: [-oːris], [-oris], [-uris], [-eris] (157 in all). The grammar derived several possible nominatives for each noun, each with its own confidence value. For example, for [-oːris] and [-oris] genitives, the possible nominatives typically included an [-oːs] nominative, an [-or] nominative, and various other possibilities, such as [-oːris] (on the basis of words like [kanis] 'dog-nom./gen.sg.', which were identical in the nominative and genitive), [-us] (like [korpus] 'body-nom.sg.'), etc. For each noun, the best possible *-r* nominative was compared against the best possible *-s* nominative, in order to gauge the model's preference for *-r* nominatives. As expected, the preference for *-r* or *-s* varied substantially from word to word, with four distinct types of words emerging (Figure 4). (Note that in the graph, bars indicate the size of a standard deviation, not the standard error)

As Figure 4 shows, agentive nouns, which are all polysyllabic and masculine or feminine, strongly favor *-r*. In fact, these words contained *-r* etymologically, and continued to have *-r* in the nominative with no variation or hypercorrections. For other polysyllabic masculine and feminine nouns, there is a slight tendency to favor *-r*, but there is strong competition from *-s*. This is the *honor* class of words, which were etymologically *-s* but changed to *-r*, with some attested variation and occasional hypercorrections of etymological *-r* to *-s* (Neue-Wagener 1902, p. 265). For the polysyllabic neuter nouns, *-s* is favored to a moderate degree; these contained *-s* etymologically, and continued to have *-s* except in some cases of contamination

[10]The two-stage nature of this model – first selecting a base and then honing the grammar – may not be crucial. In the case of Latin, and in several other cases examined, it turns out that the asymmetries between forms seen in Figure 2 are clearest when fewer words have been learned, but do not change quantitatively when the full data set is considered.

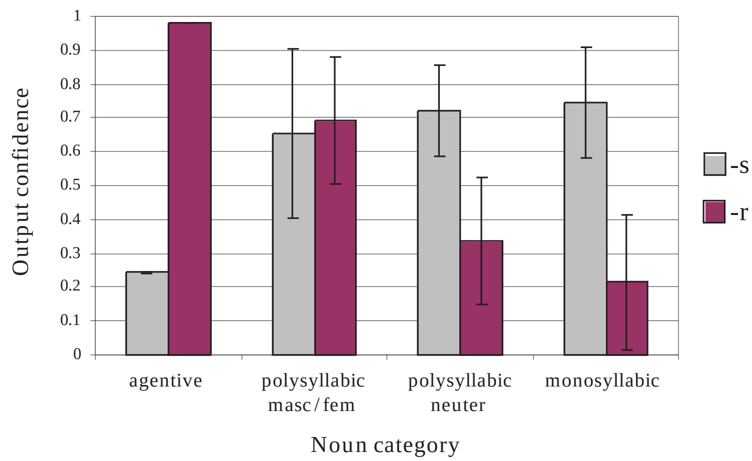


Figure 4: Preference for *-r* or *-s* in four distinct sets of words

from masculine doublets, and a few isolated other examples. Finally, monosyllables, which also contained *-s* etymologically, strongly favor *-s*; in fact, these did retain *-s* with no variation.

Why does this pattern emerge? The differing strength of *-r* and *-s* for different words is due to the fact that the system employs multiple rules, with different confidence values in different contexts. Among polysyllabic non-neuter nouns, genitives in [-oːris] frequently have nominatives in *-or*. Thus, the rule of [oːris] $\rightarrow$ [or] / $[X]_{polysyl,-neut}$___# has a relatively high confidence (.727), correctly deriving words like [soror] and [cruor], and all agentives, but failing for words like [honoːs]. Among these forms, then, there is a slight preference for *-or* in the nominative. It is also possible, however, to write a more specific rule that covers just the agentives, since these are not only polysyllabic and non-neuter, but they also all have a stem-final [s] or [t]: *doctor*, *audītor*, *cēnsor*, etc.[11] Thus, the more specific rule [oːris] $\rightarrow$ [or] / $[X\ \{s,t\}]_{polysyl,-neut}$___# is able to describe the agentives quite narrowly, and has an extremely high reliability (.980).

