2005

Taking a free ride in morphophonemic learning

John J. McCarthy
University of Massachusetts, Amherst, jmccarthy@linguist.umass.edu

Taking a Free Ride in Morphophonemic Learning*

John J. McCarthy
University of Massachusetts at Amherst. Department of Linguistics
Amherst, MA 01003 USA
jmccarthy@linguist.umass.edu

Come on and take a free ride (free ride)
(Winter 1972)

Abstract

As language learners begin to analyze morphologically complex words, they face the problem of projecting underlying representations from the morphophonemic alternations that they observe. Research on learnability in Optimality Theory has started to address this problem, and this article deals with one aspect of it. When alternation data tell the learner that some surface [B]s are derived from underlying /A/s, the learner will under certain conditions generalize by deriving *all* [B]s, even nonalternating ones, from /A/s. An adequate learning theory must therefore incorporate a procedure that allows nonalternating [B]s to take a «free ride» on the /A/ → [B] unfaithful map.

Key words: chain shift, learning, morphophonemics, opacity, Optimality Theory; Arabic, Choctaw, German, Japanese, Rotuman, Sanskrit.

Table of Contents

* Special thanks to participants in the UMass summer 2004 phonetics/phonology group (SPe/oG) for their comments and suggestions. For additional comments, I am very grateful to Shigeto Kawahara, John Kingston, Joe Pater, Alan Prince, Bruce Tesar, and an anonymous reviewer.

1. Introduction

Language learning is a central problem of linguistic research, and it is important to show that learners can in principle acquire the grammar of the ambient language from the limited data available to them. Formal studies of language learnability are concerned with determining, for a particular linguistic theory, whether and under what conditions learning can succeed. Ideally, any proposed linguistic theory will be accompanied by an explicit learning model from which learnability results can be deduced.

Optimality Theory (Prince and Smolensky 2004) has been coupled with a learning model almost from the beginning. There is by now a very substantial body of work discussing the constraint demotion learning algorithm of Tesar and Smolensky (1994), extensions, and alternatives. Although much of this literature deals with phonotactic learning, it has begun to address morphophonemics: how do learners use data from phonological alternations to work out a grammar and a lexicon of underlying representations? (See, e.g., Dinnsen and McGarrity 2004; Ota 2004; Pater 2004; Tesar *et al.* 2003; Tesar and Prince forthcoming; Tesar and Smolensky 2000: 77-84; Tessier 2004.) Morphophonemic learning presents problems that are not met with in phonotactic learning. This article describes one such problem and offers a solution to it.

The OT learnability literature standardly assumes that phonotactic learners posit identity maps from the lexicon to observed surface forms:[1] the learner hears [A], takes /A/ to be its underlying source, and seeks a grammar that will perform the identity map /A/ → [A]. If the learner never hears *[B], however, and this gap is nonaccidental, then his (or her) grammar must not map anything to *[B]. In particular, the /B/ → [B] identity map must not be sanctioned by the grammar. Instead, the grammar must treat a hypothetical /B/ input unfaithfully, mapping it to zero or to something that *is* possible in the target language.

Morphophonemic learning is different. As learners begin to analyze morphologically complex words, they discover morphophonemic alternations for which the identity map is insufficient.[2] For example, once a learner becomes aware of paradigmatic relationships like Egyptian Arabic *fihim/fihmu* 'he/they understood' and sets up underlying /fihim/, he is then committed to the unfaithful map /fihim-u/ → *fihmu*. But what happens with nonalternating forms? Is the identity map still favored?

In this article, I argue that nonalternating forms are sometimes derived by unfaithful maps as well. As I will show, there are examples with the following properties. Data from alternations indicate that *some* surface [B]s derive from underlying /A/s. Other evidence, some of it distributional and some involving opacity, argues that *all* surface [B]s, even the nonalternating ones, derive from underlying /A/s. I will propose a learning principle according to which learners who have discovered the /A/ → [B] unfaithful map from alternations will attempt to generalize it, projecting /A/ inputs for all surface [B]s, whether they alternate or not. In other

1. I will not attempt to list the many works that assume the identity map in phonotactic learning, nor will I attempt to identify the originator of this idea.
2. Alderete and Tesar (2002) argue that the identity map is insufficient even in phonotactic learning.

words, the nonalternating [B]s attempt to take a «free ride» on the independently motivated /A/ → [B] map.[3] The learner accepts this tentative hypothesis about underlying representations if it yields a grammar that is consistent (in the sense of Tesar 1997) and that is more restrictive (in the sense of Prince and Tesar 2004) than the grammar would be without this hypothesis. These two requirements, consistency and greater restrictiveness, are arguably sufficient to rule out inappropriate generalization of unfaithful maps.[4]

Section 2 and the appendix of this paper document the phenomenon: there are languages where alternations show that *some* [B]s derive from /A/s and further evidence shows that *all* [B]s must derive from /A/s. Section 3 presents and applies the proposed learning algorithm. Section 4 describes a limitation of this algorithm and proposes a modification to address this limitation.

2. Exemplification and explanation of the issue

This section describes a single example, coalescence in Sanskrit. To establish the generality of the problem, however, this article also includes an appendix with several more examples. The examples in the appendix are somewhat more complex than Sanskrit, but they also establish that the scope of the learning problem goes well beyond the details of the Sanskrit analysis.

In Sanskrit, alternations like those in (1) show that some surface long mid vowels [eː] and [oː] are derived by coalescence from /ai/ and /au/ sequences, respectively.

(1) Sanskrit coalescence (de Haas 1988; Gnanadesikan 1997; Schane 1987; Whitney 1889)

/tava indra/	taveːndra	‘for you, Indra (voc.)’
/hita upadaiʃaḥ/	hitoːpadeːʃaḥ	‘friendly advice’

Coalescence occurs in both internal and external sandhi, and it is abundantly supported by alternations. The result of coalescence is a compromise in the height of the two input vowels. The result is long because the moras associated with the input vowels are conserved in the output. These are typical properties of coalescence cross-linguistically (de Haas 1988).

Since all surface mid vowels are long in Sanskrit, all are at least potentially the result of coalescence. For many words with mid vowels, there are alternations that support a coalescent source. But it is unlikely that relevant alternation data will exist for all mid vowels, especially in the limited experience of the language learner. When presented with an [eː] and no relevant alternations, what does the learner do?

3. I have borrowed the term «free ride» from Zwicky (1970). The situation Zwicky has in mind is quite different from what I am talking about: an analyst, instead of positing a rule /A/ → [C], posits a rule /A/ → [B] that is ordered before an independently motivated rule /B/ → [C], letting the output of one rule take a free ride on the other.
4. Ricardo Bermúdez-Otero has drawn my attention to Bermúdez-Otero (2004), which discusses a similar problem from the perspective of stratal OT.

(To simplify the exposition, I will only talk about [eː] until late in section 3, when I will look at [oː] as well.) A persistent bias toward the identity map would favor positing underlying /eː/ for nonalternating [eː]. There is a good reason, however, to think that learners actually adopt a free-ride strategy, deriving nonalternating [eː] from /ai/. The reason: by deriving all [eː]s from /ai/s, we explain why Sanskrit has no short [e]s. If all mid vowels are derived by coalescence and if the product of vowel coalescence is always a long vowel, then there is no source for short [e].

The traditional approach to cases like this involves positing a morpheme structure constraint (MSC) that prohibits underlying mid vowels, short or long. Classic MSC's are not durable; they cannot restrict the output of rules. The rule of coalescence is therefore free to create (long) mid vowels. Short mid vowels are not observed in surface forms because they are barred from the lexicon by the MSC and no rule creates them. On this view, even nonalternating [eː] has to be derived by coalescence because underlying /eː/ is ruled out by the MSC.[5] This type of coalescence is said to be non-structure-preserving (de Haas 1988), since its output is something that is not allowed in the lexicon.

In OT, richness of the base (ROTB) forbids MSCs, but an abstractly similar analysis is possible. Gnanadesikan (1997: 139-153) analyzes Sanskrit with a chain shift: underlying /ai/ maps to [eː], while underlying /e(ː)/ maps to [i(ː)] or [a(ː)] —it does not matter which. Gnanadesikan assumes a constraint set similar to (2).