Outside the class of polysyllabic non-neuter nouns, the reliability of *-r* is much lower. Among polysyllabic neuters, very few nominatives end in *-r*, so the general rule [ris] $\rightarrow$ [r] / $[X]_{polysyl,+neut}$___# has a rather low confidence (.196). There are, however, a few local pockets of *-r* nominatives among the neuters, especially among those with *-aris* and *-eris* in the genitive ([kalkar]/[kalkaːris]

[11] This common phonological trait is not a coincidence: agentives were formed from the perfect passive participle (4th stem), which was generally formed by adding a [t], or in some phonological contexts, by changing a stem consonant to [s] (e.g., [kad- $\rightarrow$ [kaːs- 'fall') .

'spur-nom./gen.', [nektar]/[nektaris] 'nectar-nom./gen.', [aker]/[akeris] 'maple-nom./gen.') This is the reason why the model disfavors innovative *-r* nominatives among neuters, and why there is also a good deal of item-by-item variation among them.[12] Finally, among monosyllables, almost all *-ris* genitives had nominatives ending in *-s* – a notable exception being [fuːr] 'thief'. Thus, the model correctly learns that in this environment, the $s \sim r$ alternation is extremely robust, and final *-r* cannot compete with it.

It should be reiterated that the slight preference of the model for *-r* in words like [honor] emerges in spite of the fact that they were listed with *-s* in the training data. In other words, the grammar produces an output which is different from the existing form. Therefore, under this analysis, pre-change forms like [honoːs] would have been considered irregular, and would have had to have been listed as exceptions to the *-ris* $\sim$ *-r* pattern. Of course, if learners had perfect memories and access to all forms of all words, then they could perfectly well have memorized [honoːs] and continued to produce it, and the language would not have changed. However, in real life this is not the case, and speakers must sometimes synthesize new forms. The model is intended to predict what forms a speaker would produce in such situations, and in this case it correctly predicts errors, or overregularizations, like *honor*. It cannot, however, predict when existing forms would be unavailable, forcing speakers to use their grammars. What is missing from this model, then, is a production mechanism which uses both the lexicon and grammar to produce forms. Even so, the result in (4) is still significant, because we can assume that speakers do sometimes make overregularization errors (Marcus, Pinker, Ullman, Hollander, Rosen, and Xu 1992; Pinker 1999), and the errors that the model makes correctly mirror the attested historical change.

It is worth mentioning that the [honoːs] > [honor] change was just one of many changes that affected nominatives in the history of Latin. Numerous nouns with highly irregular nominative forms were regularized; for example, the nominative of [juppiter]/[jowis] 'Jupiter-nom./gen.' was eventually replaced by [jowis] (cf: [kanis]/[kanis] 'dog-nom./gen.'), and the nominative of [boːs]/[bowis] 'cow-nom./gen.' is attested as [bowis] (Kieckers 1960, vol. 2, §I.21.3; Kühner §63.2). The converse change, of regularizing nouns by fixing the oblique forms, gener-

[12]In fact, Kieckers (1960) points to one, possibly quite isolated example of a neuter noun with etymological *-r* being written with an *-s*: *femus* 'femur' (vol 2, §22).

ally did not occur (*[juppitris] or *[juppitriː], *[boːris], etc.).[13] Furthermore, the form of most nouns in modern Romance languages can be traced back to oblique forms in Latin. For example, Latin [peːs]/[pedis] 'foot-nom./gen.' has yielded Italian [pje**d**e], instead of the expected *[pe] (cf: Latin [treːs] > Italian [tre] 'three'); similarly [ars]/[artis] 'skill' > Ital. [ar**t**e], [floːs]/[floːris] 'flower' > Ital. [fjo**r**e], [niks]/[nivis] 'snow' > Ital. [ne**v**e], and so on. A full analysis of all of these changes is beyond the scope of this paper, but the computational model presented here does prefer many of the attested changes as regularizations in these cases as well.

4 Discussion and conclusion

This analysis captures two common intuitions about Latin nouns, and about the [honoːs] to [honor] change. The first is that oblique forms are "more revealing" about the declension of a noun than the nominative – seen, for example, in the common practice of listing both the nominative and genitive forms in dictionary entries, as the nominative alone is not considered "informative enough" to predict the entire paradigm.[14] This intuition is reflected in the current analysis by the fact that an oblique form is chosen as the base, and the remainder of the paradigm is derived from an oblique form.

The second intuition is that the change from [honoːs] to [honor] involved replacing a small, irregular, morphologized alternation with the more general default pattern of non-alternation. This analysis shares with Barr (1994) the idea that this can be captured with competing rules, but differs with respect to why different classes of words were treated differently. In Barr's system, a single small $X\ s \sim X\ ris$ rule competed with a single large $X \sim X\ is$ rule; the fact that $X\ s \sim X\ ris$ lost among polysyllabic non-neuters and was retained in other cases is attributed to extra-grammatical factors, such as salience and "degree of allomorphy," which

[13] There seems to be at least one set of cases in which a property of the nominative was extended to the remainder of the paradigm: the paradigm of words like [voːks] 'voice' originally had a long [oː] and [k] in the nominative, and a short [o] and [kʷ] elsewhere ([voːks], [vokʷis], . . . ; see Meiser 1998, p. 141 regarding vowel length, Leumann 1977, p. 148 and Kieckers 1931, p. II.13 regarding [kʷ]). The long [oː] and simple [k] of the nominative were subsequently extended to all forms: [voːks], [voːkis], ... It would be interesting to compare the relative chronology of these changes, since these nominative-driven changes may have occurred at an older stage of the language in which nominatives suffered from fewer phonological reductions, while oblique-driven changes occurred throughout the Classical and Late Latin period.

[14] The relative uninformativeness of the nominative in Latin is due, in part, to the fact that the nominative suffix for one large class of nouns lacked a vowel (-*s*), creating coda clusters that resulted in phonological simplifications (e.g., **arts* > *ars*). The oblique forms always provided a prevocalic context for the stem, resulting in far fewer neutralizations; rhotacism is a rare exception.

made the rule harder to learn or apply in different contexts. In the current system, on the other hand, the difference between different genders and word lengths is attributed to the existence of multiple versions of the rules in question, at varying levels of generality, and with differing reliability in different contexts. The use of multiple overlapping rules might be seen as unwanted redundancy in the model, but in fact cases like Latin are taken as evidence that speakers, too, have detailed knowledge about the reliability of different processes in different environments. Furthermore, an ability to assess the reliability of rules in different environments is required in any event in order for learners to locate the best rules to describe the patterns of their language.

Although this analysis makes use of several intuitions about the factors that are thought to drive paradigm leveling, it ignores certain other factors that have been proposed in the literature. Some notable factors that do not play a role in this analysis are the frequency of an allomorph within the paradigm, the token frequency of various surface forms, or the semantic naturalness of different nouns in different cases. It is useful to consider, therefore, the extent to which these other factors could provide an alternative explanation of the [honor] analogy, and whether the current model would benefit from incorporating either of these factors.

4.1 Frequency of occurrence within the paradigm

It is often suggested that the [honoːs] to [honor] change was encouraged by the fact that every form in the paradigm except the nominative singular contained [r]. I will refer to this as the "majority rule" hypothesis. Under the model proposed here, each paradigm has a single unique base, and forms are derived by grammars relating individual pairs of forms. If the most informative form had turned out to be the nominative singular, the prediction of this model is that all of the remaining forms could have been rebuilt on the basis of a single form. Thus, this model has no way to capture the majority rule intuition. It is not clear to me, however, that there is evidence that paradigm leveling is truly driven by majority rule. There are numerous of cases in which a single form seems to have driven a paradigmatic change (see, for example, Albright 2002 for a case from Yiddish), whereas it is difficult to prove that a leveling like [honoːs] > [honor] would not have happened if [s] had occurred in more slots in the paradigm. Furthermore, the majority rule hypothesis cannot explain several aspects of the Latin facts. First, forms with [r] outnumbered forms with [s] not only in the polysyllabic non-neuter nouns, but in monosyllabic and neuter nouns as well. From the point of view of paradigm-internal pressures, there is no reason that these two classes of words should have behaved differently. In addition, there are many other noun paradigms in which the nominative had a different form from the rest of the paradigm, but was not leveled (e.g., [iter] ∼ [itineris] 'road-nom./gen.sg.'). Therefore, frequency of occurrence within the paradigm does not seem to add anything to the account of the change.