(2) a. IDENT(Vowel Height) (hereafter ID(VH))
For every pair (V, V′), where V is an input vowel and V′ is its output correspondent, if V and V′ differ in height, assign a violation mark.

b. IDENT-ADJ(Vowel Height) (hereafter ID-ADJ(VH))
For every pair (V, V′), where V is an input vowel and V′ is its output correspondent, if V and V′ differ by more than one degree of height, assign a violation mark.

c. *MID
Mid vowels are prohibited.

d. *DIPH
Diphthongs are prohibited.

e. UNIFORMITY (hereafter UNIF)
No output segment has multiple correspondents in the input (≈no coalescence).

For present purposes, I will assume that this is the entire constraint set and that the only unfaithful maps that are possible are coalescence, which violates UNIF, and changes in vowel height, which violate the ID constraints.

5. The Alternation Condition of Kiparsky (1973) was specifically formulated so as not to exclude such an analysis. The Alternation Condition says, approximately, that no neutralization rule can apply to all instances of a morpheme. But coalescence in Sanskrit is not a neutralization rule if all [eː]s are derived from /ai/.

Because mid vowels are prohibited in general, *MID must dominate ID(VH), so any input /e/ or /eː/ drawn from the rich base will map unfaithfully to a high or low vowel:

(3) *MID » ID(VH)

/be(ː)/	*MID	ID(VH)
→ bi(ː) *or* ba(ː)		1
~ be(ː)	$_1$ W	L

In this comparative tableau (Prince 2002), the winning candidate's violation marks are displayed as integers in the first row below the constraints. Subsequent rows compare the winner with various losers, which are preceded by a tilde. The W indicates that *MID favors the winner over this loser; the L says that ID(VH) favors the loser over the winner. (The loser's number of violation marks is also shown in small print.) This tableau is telling us that winner-favoring *MID dominates loser-favoring ID(VH).

Coalescence of diphthongs requires a ranking where *DIPH dominates both UNIF and ID(VH):

(4) *DIPH » UNIF, ID(VH)

/va$_1$-i$_2$/	*DIPH	UNIF	ID(VH)
→ ve:$_{1,2}$		1	2
~ va$_1$i$_2$	$_1$ W	L	L

The winning candidate in (4) incurs two marks from ID(VH) because each of its input vowels stands in correspondence with an output vowel of a different height.

For Gnanadesikan, the key to analyzing Sanskrit is the constraint ID-ADJ(VH), since it blocks the *MID-satisfying map of /ai/ to [iː] or [aː] that would otherwise be expected. The idea is that fusing /ai/ to create a vowel that is either high or low establishes correspondence between one of the input vowels and an output vowel that is two steps away from it in height. ID-ADJ(VH) specifically excludes this map even though it offers satisfaction of *MID:

(5) ID-ADJ(VH) » *MID

/va$_1$-i$_2$/	ID-ADJ(VH)	*MID
→ ve:$_{1,2}$		1
~ viː$_{1,2}$ *or* vaː$_{1,2}$	$_1$ W	L

As the subscripts indicate, the candidates in (5) are products of segmental coalescence, not segmental deletion. Therefore they owe allegiance, IDENT-wise, to both of their input parents.

This analysis is a typical chain shift.[6] Underlying /ai/ becomes [eː], whereas underlying /eː/ and /e/ become something else, either a high vowel or a low one (it does not matter which). The constraint ID-ADJ(VH) blocks the fell-swoop map in which /ai/ coalesces to form a high or low vowel. In any chain shift /A/ → [B] and /B/ → [C], there must be a faithfulness constraint that is violated by the forbidden /A/ → [C] map but not by either of the permitted maps (Moreton 2000). ID-ADJ(VH) is such a constraint.

No matter how Sanskrit is analyzed —with a MSC against mid vowels or the ranking ⟦*MID » ID(VH)⟧ in (3)— it presents a learning problem. Learners who hear nonalternating mid vowels will assume that they are derived from underlying mid vowels. How are they able to override this natural bias toward the identity map and instead learn a grammar that excludes underlying mid vowels with a MSC or, what amounts to almost the same thing, maps all of them unfaithfully? To see this problem in more concrete terms, we will look at how learning of an OT grammar is accomplished (there being no extant proposals for the learning of MSCs).

The learner is armed with the learning procedure known as biased multirecursive constraint demotion (BMCD) (Prince and Tesar 2004). In multirecursive constraint demotion (Tesar 1997), learning proceeds on the basis of a comparative tableau like those given above, though with multiple inputs considered simultaneously. (See Prince 2002 on comparative tableaux in constraint demotion.) Constraints are ranked one at a time starting from the top of the hierarchy. A constraint is rankable if it favors no losers among the candidates that have not yet been accounted for. (A losing candidate has been accounted for if it is dispreferred relative to the winner by some constraint that has already been ranked.) Once some constraint has been ranked, the tableau with this new ranking is then submitted to the same procedure, recursively. The procedure terminates when all constraints have been ranked or when there are irreducible conflicts (as in (16) below), in which case no grammar can be found.

The bias in BMCD favors grammars that are more restrictive (Prince and Tesar 2004; see also Hayes 2004, and Itô and Mester 1999). The bias says which constraints to rank first (=higher) when there are two or more constraints that favor no losers. By preference, markedness constraints are ranked higher. If there are no rankable markedness constraints, the ranking preference perforce goes to a faithfulness constraint. When several faithfulness constraints are available for ranking, the algorithm chooses the one that, by accounting for certain candidates, allows a markedness constraint to be ranked on the next recursive pass. (This is an intentional simplification; see Prince and Tesar 2004 for further details that are not important in the current context.)

At the earliest stages of learning, there is presumably little or no awareness of morphological structure and no awareness of alternations, so only the phonotactics are being learned (on this assumption, see Hayes 2004; Pater and Tessier 2003; Prince and Tesar 2004). The phonotactic learner seeks a grammar that performs identity maps from the inferred lexicon to the observed surface forms. In the case of Sanskrit,

6. On chain shifts in acquisition, see Dinnsen (2004); Dinnsen and Barlow (1998); Dinnsen, O'Connor, and Gierut (2001).

the ambient language contains words with the vowels [i], [iː], [a], [aː], and [eː]. A tableau is given in (6); all constraints are assumed to be unranked, though the algorithm works equally well if they have some ranking imposed on them already.

(6) Sanskrit at onset of phonotactic learning (hypothetical examples)

lexicon	cands.	*MID	*DIPH	ID(VH)	ID-ADJ(VH)	UNIF
/ba/	→ ba					
(similarly /baː/)	~ bi			$_{1}$ W	$_{1}$ W	
	~ be	$_{1}$ W		$_{1}$ W		
/bi/	→ bi					
(similarly /biː/)	~ ba			$_{1}$ W	$_{1}$ W	
	~ be	$_{1}$ W		$_{1}$ W		
/beː/	→ beː	1				
	~ biː *or* baː	L		$_{1}$ W		

I will refer to this as a «support tableau», after Tesar and Prince (forthcoming), since it provides the support for the rankings found by BMCD. In this tableau, the initially rankable constraints (i.e., the constraints that favor no losers) are *DIPH, ID(VH), ID-ADJ(VH), and UNIF. The bias in BMCD favors ranking markedness constraints first, and of these only *DIPH is a markedness constraint, so it goes in the top rank of the constraint hierarchy. But because *DIPH favors no winners either, ranking it does not account for any losing candidates.

On the next recursive pass through BMCD, no rankable markedness constraints can be found; *MID is the only remaining markedness constraint, and it favors the loser from the input /beː/. Therefore, BMCD must rank a faithfulness constraint, and it selects the faithfulness constraint ID(VH) (see (7)), because adding ID(VH) to the ranking accounts for the loser from input /beː/, and this will allow a markedness constraint, *MID, to be ranked on the next pass.

(7) Support tableau (6) after second ranking pass

lexicon	cands.	*DIPH	ID(VH)	*MID	ID-ADJ(VH)	UNIF
/ba/	→ ba					
(similarly /baː/)	~ bi		$_{1}$ W		$_{1}$ W	
	~ be		$_{1}$ W	$_{1}$ W		
/bi/	→ bi					
(similarly /biː/)	~ ba		$_{1}$ W		$_{1}$ W	
	~ be		$_{1}$ W	$_{1}$ W		
/beː/	→ beː			1		
	~ biː *or* baː		$_{1}$ W	L		

All of the loser rows have been shaded to indicate that these candidates have been accounted for by ranking ID(VH). Since no unaccounted-for candidates remain, *MID is obviously free to be ranked. Because of the bias in BMCD, the remaining faithfulness constraints go below *MID, yielding the result in (8).