4.2 Token frequency of different paradigm members

A natural hypothesis, pursued by Mańczak (1958) and others, is that less frequent forms are often rebuilt on the basis of more frequent forms within the paradigm. Could it be the case that the nominative was actually a lot less frequent than the oblique forms in Latin? This is especially relevant in Latin because many or most of the words affected by the [honor] analogy were inanimate or abstract nouns, which are perhaps more frequent in oblique forms than in the nominative.

In order to get a rough (and very informal) estimate of the relative frequency of case forms for different nouns, I performed some counts on the complete works of Cicero, as found in the Perseus Digital Library (http://www.perseus.tufts.edu/). As (15) shows, it is true that among singular forms, nominatives rarely constitute the majority of tokens for any noun. This may help to explain why nominative forms were open to rebuilding in Latin – they were, on the whole, not so frequent that they could always be reliably memorized and retrieved.

(15) Distribution of singular tokens for some Latin nouns

Noun	Total Sg.	Nom.	Gen.	Acc.	Abl.
Polysyllabic, non-neuter					
hono:s/*honor* 'honor'	285	19%	25%	18%	32%
labo:s/*labor* 'work'	163	17%	21%	37%	35%
odo:s/*odor* 'odor'	4	50%	0%	0%	50%
Monosyllabic					
flo:s 'flower'	16	25%	0%	56%	19%
mo:s 'custom'	146	18%	2%	17%	63%
o:s 'mouth'	65	29%	9%	6%	55%
Neuter[15]					
corpus 'body'	174	21%	47%	–	28%
onus 'burden'	40	45%	40%	–	15%
tempus 'time'	935	32%	15%	–	51%
Masculine, agentive					
re:x 'king'	207	23%	21%	27%	18%
homo: 'man'	1049	19%	23%	35%	12%
sena:tor 'senator'	43	33%	23%	28%	14%

[15] A problem arises in counting frequencies for neuters, since the nom. and acc. forms are identical. The hypothesis being tested here is that the frequency of [s] forms in the paradigm determines their susceptibility to leveling, so I have counted all *s* forms in the nom. column, to facilitate comparison with the masc. and fem. nouns.

What these counts cannot explain, however, is why the change should have been restricted only to the non-neuter polysyllabic nouns. The nominative does not seem to be less frequent in this class of nouns that in any other class. Furthermore, there is apparently not even a difference between masculine agentive nouns like 'king', 'man' and 'senator', and inanimate, abstract nouns like 'honor' or 'custom'. Thus, a frequency-based account can explain only the direction, but not the details of the [honoːs] > [honor] change.

4.3 Semantics and local markedness

Another intuition, related to token frequency, but logically distinct from it, is that the semantics of particular lexical items make them more "natural" in some case forms than in others. Tiersma (1982), for example, shows that singular forms in Frisian have been rebuilt on the basis of plural forms, but just for those nouns which occur more naturally in the plural than in the singular (such as 'teeth' or 'geese'). He refers to this phenomenon as *local markedness*. On the whole, we would expect local markedness to be reflected in token frequency, which is much easier to measure (see above). I have no estimate of the naturalness of the nominative forms which changed from [-oːs] to [-or], but I see no reason why this would fare any better than token frequency as an explanation of the change. Nouns like [onus] 'burden', [korpus] 'body', and [floːs] 'flower' seem to me to be just as "non-agentive" as [honoːs] 'duty', [odoːs] 'odor', or [arboːs] 'tree'. It appears that the class of nouns that changed is best defined by prosodic and morphological properties, and adding a sensitivity to frequency or semantics would not improve the model's predictions in this case.

4.4 Leveling vs. extending alternations

The analysis of paradigm leveling proposed here relies on a strong pre-existing pattern of non-alternation in the lexicon – in this case, the non-alternation of [r]. This proposal immediately raises two related questions: first, if paradigm uniformity is really just the extension of an existing pattern of non-alternation, then what happens when the dominant pattern is alternation? Why does there seem to be a universal tendency towards leveling?