(8) Sanskrit ranking after phonotactic learning
*DIPH » ID(VH) » *MID » ID-ADJ(VH), UNIF

The grammar in (8) is consistent with the observed vowel system of Sanskrit, but it is also consistent with a proper superset of Sanskrit in which short mid vowels are permitted: *[be]. (We can infer this because, if this grammar is given the input /be/, it yields the output *[be].) This grammar, then, is insufficiently restrictive, allowing things that Sanskrit does not allow. This is an instance of the Subset Problem (Angluin 1980; Baker 1979): learners presented with only positive evidence cannot proceed from a less restrictive grammar to a more restrictive one. The bias in BMCD is intended to address this problem by ensuring that learners always favor the most restrictive grammar that is consistent with the primary data, but BMCD is insufficient in this case (see also Alderete and Tesar 2002). The goal of this article is to explain how learners find the more restrictive grammar in cases like Sanskrit.

Another way to look at the problem with (8) is that Sanskrit is harmonically incomplete relative to the markedness constraints provided: it lacks short mid vowels, yet there is no markedness constraint against them. (On harmonic (in)completeness, see Prince and Smolensky 2004: 164, 219-220.) As Gnanadesikan's analysis shows, Sanskrit's harmonically incomplete system can be derived by judiciously deploying the faithfulness constraint ID-ADJ(VH), but the phonotactic learner has as yet no evidence from alternations that show mid vowels being derived from underlying diphthongs. Worse yet, even with that evidence, the morphophonemic learner also fails to acquire the correct grammar, as we will see shortly.

When a system is unlearnable because it is harmonically incomplete relative to a set of markedness constraints, there is a straightforward solution: make it harmonically complete, and therefore learnable with BMCD, by enlarging the set of markedness constraints. If the universal constraint component CON includes a constraint against *short* mid vowels, then BMCD will without difficulty find a grammar that is no less restrictive than Sanskrit requires. Such a constraint is not implausible as a narrowly tailored solution for Sanskrit,[7] but it is no help in dealing with the larger problem of learning similar systems. In the appendix of this article (see also McCarthy 2004), I document additional cases of the same learning problem and show for each of them that they cannot be solved by the expedient of introducing an additional markedness constraint. Because it is simple and familiar, I will stick with the Sanskrit example to exemplify the proposal made here, but bear in mind that any putative alternative solutions, to be of any real interest, need to deal with the cases in the appendix as well.

7. I am grateful to John Kingston for raising this issue.

As language development proceeds, the Sanskrit learner begins to attend to morphological structure and this leads, by a process I will not be discussing (though see the references in section 1), to the discovery of unfaithful maps from input to output. Let us say, for instance, that the learner has realized that the isolation forms *tava* and *indra* are components of the phrase *tave:ndra*. The learner adds this new information about /ai/ → [eː] unfaithful maps to the support tableau, yielding (9).

(9) Sanskrit support tableau after some morphological analysis

lexicon	cands.	*DIPH	ID(VH)	*MID	ID-ADJ(VH)	UNIF
/ba/	→ ba					
(similarly /baː/)	~ bi		$_1$ W		$_1$ W	
	~ be		$_1$ W	$_1$ W		
/bi/	→ bi					
(similarly /biː/)	~ ba		$_1$ W		$_1$ W	
	~ be		$_1$ W	$_1$ W		
/beː/	→ beː			1		
	~ biː *or* baː		$_1$ W	L		
/va$_1$+i$_2$/	→ veː$_{1,2}$		2	1		1
	~ va$_1$i$_2$	$_1$ W	L	L		L
	~ viː$_{1,2}$ *or* vaː$_{1,2}$		$_1$ L	L	$_1$ W	$_1$

Three comments on this tableau:

(i) Recall that I am making the simplifying assumption that there is coalescence but no deletion (because MAX » UNIF). The correspondence relations in the /va$_1$+i$_2$/ rows presuppose this: the vowels in all candidates bear the indices of their input correspondents /a$_1$/ and /i$_2$/. ID(VH) assigns a mark for every pair of corresponding segments that differ in height. The candidate [veː$_{1,2}$] has two such pairs, (a$_1$, eː$_{1,2}$) and (i$_2$, eː$_{1,2}$), whereas [viː$_{1,2}$] has only the pair (a$_1$, iː$_{1,2}$). Unlike [veː$_{1,2}$], however, [viː$_{1,2}$] violates ID-ADJ(VH) because the pair (a$_1$, iː$_{1,2}$) contains corresponding vowels that differ by two degrees of height and not just one.

(ii) This tableau contains the /beː/ → [beː] identity map because there are forms that never alternate and/or forms whose alternations have escaped the learner's experience or attention. This lingering identity map is the crux of the learning problem: because it is there in the tableau, BMCD continues to find a grammar that is insufficiently restrictive.

(iii) This tableau preserves the ranking (8) that was acquired during phonotactic learning. With the added information about /va$_1$+i$_2$/ → [veː$_{1,2}$], this ranking is now incorrect —the bottom row has loser-favoring constraints dominating the highest-ranking winner-favoring constraint. Because of this discrepancy between the ranking and the result, a new round of BMCD is required.

On the first pass of (9) through BMCD, the only markedness constraint that favors no losers is again *DIPH, so it is ranked at the top of the hierarchy:

(10) Support tableau (9) after first ranking pass. Ranking: ⟦*DIPH » rest⟧

lexicon	cands.	*DIPH	*MID	ID(VH)	ID-ADJ(VH)	UNIF
/ba/	→ ba					
(similarly /baː/)	~ bi			$_1$ W	$_1$ W	
	~ be		$_1$ W	$_1$ W		
/bi/	→ bi					
(similarly /biː/)	~ ba			$_1$ W	$_1$ W	
	~ be		$_1$ W	$_1$ W		
/beː/	→ beː		1			
	~ biː *or* baː		L	$_1$ W		
/va$_1$+i$_2$/	→ veː$_{1,2}$		1	2		1
	~ va$_1$i$_2$	$_1$ W	L	L		L
	~ viː$_{1,2}$ *or* vaː$_{1,2}$		L	$_1$ L	$_1$ W	$_1$

Now there are no markedness constraints that favor no losers, so the only choice is to rank a faithfulness constraint. The only rankable faithfulness constraint is ID-ADJ(VH):

(11) Support tableau (10) after second ranking pass. Ranking: ⟦*DIPH » ID-ADJ(VH) » rest⟧

lexicon	cands.	*DIPH	ID-ADJ(VH)	*MID	ID(VH)	UNIF
/ba/	→ ba					
(similarly /baː/)	~ bi		$_1$ W		$_1$ W	
	~ be			$_1$ W	$_1$ W	
/bi/	→ bi					
(similarly /biː/)	~ ba		$_1$ W		$_1$ W	
	~ be			$_1$ W	$_1$ W	
/beː/	→ beː			1		
	~ biː *or* baː			L	$_1$ W	
/va$_1$+i$_2$/	→ veː$_{1,2}$			1	2	1
	~ va$_1$i$_2$	$_1$ W		L	L	L
	~ viː$_{1,2}$ *or* vaː$_{1,2}$		$_1$ W	L	$_1$ L	$_1$

The remaining markedness constraint, *MID, still favors one loser, so it is not yet rankable. Both of the remaining faithfulness constraints, ID(VH) and UNIF, are rankable, however. The bias in BMCD favors ranking ID(VH) first because this will allow *MID to be ranked on the next round. Once ID(VH) has been ranked, *MID is ranked next, and UNIF goes at the bottom of the hierarchy. The final result of BMCD is given in (12).

(12) Sanskrit ranking after morphophonemic learning
*DIPH » ID-ADJ(VH) » ID(VH) » *MID » UNIF

The problem with (12) is that it is insufficiently restrictive. Like the ranking in (8) that emerged from phonotactic learning, the ranking in (12) produces surface short mid vowels if presented with input /e/ from the rich base. Gnanadesikan's chain-shift analysis, which derives all surface mid vowels from diphthongs and maps all underlying mid vowels onto high or low vowels, cannot be learned using just BMCD and this constraint set. It is unlearnable because the lingering presence of the /beː/ → [beː] identity map in the support tableau forces the ranking ID(VH) » *MID —exactly the ranking that makes this the grammar of the superset language. This result confirms what Alderete and Tesar (2002) have argued on independent grounds: there are circumstances where a commitment to the identity map can lead to a superset language. An adequate learning theory needs to explain how learners get to the more restrictive grammar that is required for Sanskirt and the other languages discussed in the appendix.

3. The Free-Ride Learning Algorithm

This section presents the proposal at three different levels of detail. The first is a brief overview; the second is a detailed textual presentation interspersed with worked examples; and the third is a formal statement of the algorithm in (24).

The morphophonemic learning problem in Sanskrit can be described as follows. Learners can infer from alternations that *some* [eː]s are derived from /a-i/ sequences. To learn a sufficiently restrictive grammar, however, learners must generalize from alternating [eː]s to nonalternating [eː]s, eliminating the /eː/ → [eː] identity map from the support tableau and instead deriving *all* [eː]s from /a-i/ sequences, even in forms with no (observed) alternations.

The essence of my proposal is that learners, whenever alternations lead them to discover a new unfaithful map, always attempt to generalize that map across the entire support tableau and thereby across the entire language. This is the free ride: nonalternating [eː]s take a free ride on the /ai/ → [eː] map that was discovered from alternations. Because this move eliminates the /eː/ → [eː] identity map, the revised support tableau has fewer maps, and this allows BMCD to find a more restrictive grammar. If the identity map bespeaks economy of derivation, the free ride bespeaks profligacy, but profligacy that leads to greater restrictiveness where it really matters, in the grammar. (On the decidedly harder problem of contextually restricted free rides —e.g., all *word-final* vowels are underlying long because some are— see section 4.)

Sometimes the free ride is a trip to nowhere. There are two situations where generalizing in this way does not work out. If other similar identity maps remain in the support tableau after taking the free ride, then BMCD may yield a grammar that is unchanged, so the exercise was pointless. Or if the generalized unfaithful map was authentically neutralizing, then BMCD will fail to find any grammar. (That is, BMCD will be unable to rank all constraints.) In either of these situations, the free-ride generalization is rejected.

Some formal development is necessary before we can put these ideas into practice. We require a definition of an unfaithful map, a statement of how to alter the support tableau in order to take the free ride, and a way of deciding whether or not the free ride is appropriate.

The notion «unfaithful map» appears in some of my previous work on opacity (McCarthy 2003a, 2003b), but to get a solid handle on what this means requires some basic changes in the theory of correspondence.[8] Under the original formalization of correspondence in McCarthy and Prince (1995, 1999), some kinds of unfaithfulness simply do not involve maps from input to output: a deleted segment is not in the range of the correspondence relation $\mathfrak{R}$, and an epenthetic segment is not in the domain of $\mathfrak{R}$. Deletion and epenthesis, then, are not input $\rightarrow$ output maps, though other kinds of unfaithfulness are input $\rightarrow$ output maps. A coherent definition of «unfaithful map» is elusive in the original correspondence model.

The necessary revisions of correspondence theory are proposed by McCarthy and Wolf (2005) for a very different reason, to support a novel characterization of the null output. The main idea is that $\mathfrak{R}$ is not a relation between the input and output directly, but rather it is a relation between a concatenative decomposition of the input and a concatenative decomposition of the output. «Concatenative decomposition» is defined in the following quotation:

> Concatenative Decomposition
>
> A concatenative decomposition of a string S is a sequence of strings $\langle d_i \rangle_{j \le i \le k}$ such that $d_j{}^\frown \ldots {}^\frown d_k = S$.
>
> The concatenative decompositions of a given string are numerous indeed, because any of the d_i may correspond to the empty string *e*, which has the property that $s^\frown e = e^\frown s = s$, for any string s. Compare the role of 0 in addition: 3+0 = 0+3 = 0+3+0 = 3. All these refer to the same number, but all are distinct as expressions. The notion «concatenative decomposition» allows us to distinguish among the different ways of expressing a string as a sequence of binary concatenations. (McCarthy and Prince 1993: 89-90)

In any input $\rightarrow$ output mapping $x \mathfrak{R} y$, *x* and *y* are strings taken from the concatenative decompositions of the input and output, respectively. The strings *x* and *y* may be empty, they may be monosegmental, or they may be multisegmental. When a segment deletes, a monosegmental string in a concatenative decomposition of the input is in correspondence with an empty string in a concatenative decomposition

8. I am grateful to Bruce Tesar for raising this issue.

of the output: <p, a, t> → <p, a, Ø>. (The empty string is written as Ø rather than the more usual *e* to avoid confusion with the segment [e].) Segmental epenthesis is the dual of deletion: <Ø, a, p> → <ʔ, a, p>. Segmental coalescence, as in Sanskrit, places a multisegmental input string in correspondence with a monosegmental output string: <p, ai> → <p, eː>. The (input, output) pairs (t, Ø) (short for *t*ℜØ), (Ø, ʔ), and (ai, eː) are all examples of unfaithful mappings. They are simply pairs of corresponding strings that violate at least one faithfulness constraint. (For further details, see McCarthy and Wolf 2005.)

In morphophonemic learning, alternations among morphologically related forms lead the learner to posit underlying representations that require unfaithful maps. For example, the Sanskrit learner who has concluded that /tava indra/ underlies *taveːndra* will infer a correspondence relation that includes the map (a_3i_4, $eː_{3,4}$). This map is unfaithful because it violates ID(VH) and UNIF.

I assume that learners maintain a set of unfaithful maps 𝔏, which they have learned from analyzing paradigmatic alternation data.[9] Whenever the lexicon column of the support tableau changes, 𝔏 is updated, and if it changes, the free-ride learning algorithm (hereafter, FRLA) is invoked.

The free-ride procedure is roughly as follows (see (24) for a more explicit statement). Assume that the learner has a support tableau τ_o (*o* for old) derived from previous learning and a list 𝔏 of (alternation-supported) unfaithful maps in τ_o. Start with a member of 𝔏, the unfaithful map (A, B). Make a copy of τ_o called τ_n (*n* for new) and locate all (B, B) identity maps in τ_n. Perform «surgery» (Tesar *et al.* 2003) on τ_n, changing the underlying form of any morpheme that undergoes the (B, B) map, replacing it with a (A, B) map, and updating the faithfulness assessments accordingly.

Now τ_n has the same outputs as τ_o but with a different lexicon. Submit τ_n to BMCD. If BMCD fails to find a grammar for τ_n, discard τ_n and go back to τ_o. If BMCD finds a grammar for τ_n, compare it with the grammar for τ_o. Of these two choices, the learner adopts the one that leads to a more restrictive grammar. In other words, the learner chooses the lexicon that is compatible with the subset language.

This process continues with the other unfaithful maps in 𝔏 and with all of the subsets of the unfaithful maps in 𝔏. If a free ride is warranted, then the support tableau at the end of the process will have different underlying representations than at the beginning of the process, and the grammar will be more restrictive.

BMCD's ability to detect inconsistency plays a key role in this process. Recall that multirecursive constraint demotion, of which BMCD is a special case, fails to find a ranking when there are two or more constraints left to rank and all favor at least one loser, so none can be ranked. If no grammar can be found under the spe-

9. There is a systematic ambiguity in the term «unfaithful map». In a correspondence relation, an unfaithful (or faithful) map is a pairing of specific segments in the input and output strings. In 𝔏, however, an unfaithful map is a pairing of segment types, lifted out of the particulars of the word that instantiates that map. Typographically, unfaithful maps of the former type are sometimes supplied with segmental subscripts, whereas those of the latter type have no subscripts. In actual practice, there is little danger of mixing up the two senses of the term.

cific assumptions about underlying representation that τ_n embodies, then τ_n must have taken the free ride too far; it has eliminated at least one ineliminable identity map. By using BMCD in this way, I am directly following Tesar *et al.* (2003) and Tesar and Prince (forthcoming), who use inconsistency-detection to reject hypothesized underlying representations for alternating morphemes, and less directly following Kager (1999: 334), who invokes inconsistency-detection to trigger modifications in the lexicon.

If BMCD succeeds in finding a ranking for τ_n, a decision must be made: which hypothesis about the lexicon is superior, the one embodied by τ_n or the one embodied by τ_o? As the discussion of Sanskrit showed, we seek the lexicon that allows for a more restrictive grammar, thereby avoiding the Subset Problem. The choice between τ_n and τ_o therefore requires a metric for grammar restrictiveness, and one exists in the form of the «r-measure» of Prince and Tesar (2004). It is defined in (13).[10]

(13) R-measure (Prince and Tesar 2004: 252)
The r-measure for a constraint hierarchy is determined by adding, for each faithfulness constraint in the hierarchy, the number of markedness constraints that dominate that faithfulness constraint.