As an example of a language with a dominant pattern of alternation, consider a previous stage of Korean (Martin 1992):

(16)

/___#	Example	/___V	(ACC -*iɭ*)	gloss
[t̚]	[nat̚]	[d]	[nadɨɭ]	'grain'
[t̚]	[pat̚]	[tʰ]	[patʰɨɭ]	'field'
[t̚]	[tʃʌt̚]	[dʒ]	[tʃʌdʒɨɭ]	'milk'
[t̚]	[kot̚]	[tʃʰ]	[kotʃʰɨɭ]	'flower'
[t̚]	[ot̚]	[s]	[osɨɭ]	'clothing'

As (16) shows, all stem-final coronal obstruents alternate with [t̚] word-finally. As with Latin, this alternation could be expressed as the result of a markedness constraint against manner and laryngeal specifications in coda position (favoring [t̚]) outranking faithfulness constraints (which preserve underlying contrasts intervocalically). If there was a universal pressure for uniform exponence constraints to move above IO-Faithfulness constraints, then we would expect that paradigmatic changes in Korean should bring Korean closer to non-alternating paradigms, perhaps as in (17). (The intervocalic voicing of /t/→[d] is a completely predictable process in Korean.) Note that although the phonotactics of Korean rule out a completely non-alternating paradigm ([nat̚] ∼ *[natɨɭ], or *[nad] ∼ [nadɨɭ]), we may assume that the relatively minor, predictable allophonic alternation between [t̚] and [d] better satisfies Uniform Exponence than a [t̚] ∼ [s] or [t̚] ∼ [tʃʰ] alternation, just as the shortening of final /oːr/→[or] in Latin is assumed to be a less serious violation of Uniform Exponence than a [s] ∼ [r] alternation is.

(17) Expected Korean paradigm leveling:

/___#		/___V		
[t̚]	[nat̚]	[d]	[nadɨɭ]	'grain'
[t̚]	[pat̚]	[d]	[padɨɭ]	'field'
[t̚]	[tʃʌt̚]	[d]	[tʃʌdɨɭ]	'milk'
[t̚]	[kot̚]	[d]	[kodɨɭ]	'flower'
[t̚]	[ot̚]	[d]	[odɨɭ]	'clothing'

In fact, the attested change in Korean noun paradigms is quite different. As it turns out, the majority of coronal obstruent-final stems contained [s] or [tʃʰ] etymologically (i.e., most were like [kot̚]/[kotʃʰɨɭ] or [ot̚]/[osɨɭ]), and many Korean noun paradigms are being rebuilt to contain [s] or [tʃʰ]:

(18) Actual change in Korean Paradigms: (Martin 1992; Hayes 1995)

/__#		/__V		
[t˺]	[nat˺]	[s]	[nasɨl̢], ??[nadɨl̢]	'grain'
[t˺]	[pat˺]	[tʰ],[tʃʰ],[s]	[patʰɨl̢], [patʃʰɨl̢], [pasɨl̢]	'field'
[t˺]	[tʃʌt˺]	[s],[dʒ]	[tʃʌsɨl̢], [tʃʌdʒɨl̢]	'milk'
[t˺]	[kot˺]	[tʃʰ],[s]	[kotʃʰɨl̢], [kosɨl̢]	'flower'
[t˺]	[ot˺]	[s]	[osɨl̢]	'clothing'

Although there is a considerable amount of word-by-word and speaker-by-speaker variation, it is clear that the restructuring underway in Korean is introducing, not eliminating alternations. For the most part, the dominant alternations of [t˺] ∼ [s] and [t˺] ∼ [tʃʰ] are coming to replace other, arguably less drastic alternations like [t] ∼ [d].

The explanatory challenge, therefore, is to explain why in some cases a pattern of alternation is extended (as in Korean), while in other cases, alternations are eliminated (as in Latin). The model of paradigm learning advocated in this paper always extends the strongest pattern, regardless of whether it is alternating or uniform. The reranking of paradigm uniformity constraints, on the other hand, can explain only leveling; the spread of alternations would have to be handled by other means, such as anti-correspondence constraints (Hayes 1999), leaving us with no explanation for why sometimes paradigm uniformity wins out, and sometimes anti-correspondence wins out.