Grammars with higher r-measures are in general more restrictive, since they grant more power to markedness constraints. Prince and Tesar note that grammars with equal r-measures may also differ in restrictiveness, but this as-yet imperfect metric appears to be sufficient for present purposes. In short, τ_n beats τ_o if τ_n supports a grammar with a higher r-measure.

The following examples briefly illustrate the FRLA. They are simplified, with artificially restricted support tableaux and a single unfaithful map. Three outcomes are considered: (i) giving a free ride to an identity map produces a support tableau for which no grammar can be found (devoicing in German); (ii) the free ride leads to a grammar that is the same as the one without the free ride (hypothetical Sanskrit′ with short mid vowels); and (iii) the free ride leads to a more restrictive grammar (real Sanskrit).

(i) *No grammar is found.* If taking a free ride leads to a demonstrably wrong surmise about underlying representations, no grammar will be found by BMCD, and the offending support tableau is rejected.

This happens when a process of neutralization is the source of the unfaithful map. Imagine, for example, a learner of German whose first morphophonemic discovery is that /raːd/ is the underlying form of [raːt] ‘wheel’, so $\mathfrak{L}$={(d, t)}. This learner has the support tableau in (14) prior to applying the FRLA:

10. I am grateful to Joe Pater for suggesting that I use the r-measure here.

(14) German morphophonemic learning – τ_o

lexicon	cands.	*VcdObst	*VcdObstCoda	Id(voice)
/daŋk/ 'thanks'	→ daŋk	1		
	~ taŋk	L		$_1$ W
/taːt/ 'deed'	→ taːt			
	~ daːd	$_2$ W	$_1$ W	$_2$ W
	~ daːt	$_1$ W		$_1$ W
/raːd/ 'wheel'	→ raːt			1
	~ raːd	$_1$ W	$_1$ W	L

Applying BMCD to (14) yields the ranking ⟦*VcdObstCoda » Id(voice) » *VcdObst⟧.

The FRLA makes a copy of (14) in which all (t, t) identity maps have been replaced by the (d, t) unfaithful map in $\mathfrak{L}$. The support tableau with this free ride is given in (15).

(15) Free ride on (d, t) in German – τ_n

lexicon	cands.	*VcdObst	*VcdObstCoda	Id(voice)
/daŋk/	→ daŋk	1		
	~ taŋk	L		$_1$ W
/daːd/	→ taːt			2
	~ daːd	$_2$ W	$_1$ W	L
	~ daːt	$_1$ W		$_1$ L
/raːd/	→ raːt			1
	~ raːd	$_1$ W	$_1$ W	L

The difference between (14) and (15) is that (14) derives [taːt] by identity maps from /taːt/, while (15) derives it from /daːd/ by taking a free ride on the /d/ → [t] unfaithful map. Obviously, (15) is not the outcome we seek: because /d/ → [t] is neutralizing, it is a grievous error to attempt to derive all surface [t]s from underlying /d/s. The FRLA creates (15) as a hypothesis about the German lexicon, but this hypothesis is soon rejected when it turns out that no constraint ranking is consistent with it.

On (15)'s first pass through BMCD, *VcdObstCoda is placed at the top of the constraint hierarchy, thereby disposing of two losing candidates, as shown in (16).

(16) Support tableau (15) after first ranking pass

lexicon	cands.	*VcdObstCoda	*VcdObst	Id(voice)
/daŋk/	→ daŋk		1	
	~ taŋk		L	1 W
/da:d/	→ ta:t			2
	~ da:d	1 W	2 W	L
	~ da:t		1 W	1 L
/ra:d/	→ ra:t			1
	~ ra:d	1 W	1 W	L

But now it is clear that *VcdObst and Id(voice) place antithetical demands on ranking, since both favor some unaccounted-for losers. This conflict is irreducible, so no further ranking is possible. BMCD has failed to find a grammar for German under the specific assumptions about underlying representations embodied in (15). This failure is unsurprising: devoicing is a neutralization process in German, since the [d]/[t] contrast is maintained in onset position, so there is no hope of deriving all [t]s from /d/s. Because BMCD fails to find a grammar, this new support tableau is discarded, as is the hypothesis it embodies about the German lexicon. Instead, the original support tableau (14) is correctly retained by the learner.[11, 12]

Taking a free ride on a neutralization process is a bad choice. The FRLA requires learners to consider this choice, but through BMCD it also supplies a way of rapidly detecting the error and recovering from it.

11. If the learner of German has insufficient phonotactic data at the point when the FRLA is activated, then he is in danger of reaching the wrong conclusion about the lexicon. Consider, for example, the effect of removing /daŋk/ from the support tableau, so there are no words in the support with surface syllable-initial voiced obstruents. In that case, a grammar —the wrong one— will be found by BMCD. (I am grateful to Anne-Michelle Tessier for raising this point.)

    In reality, there is probably nothing to worry about. Morphophonemic learning requires prior morphological analysis, whereas phonotactic learning can precede all morphological analysis. Therefore, morphophonemic learning is unlikely to begin before phonotactic learning is already quite far advanced. It is therefore reasonable to assume that the morphophonemic learner has already accumulated data that are sufficiently representative of the language's phonotactic possibilities.

12. The German example indicates the need to say more about how FRLA interfaces with other aspects of learning. Bruce Tesar points out that the two losers from input /ta:t/ in support tableau (14) contribute no information about ranking, since no constraint favors them; hence, the standard approach to identifying informative losers, error-driven learning (Tesar 1998), would not include them in the support. Since these losers are informative in tableau (15), Tesar suggests that FRLA needs to apply error-driven learning to τ_n to find any losers that were uninformative in τ_o but have become informative in τ_n as a result of the change in underlying representations.

(ii) *A grammar is found, but the ranking is unchanged.* In this situation, nothing is gained in restrictiveness by taking the free ride. I assume that the lingering bias toward the identity map disfavors this use of the free ride. (See footnote 13 on why this is only an assumption rather than a solid fact.)

As an illustration of this possibility, imagine a language Sanskrit′ that is identical to real Sanskrit except that the inventory includes short mid vowels. Suppose the learner of Sanskrit′ is in the early stages of morphophonemic analysis and has just discovered the coalescence alternation, so $\mathfrak{A}$ = {(ai, eː)}. The pre-free-ride support tableau τ_o is given in (17). The FRLA produces the support tableau (18), which is Sanskrit′ with all (eː, eː) identity maps replaced by (ai, eː).

(17) Sanskrit′ without free ride – τ_o

lexicon	cands.	*MID	*DIPH	ID(VH)	ID-ADJ(VH)	UNIF
/ba/	→ ba					
(similarly /baː/)	~ bi			$_1$ W	$_1$ W	
	~ be	$_1$ W		$_1$ W		
/bi/	→ bi					
(similarly /biː/)	~ ba			$_1$ W	$_1$ W	
	~ be	$_1$ W		$_1$ W		
/be/	→ be		1			
	~ biː *or* ba	L		$_1$ W		
/beː/	→ beː	1				
	~ biː *or* baː	L		$_1$ W		
/va$_1$+i$_2$/	→ veː$_{1,2}$	1		2		1
	~ va$_1$i$_2$	L	$_1$ W	L		L
	~ viː$_{1,2}$ *or* vaː$_{1,2}$	L		$_1$ L	$_1$ W	$_1$

(18) Sanskrit′ with (ai, eː) free ride – τ_n

lexicon	cands.	*Mid	*Diph	Id(VH)	Id-Adj(VH)	Unif
/ba/	→ ba					
(similarly /baː/)	~ bi			$_1$ W	$_1$ W	
	~ be	$_1$ W		$_1$ W		
/bi/	→ bi					
(similarly /biː/)	~ ba			$_1$ W	$_1$ W	
	~ be	$_1$ W		$_1$ W		
/be/	→ be	1				
	~ biː *or* ba	L		$_1$ W		
/ba$_1$i$_2$/	→ beː$_{1,2}$	1		2		1
	~ ba$_1$i$_2$	L	$_1$ W	L		L
	~ biː$_{1,2}$ *or* baː$_{1,2}$	L		$_1$ L	$_1$ W	$_1$
/va$_1$+i$_2$/	→ veː$_{1,2}$	1		2		1
	~ va$_1$i$_2$	L	$_1$ W	L		L
	~ viː$_{1,2}$ *or* vaː$_{1,2}$	L		$_1$ L	$_1$ W	$_1$

These two support tableaux, old and new, get identical grammars from BMCD: ⟦*Diph » Id-Adj(VH) » Id(VH) » *Mid » Unif⟧. Nothing is gained in restrictiveness by deriving nonalternating [eː] from /ai/, since the lingering /e/ → [e] identity map in (18) will not allow *Mid to dominate Id(VH). In situations like this, where restrictiveness is not at stake, we may surmise that the learner does not take the pointless free ride. Therefore, the original support tableau (17) is retained and (18) is discarded.[13] In Sanskrit′, nonalternating [eː] is derived from /eː/, not /ai/, because [e] exists and is derived from /e/.

(iii) *A more restrictive grammar is found.* This is the situation in real Sanskrit. Assume that the morphophonemic learner has already arrived at the stage represented by the support tableau (19), ranked according to (12).

13. It is only an assumption rather than an established fact that learners choose (17) over (18), favoring the identity map when all else is equal. Ordinary linguistic data is unhelpful in such situations, though other evidence may be relevant. For related discussion, see e.g. Kiparsky (1973) or Dresher (1981) on the Alternation Condition.

(19) Sanskrit without free ride – τ_o

lexicon	cands.	*DIPH	ID-ADJ(VH)	ID(VH)	*MID	UNIF
/ba/	→ ba					
(similarly /baː/)	~ bi		$_1$ W	$_1$ W		
	~ be			$_1$ W	$_1$ W	
/bi/	→ bi					
(similarly /biː/)	~ ba		$_1$ W	$_1$ W		
	~ be			$_1$ W	$_1$ W	
/beː/	→ beː				1	
	~ biː *or* baː			$_1$ W	L	
/va$_1$+i$_2$/	→ veː$_{1,2}$			2	1	1
	~ va$_1$i$_2$	$_1$ W		L	L	L
	~ viː$_{1,2}$ *or* vaː$_{1,2}$		$_1$ W	$_1$ L	L	$_1$

At this point, $\mathfrak{L}$ = {(ai, eː)}. A copy of (19) is made, substituting the (ai, eː) unfaithful map for all (eː, eː) identity maps. The table with this free ride is shown in (20).

(20) Sanskrit with (ai, eː) free ride – τ_n

lexicon	cands.	*DIPH	ID-ADJ(VH)	ID(VH)	*MID	UNIF
/ba/	→ ba					
(similarly /baː/)	~ bi		$_1$ W	$_1$ W		
	~ be			$_1$ W	$_1$ W	
/bi/	→ bi					
(similarly /biː/)	~ ba		$_1$ W	$_1$ W		
	~ be			$_1$ W	$_1$ W	
/ba$_1$i$_2$/	→ beː$_{1,2}$			2	1	1
	~ ba$_1$i$_2$	$_1$ W		L	L	L
	~ biː$_{1,2}$ *or* baː$_{1,2}$		$_1$ W	$_1$ L	L	$_1$
/va$_1$+i$_2$/	→ veː$_{1,2}$			2	1	1
	~ va$_1$i$_2$	$_1$ W		L	L	L
	~ viː$_{1,2}$ *or* vaː$_{1,2}$		$_1$ W	$_1$ L	L	$_1$

BMCD is applied to the new support tableau (20), yielding the ranking in (21).

(21) Ranking from τ_n (20)
*DIPH » ID-ADJ(VH) » *MID » ID(VH), UNIF

Since a grammar has been found, it needs to be compared for restrictiveness with the grammar derived from the original support tableau (19). We already saw that ranking in (12), and it is repeated in (22):

(22) Ranking of τ_o (19)
*DIPH » ID-ADJ(VH) » ID(VH) » *MID » UNIF

These grammars differ in restrictiveness. The grammar in (21) for the new support tableau (20) is more restrictive because it contains the ranking ⟦*MID » ID(VH)⟧, which effectively denies faithful treatment to input mid vowels. The grammar in (22) for the old support tableau (19) has the opposite ranking of these two constraints and thereby permits input mid vowels to emerge faithfully. This grammar describes a proper superset of Sanskrit.

The rankings in (21) and (22) are competing descriptions of Sanskrit based on different assumptions about the lexicon. Both grammars are consistent with the primary data, but (21) is more restrictive. Formally, the FRLA favors (21) because (21) has a higher r-measure (see (13)): the r-measure of (21) is 5, while the r-measure of (22) is 4.[14] Since (21) is more restrictive, the learner adopts (21)'s support tableau (20) and the attendant hypotheses about underlying representations.

In the discussion thus far, I have pretended that a single unfaithful map, learned from alternations, is sufficient with the FRLA to lead learners to the more restrictive grammar (21). This is a simplification; Sanskrit learners cannot get to (21) until $\mathfrak{L}$ contains at least two unfaithful maps, (ai, eː) and (au, oː). The reason for this is that taking a free ride on only one of these unfaithful maps will leave the other identity map in the support tableau. For instance, if $\mathfrak{L}$ = {(ai, eː)}, then the support tableau in which (eː, eː) has been eliminated will still contain some (oː, oː) identity maps. And the presence of (oː, oː) in a tableau is sufficient to force ID(VH) to dominate *MID, as in the less restrictive grammar (22).

This matter is already addressed in the FRLA. The learner constructs a support tableau for each of the subsets of $\mathfrak{L}$ (including $\mathfrak{L}$ itself, of course). If the learner has not yet encountered or attended to alternation data that support the (au, oː) unfaithful map, then $\mathfrak{L}$ will not yet contain that map and the less restrictive grammar (22) will persist. Eventually, when the learner does get a handle on this other unfaithful map, he will be presented with a support tableau that, by virtue of two free rides, contains no (eː, eː) or (oː, oː) identity maps, and BMCD will find the more restrictive grammar in (21).

14. The r-measure of (21), for instance, is the sum of the number of markedness constraints that dominate ID-ADJ(VH) (that is, 1) and the number of markedness constraints that dominate ID(VH) and UNIF (that is, 2 each). The maximum r-measure for this constraint set is 6.

The FRLA is defined in this way so that grammar learning itself, rather than some special extragrammatical mechanism for data reduction, is responsible for discovering the generalization that all surface mid vowels derive from diphthongs. By itself, an unfaithful map is the meagerest sort of generalization: (A, B) pairs the input segment (or string) /A/ with output [B], abstracting away only from the actual form(s) in which this pairing occurs (see footnote 9). Unfaithful maps do not include information about the context in which they occur, nor do they state higher-order linguistic generalizations —e.g., there is no procedure for generalizing the (ai, eː) map to all mid vowels. Rather, the capacity for learning this or any other linguistic generalization is reserved to grammar learning by BMCD. In this way, we avoid an unwelcome redundancy, where the learning mechanism preprocesses the data in a way that is closely attuned to what the constraints are looking for.

A final point. Once the learner has successfully adopted a free ride, he ought to use this knowledge to figure out the underlying representations of newly-encountered words. A free ride is possible only when unfaithful mappings are biunique (that is, invertible): whenever the learner hears a novel word with [eː], he can safely infer underlying /ai/ without waiting to hear other members of the paradigm. As yet, however, I have no concrete proposal for how learners might exploit this undoubtedly useful information.

This section concludes with a formal statement of the FRLA. After the definitions in (23), (24) gives the algorithm in Algol/C-inspired pseudocode. (In definition (23), the distinction between phonological strings and their concatenative decompositions has been elided for clarity. In (24), the /* */ annotations surround comments.)

(23) Definitions

a. (Support) tableau
A (support) tableau τ is a set of ordered 6-tuples (input, winner, $\Re_{(\text{input, winner})}$, loser, $\Re_{(\text{input, loser})}$, tableau-row), where «input» is an underlying representation, «winner» and «loser» are output candidates, $\Re_{(\text{input, winner})}$ and $\Re_{(\text{input, loser})}$ are their respective correspondence relations with the input, and «tableau-row» is an ordered *n*-tuple of elements chosen from {W, L, Ø}, with one element for each constraint in CON.

b. Unfaithful map in τ
An unfaithful map in the (support) tableau τ is a member (α_i, β_i) of $\Re_{(\text{input, winner})}$ in some member of τ, where $\alpha_i \Re \beta_i$ violates some faithfulness constraint in CON. The set of all unfaithful maps in a tableau τ is designated by $\mathfrak{L}_\tau$.