The contrast between Latin and Korean may also help to shed light on why there seems to be a typological bias towards paradigm leveling; in Latin, rhotacism affects just one segment (/s/), so the only words which exhibit the alternation are those with stem-final [s]. In Korean, on the other hand, coda neutralizations affect all obstruents, so almost all lexical items exhibit alternations. I conjecture that morphophonemic alternations typically affect only a smallish subset of the phonemic inventory; rhotacism affects just /s/, umlaut affects only back vowels, palatalization tends to affect just coronals (and often just coronal stridents), and so on. The end result is that alternations tend to be the minority pattern, and there will be a tendency to generalize non-alternation, or level.

4.5 Conclusion

In this paper, I have argued that the Latin [hoнoːs] > [honor] change was caused by more than simply a sporadic pressure for paradigm uniformity or uniform exponence constraints to assert themselves over IO-Faithfulness constraints. I have shown that the spread of [r] to nominative forms did more than just create uniform paradigms; it also extended a pattern of non-alternation that was already dominant in the lexicon. Details of the change, such as its restriction to polysyllabic nouns and non-neuters reflect the fact that these were especially strong contexts for [r] stems. Furthermore, the “backwards” direction of the leveling, with oblique forms influencing the nominative singular, can be explained by a particular model of Latin noun paradigms, in which an oblique form served as the base, and nominative forms were derived from oblique forms by rules operating on surface forms.

More generally, this result provides evidence for a model of paradigm learning in which learners choose the base form that is “the most informative” – i.e., that preserves the most distinctions between classes of words, and allows the remainder of the paradigm to be predicted with the greatest accuracy and confidence. This echoes a proposal by Lahiri and Dresher (1984) that certain forms in the paradigm “matter more than others” to learners when they are determining what class a word belongs to. The remainder of the paradigm is then derived using a grammar of gradient morphological rules, providing a measure of the relative strength of different patterns, and allowing different patterns to have different productivity in various contexts. The prediction of this model is that distinctions that are preserved in the base form will be easily learned and maintained, whereas distinctions that are neutralized in the base form may be lost by leveling or regularization.

A Metrics for base selection

The table below lists all of the effectiveness measures of each of the six candidates for base status, based on the 494 most frequent Latin nouns. Rows indicate the input forms, and columns indicate the output forms; for example, the average winner confidence for the *nom.→gen.* mapping is 0.76, in the upper left.

	↓In/Out→	nom.sg.	gen.sg.	dat.sg.	acc.sg.	abl.sg.	nom.pl.	**sum**
Avg winner conf.:	nom.sg.		0.76	0.70	0.86	0.82	0.76	**0.78**
Avg margin:			0.66	0.59	0.77	0.68	0.65	**0.67**
Avg competitors:			5.68	6.44	3.57	5.76	6.15	**5.52**
Percent correct:			0.82	0.77	0.87	0.85	0.80	**0.82**
No. of constraints:			2341	2618	1987	2340	2308	**2319**
Avg grammar conf.:			0.35	0.31	0.45	0.33	0.33	**0.35**
Avg winner conf.:	gen.sg.	0.88		0.97	0.94	0.96	0.95	**0.94**
Avg margin:		0.69		0.92	0.91	0.91	0.79	**0.84**
Avg competitors:		3.16		1.92	1.48	2.89	2.38	**2.37**
Percent correct:		0.93		0.97	0.96	0.96	0.94	**0.95**
No. of constraints:		1877		2781	2253	2609	2618	**2428**
Avg grammar conf.:		0.45		0.70	0.67	0.69	0.53	**0.61**
Avg winner conf.:	dat.sg.	0.86	0.97		0.94	0.96	0.95	**0.94**
Avg margin:		0.69	0.93		0.89	0.91	0.81	**0.85**
Avg competitors:		3.68	1.83		2.99	2.69	3.25	**2.89**
Percent correct:		0.92	0.98		0.96	0.97	0.96	**0.96**
No. of constraints:		1920	2418		2538	2805	3377	**2612**
Avg grammar conf.:		0.43	0.81		0.59	0.65	0.43	**0.58**
Avg winner conf.:	acc.sg.	0.89	0.91	0.90		0.91	0.90	**0.90**
Avg margin:		0.76	0.82	0.80		0.79	0.77	**0.79**
Avg competitors:		3.29	5.40	7.01		3.79	5.39	**4.98**
Percent correct:		0.95	0.94	0.92		0.93	0.92	**0.93**
No. of constraints:		1504	2045	2460		2467	2465	**2188**
Avg grammar conf.:		0.55	0.73	0.60		0.60	0.53	**0.60**
Avg winner conf.:	abl.sg.	0.86	0.96	0.96	0.95		0.94	**0.93**
Avg margin:		0.68	0.93	0.90	0.91		0.85	**0.85**
Avg competitors:		2.76	3.36	1.65	1.29		2.45	**2.30**
Percent correct:		0.92	0.97	0.97	0.97		0.95	**0.95**
No. of constraints:		1829	2377	2514	2087		2398	**2241**
Avg grammar conf.:		0.45	0.78	0.72	0.72		0.56	**0.65**
Avg winner conf.:	nom.pl	0.84	0.94	0.92	0.92	0.92		**0.91**
Avg margin:		0.69	0.76	0.81	0.86	0.81		**0.79**
Avg competitors:		2.61	1.77	3.24	2.11	2.95		**2.54**
Percent correct:		0.89	0.94	0.94	0.94	0.94		**0.93**
No. of constraints:		1424	1855	2260	1955	2010		**1901**
Avg grammar conf.:		0.50	0.69	0.56	0.66	0.62		**0.61**

References

Albright, A. (2002). Base selection in analogical change: A German/Yiddish comparison. Talk presented at the 28th Annual Meeting of the Berkeley Linguistics Society, Feb. 15-18.

Albright, A. and B. Hayes (1999). An Automated Learner for Phonology and Morphology. Unpublished UCLA ms, http://www.linguistics.ucla.edu/people/hayes/learning/learning.htm.

Albright, A. and B. Hayes (2000). Distributional encroachment and its consequences for morphological learning. In A. Albright and T. Cho (Eds.), *UCLA Working Papers in Linguistics 4 (Papers in Phonology 4)*, pp. 179–190.

Allen, J. H. (1903). *Allen and Greenough's New Latin Grammar for Schools and Colleges*. Boston: Ginn and Company.

Barr, R. (1994). *A Lexical Model of Morphological Change*. Ph. D. thesis, Harvard University.

Boersma, P. and B. Hayes (2001). Empirical Tests of the Gradual Learning Algorithm. *Linguistic Inquiry 32*(1), 45–86.

Bybee, J. (1985). *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins Publishing Company.

Delatte, L., E. Evrard, S. Govaerts, and J. Denooz (1981). *Dictionnaire fréquentiel et Index inverse de la langue latine*. Université de Liège: Laboratoire d'analyse statistique des langues anciennes.

Friederici, A. D. and J. E. Wessels (1993). Phonotactic knowledge of word boundaries and its use in infant speech perception. *Perception and Psychophysics 54*, 287–295.

Hayes, B. (1995). On what to teach the undergraduates: Some changing orthodoxies in phonological theory. *Linguistics in the Morning Calm 3*, 59–77.

Hayes, B. (1999). Phonological Restructuring inYidiny and its Theoretical Consequences. In B. Hermans and M. Oostendorp (Eds.), *The Derivational Residue in Phonological Optimality Theory*, pp. 175–205. Amsterdam: John Benjamins.

Hayes, B. (to appear). Phonological acquisition in Optimality Theory: The early stages. In R. Kager, J. Pater, and W. Zonneveld (Eds.), *Fixing Priorities: Constraints in Phonological Acquisition*. Cambridge University Press.

Hock, H. H. (1991). *Principles of Historical Linguistics* (2nd ed.). The Hague: Mouton de Gruyter.