(24) Free-ride learning algorithm
Given: an initial support tableau τ and a set $\mathfrak{L}_\tau$ of all unfaithful maps in τ.

FRLA(τ, $\mathfrak{L}_\tau$)
  $\tau_o := \tau$ /*initialize result*/
  For every $p_i \in \wp(\mathfrak{L}_\tau)$ /*$\wp(\mathfrak{L}_\tau)$=powerset of $\mathfrak{L}_\tau$*/
    $\tau_n :=$ **Subst**(τ_o, p_i) /*substitute unfaithful maps in $\mathfrak{L}$ subset p_i for identity maps in τ_o*/
    if **BMCD**(τ_n) = undefined /*if BMCD fails, go to next p_i*/
      continue
    else if **R-measure**(**BMCD**(τ_n)) > **R-measure**(**BMCD**(τ_o))
      $\tau_o := \tau_n$ /*if τ_n has a higher r-measure, then it is now the one to beat*/
    endif
  endfor
  return(τ_o)

Remarks:

- **FRLA**(τ, $\mathfrak{L}$) accepts as input a support tableau with a list of unfaithful maps contained in it. It returns a support tableau that may or may not be different from its input.
- **Subst**(τ, $\mathfrak{L}$) accepts as input a support tableau and a list of unfaithful maps. For every unfaithful map (α, β) in $\mathfrak{L}$, it locates any winners in τ with (β, β) identity maps (i.e., it checks whether (β, β) $\in \mathfrak{R}_{(\text{input, winner})}$). The morphemes underlying these winners are altered by substituting α for β. As in «surgery» (Tesar *et al.* 2003), these changes are carried over to all forms in τ that contain the affected morphemes, and constraint assessments are updated to reflect the altered inputs.
- **BMCD**(τ) is defined by Prince and Tesar (2004). It accepts a support tableau as input and returns a constraint hierarchy or *undefined* if no hierarchy can be found.
- **R-measure**($\mathcal{H}$) is defined by Prince and Tesar (2004) and in (13) above. It accepts as input a constraint hierarchy (with constraints tagged as markedness or faithfulness), and it returns a positive integer or zero.

The complexity of this learning algorithm is, in the worst case, proportional to the cardinality of $\wp(\mathfrak{L})$ —that is, 2^N, where N is the number of unfaithful maps in $\mathfrak{L}$. Exponentially increasing complexity looks bad, but in this case it is not as bad as it seems. As the learning problem scales up from examples of single processes to whole languages, or as the learner processes more alternation data, N grows very slowly. Most segments are mapped faithfully most of the time; alternation-supported unfaithful maps are more eye-catching, but they are also very much the exception. There is no fixed upper limit on N, but in practice languages with more than, say, 20 distinct alternation-supported unfaithful maps are probably not common. There is, then, little danger of a combinatorial explosion in more realistic learning situations than have been contemplated here.

Furthermore, there may be ways of improving the FRLA's efficiency in actual practice. The various members of $\wp(\mathfrak{L})$ are not equally likely to yield a grammar on their respective passes through the main loop of FRLA(τ, $\mathfrak{L}$). Because finding a grammar is roughly equivalent to finding a linguistically significant generalization, grammars are more likely to be found for free rides on those subsets of $\mathfrak{L}$ that (i) violate the same faithfulness constraint(s) and (ii) are maximal subject to (i). For example, the Sanskrit learner would be well advised to consider p_i={(ai, eː), (au, oː)} on an earlier pass through the loop than p_i={(ai, eː)} or p_i={(au, oː)}, since (ai, eː) and (au, oː) violate the same faithfulness constraint, UNIF. Once the free ride has been found, FRLA continues in search of other free rides, but the complexity of the problem is reduced significantly: because there are no remaining (eː, eː) or (oː, oː) identity maps in the support tableau, it is no longer necessary to examine members of $\wp(\mathfrak{L})$ that include maps with [eː] or [oː] as output.[15]

4. Limitations of and Extensions to the FRLA

The FRLA has a significant limitation: it works for across-the-board free rides, but it does not work for contextually restricted free rides. Colloquial Arabic word-final vowels are a case in point (for further analytic details, see McCarthy 2004). Word-final vowels are always short, though vowel length is contrastive in other positions. Alternations like [ˈjilqa]~[ˈjilqaːni] 'he will find ~ he will find me' are straightforwardly accounted for under the assumption that *some* word-final short vowels are derived from underlying long vowels: the mapping of unsuffixed /jilqaː/ to [ˈjilqa] is a familiar effect of NONFINALITY (NONFIN) and WEIGHT-TO-STRESS (WSP), which, by dominating ID(long), rule out *[jilˈqaː] and *[ˈjilqaː], respectively.

The naive expectation is that Arabic would have a contrast between final long vowel stems like /jilqaː/ and final short vowel stems like hypothetical /jitba/. The contrast would be neutralized word-finally —[ˈjilqa] and [ˈjitba]— and preserved before suffixes —[jilˈqaːni] and [ˈjitbani]. In reality, though, there is no such contrast: no stems behave like hypothetical *[ˈjitba]~[ˈjitbani], with a stem-final vowel that is short before suffixes. Every word-final vowel, although short on the surface, must be derived from an underlying long vowel, since all word-final short vowels alternate with long vowels when suffixed. Underlying stem-final short vowels served up by the rich base are presumably deleted, given the language's propensity for syncope. Deletion of underlying final short vowels is accomplished by the ranking ⟦FINAL-C » MAX(V)⟧ (see (25)). (FINAL-C requires every word to end in a consonant.) Final underlying long vowels are protected from deletion, however, by MAX(Vː) (see (26)), for which there is solid support elsewhere in the language.[16]

15. Bruce Tesar raises a related issue: could there be effects of the order in which the unfaithful mappings in $\mathfrak{L}$ are tried out? Schematically, if $\mathfrak{L}$ = {(A, B), (C, D)}, is it possible for the discovery of a free ride on (A, B) to block the discovery of a free ride on (C, D)? I have been unable to construct an example with this property, but of course this is no guarantee that the problem never arises.
16. On why there are no stems like *[ˈjitb]~[ˈjitbani], see McCarthy (2004).

(25) FINAL-C » MAX(V) (hypothetical example)

/jitba/	FINAL-C	MAX(V)
→ ˈjitb		1
~ ˈjitba	1 W	L

(26) MAX(Vː) » FINAL-C

/jilqaː/	MAX(Vː)	FINAL-C
→ ˈjilqa		1
~ ˈjilq	1 W	L

Colloquial Arabic therefore has a chain shift, but a chain shift that is limited to word-final position: /Vː#/ → [V#] and /V#/ → [Ø#]. Alternation data tell the learner that some [V#]s derive from /Vː#/s, and the FRLA ought to get the learner over the next hump to conclude that all [V#]s derive from /Vː#/s, even when they occur in nonalternating, never-suffixed stems like [ˈhuwwa] 'he'. If the learner fails to get over that hump, then he will acquire a grammar that maps /huwwa/ faithfully to [ˈhuwwa]. This less restrictive grammar includes the ranking ⟦MAX(V) » FINAL-C⟧ and wrongly allows stems like *[ˈjitba]~[ˈjitbani].

Unfortunately, the FRLA in its present form is not up to this task. It can deal with an across-the-board free ride like Sanskrit's, but not with a contextually limited one like Arabic's. Alternation data tell the learner that there is a (aː, a) unfaithful map in the language. But replacing all (a, a) identity maps with (aː, a) maps is a gross error: it would entail deriving [ˈkatab] 'he wrote' from /kaːtaːb/, and that is impossible because vowel length is contrastive in nonfinal open syllables, as shown by examples like [ˈkatab] 'wrote' and [ˈkaːtib] 'writing'. What is required is a way for learners to discover contextually limited free rides like the one in Colloquial Arabic: all (a, a) identity maps should be replaced by (aː, a), but only word-finally. Identity maps in other positions must be left intact.

There are two ways to fix this problem with the FRLA. One way is to greatly increase the power of the currently quite limited theory of unfaithful maps, so that they can be distinguished by the contexts in which the maps occur. For reasons already discussed (see the end of section 3), this seems like a poor idea: it requires a special auxiliary learning mechanism, independent of the grammar, that is more or less able to reproduce the interactional possibilities of NONFIN and WSP. The contexts in which unfaithful maps occur are defined by the grammar, and any special learning mechanism charged with discovering and recording those contexts would simply duplicate the effect of the grammar.