Jusczyk, P. W., A. Friderici, J. M. Wessels, V. Y. Svenkerud, and A. M. Jusczyk (1993). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language 32*, 402–420.

Jusczyk, P. W., P. A. Luce, and J. Charles-Luce (1994). Infants sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language 33*, 630–645.

Kenstowicz, M. (1995). Base-identity and uniform exponence: Alternatives to cyclicity. In J. Durand and B. Laks (Eds.), *Current Trends in Phonology: Models and Methods*. CNRS and University of Salford Publications. Available for download from the Rutgers Optimality Archive, http://roa.rutgers.edu.

Kiparsky, P. (1997). Covert generalization. *Stanford Univ. ms*.

Klausenburger, J. (1979). *Morphologization: Studies in Latin and Romance morphophonology*. Tübingen: Max Niemeyer.

Kühner, R. (1912). *Ausführliche Grammatik der Lateinischen Sprache* (2nd ed.). Hannover: Verlag Hahnsche Buchhandlung.

Kuryłowicz, J. (1947). The nature of the so-called analogical processes. *Diachronica 12*(1), 113–145.

Lahiri, A. and B. E. Dresher (1984). Diachronic and synchronic implications of declension shifts. *The Linguistic Review 3*, 141–163.

Leumann, M. (1977). *Lateinische Laut- und Formenlehre*. Munich: C. H. Beck'sche Verlagsbuchhandlung (Oscar Beck).

Mańczak, W. (1958). Tendences générales des changements analogiques. *Lingua 7*, 298–325 and 387–420.

Marcus, G., S. Pinker, M. Ullman, M. Hollander, T. Rosen, and F. Xu (1992). Overregularization in language acquisition. *Monographs of the Society for Research in Child Development*.

Martin, S. E. (1992). *A Reference Grammar of Korean*. Tokyo, Japan: Charles E. Tuttle Company.

McCarthy, J. (1998). Morpheme structure constraints and paradigm occultation. In M. C. Gruber, D. Higgins, K. Olson, and T. Wysocki (Eds.), *CLS 32, vol. II: The Panels*, Chicago. Chicago Linguistic Society.

Meiser, G. (1998). *Historische Laut- und Formenlehre der lateinischen Sprache*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Mikheev, A. (1997). Automatic Rule Induction for Unknown-Word Guessing. *Computational Linguistics 23*(3), 405–423.

Neue, C. F. and C. Wagener (1902). *Formenlehre der lateinischen Sprache*, (3rd ed.), Volume 1 (Das Substantivum). Leipzig: O.R. Reisland.

Pinker, S. (1999). *Words and Rules: The Ingredients of Language*. New York: Basic Books.

Pinker, S. and A. Prince (1988). On language and connectionism: Analysis of a Parallel Distributed Processing model of language acquisition. *Cognition 28*, 73–193.

Skutsch, O. (1985). *The Annals of Q. Ennius*. Oxford: Clarendon Press.

Smolensky, P., L. Davidson, and P. Jusczyk (to appear). Optimality in Language Acquisition I: Phonology. In P. Smolensky and G. Legendre (Eds.), *Architecture of the Mind/Brain: Neural computation, optimality, and universal grammar in Cognitive Science*, Chapter 16.

Steriade, D. (1994). Lexical conservatism. In *Linguistics in the Morning Calm, Selected Papers from SICOL 1997, Linguistic Society of Korea*. Hanshin Publishing House.

Steriade, D. (2000). Paradigm Uniformity and the Phonetics/Phonology boundary. In J. Pierrehumbert and M. Broe (Eds.), *Papers in Laboratory Phonology, vol. 6*. Cambridge University Press.

Tiersma, P. (1982). Local and general markedness. *Language 58*(4), 832–849.

Wetzels, L. (1984). Paradigm leveing in Latin and Old French: A critical view of rule inversion. In P. Baldi (Ed.), *Papers from the XXth Linguistic Symposium on Romance Languages*, Volume 26 of *Amsterdam Studies in the Theory and History of Linguistic Science*, Amsterdam. John Benjamins Publishing Company.

Adam Albright
Dept. of Linguistics, UCLA
Los Angeles, CA 90095-1543
Email: aalbrigh@ucla.edu