The other way of improving the FRLA is to provide it with a wider range of hypotheses about the lexicon. One of those hypotheses is a lexicon where only word-final [a]s take a free ride on (aː, a). That is, the learner must consider lexicons in which various subsets of the (a, a) identity maps in the support tableau have been replaced by (aː, a). These possibilities are listed in (27) for a dataset that contains a form with underlying /aː/ determined from alternations ([ˈjilqa]~[jilˈqaːni]), two forms with underlying nonfinal /aː/ ([kiˈtaːb], [ˈkaːtib]), and two forms with nonalternating surface [a] that might be eligible for a free ride ([ˈhuwwa], [ˈkatab]). The input segments that take a free ride are underlined.

(27) Hypothesized lexicons

a. /jilqaː/, /kitaːb/, /kaːtib/, /huwwa/, /katab/
b. /jilqaː/, /kitaːb/, /kaːtib/, /huwwa/, /kaːtab/
c. /jilqaː/, /kitaːb/, /kaːtib/, /huwwa/, /kataːb/
d. /jilqaː/, /kitaːb/, /kaːtib/, /huwwa/, /kaːtaːb/
e. /jilqaː/, /kitaːb/, /kaːtib/, /huwwaː/, /katab/
f. /jilqaː/, /kitaːb/, /kaːtib/, /huwwaː/, /kaːtab/
g. /jilqaː/, /kitaːb/, /kaːtib/, /huwwaː/, /kataːb/
h. /jilqaː/, /kitaːb/, /kaːtib/, /huwwaː/, /kaːtaːb/

Actual surface forms

[ˈjilqa], [kiˈtaːb], [ˈkaːtib], [ˈhuwwa], [ˈkatab]

Lexicon (27a) has no free rides. Lexicon (27h) has an across-the-board free ride: all [a]s everywhere are derived from /aː/. Lexicon (27e) is the desired result: only word-final [a] takes a free ride on the (aː, a) unfaithful map. Other lexicons represent other choices about which [a]s to grant a free ride to.

When presented with this wider range of possible lexicons, the FRLA will select the correct one, (27e). The short explanation is that BMCD cannot find a grammar for (27b, c, d, f, g, h), and (27e) is preferred to (27a) because (27e) allows for a grammar with a higher r-measure. The long explanation is that no grammar can be found for (27b, c, d, f, g, h) because no constraint ranking can derive [ˈkatab] from /kaːtab/, /kataːb/, or /kaːtaːb/ while also preserving the long vowels of /kitaːb/ and /kaːtib/. For example, (28) is a support tableau that presupposes the lexicon in (27h), where all [a]s are derived from /aː/.

(28) Across-the-board (aː, a) free ride (= lexicon (27h))[17]

lexicon	cands.	NonFin	WSP	Id(long)	Final-C	Max(Vː)	Max(V)
/jilqaː/	→ ˈjilqa			1	1		
	~ ˈjilqaː		1 W	L	1		
	~ jilˈqaː	1 W		L	1		
	~ ˈjilq			L	L	1 W	1 W
/kitaːb/	→ kiˈtaːb						
	~ ˈkitab			1 W			
/kaːtib/	→ ˈkaːtib						
	~ ˈkatib			1 W			
/huwwaː/	→ ˈhuwwa			1	1		
	~ ˈhuwwaː		1 W	L	1		
	~ huwˈwaː	1 W		L	1		
	~ ˈhuww			L	L	1 W	1 W
/kaːtaːb/	→ ˈkatab			2			
	~ kaˈtaːb			1 L			
	~ ˈkaːtab			1 L			
	~ ˈkaːtaːb		1 W	L			

The first two /kaːtaːb/ loser rows contain L but no W, so there is no ranking for BMCD to find. Other markedness constraints could be introduced to deal with this, but then they would affect /kitaːb/ and /kaːtib/.

BMCD succeeds in finding a grammar for lexicons (27a) and (27e). The support tableau for the latter is given in (29).

17. The constraint assessments in the support tableau (28) makes certain standard assumptions about Arabic prosody (see, e.g., Hayes 1995, where they go under the rubric of final consonant extrametricality): final superheavy syllables actually consist of a heavy syllable followed by a word appendix, which is sufficient to satisfy NonFin; and word-final consonants do not project a mora, so final CVC is light. The assessments also presuppose that Id(long) is vacuously satisfied when a long vowel is deleted, and that heavy syllables that precede the main stress satisfy WSP by virtue of secondary stress. None of these assumptions is essential to the main argument.

(29) Word-final (aː, a) free ride (= lexicon (27e))

lexicon	cands.	NonFin	WSP	Id(long)	Final-C	Max(Vː)	Max(V)
/jilqaː/	→ ˈjilqa			1	1		
	~ ˈjilqaː		1 W	L	1		
	~ jilˈqaː	1 W		L	1		
	~ ˈjilq			L	L	1 W	1 W
/kitaːb/	→ kiˈtaːb						
	~ ˈkitab			1 W			
/kaːtib/	→ ˈkaːtib						
	~ ˈkatib			1 W			
/huwwaː/	→ ˈhuwwa			1	1		
	~ ˈhuwwaː		1 W	L	1		
	~ huwˈwaː	1 W		L	1		
	~ ˈhuww			L	L	1 W	1 W
/katab/	→ ˈkatab						
	~ kaˈtaːb			1 W			
	~ ˈkaːtab			1 W			
	~ ˈkaːtaːb		1 W	2 W			

The key point is that (27e) assigns /katab/ as the underlying form of [ˈkatab]. BMCD finds the grammar ⟦NonFin, WSP » Max(Vː) » Final-C » Id(long), Max(V)⟧, with an r-measure of 8.[18]

When BMCD is applied to the support tableau for the lexicon with no free rides, (27a), it finds a different grammar, ⟦NonFin, WSP » Max(Vː) » Max(V) » Final-C » Id(long)⟧. This grammar has a lower r-measure, 7, and is in fact less restrictive, since it allows stems of the forbidden *[ˈjitba]~[ˈjitbani] type. The FRLA therefore favors the lexicon in (27e), in which only word-final [a] gets a free ride on (aː, a).

The general idea, then, is that the learner discovers a contextually limited free ride on the unfaithful map (A, B) by testing all of the logically possible combinations of (B, B) identity maps throughout the support tableau, relying on BMCD to determine which of these possibilities is linguistically meaningful. BMCD can do this because the random, linguistically meaningless combinations of free rides are not analyzeable with the given constraint-set, so BMCD quickly fails to find a grammar. In Arabic, the free ride is limited to word-final position because that is

18. Max(Vː) is ranked before Max(V) under the assumption that their ranking is fixed universally (cf. Beckman 1998). Furthermore, additional data involving syncope in other positions require this ranking (McCarthy 2004).

where the length contrast is neutralized, as a result of NONFIN and WSP. The learner discovers this by considering various lexicons in which the free ride on the shortening process occurs in word-final position and in other, sometimes random positions as well. Most of this space of possibilities is occupied by lexicons for which no grammar is possible.

The worst-case complexity of the FRLA as defined in section 3 is exponential in the size of $\mathfrak{L}$. The complexity of the system I have just sketched is exponential in the number of morphemes in the support tableau —a much more serious matter. The size of the support tableau (i.e., the number of winner~loser pairs) that is required for ranking is, in the worst case, quadratic in the number of constraints that must be ranked (Tesar and Smolensky 2000: 99). The overall complexity of this system is therefore proportional to 2^{N^2}, where N is the number of constraints in CON. This is deeply problematic, particularly since a reasonable value for N is probably an order of magnitude larger than a reasonable value for the cardinality of $\mathfrak{L}$.

It may prove possible to avoid this combinatorial explosion with an iterative procedure for lexicon building. Start with a support tableau that includes only the forms that are not eligible for the (A, B) free ride, either because they contain no [B]s or because any [B]s that they do contain are known from alternations to be derived from /A/s. (In the Arabic example above, this would mean starting with just /jilqaː/, /kitaːb/, and /kaːtib/, the three constants among the lexicons in (27).) Now, take each of the remaining [B]-containing forms one at a time and construct its possible underlying representations with and without the free ride. (In Arabic, these possibilities include /huwwaː/ and /huwwa/ for [ʰuwwa] and /kaːtaːb/, /kaːtab/, /kataːb/, and /katab/ for [ʰkatab].) Add each of these possibilities separately to the known, constant support tableau and apply BMCD, which will determine which possibilities work at all and whether any lead to a more restrictive grammar. The result will be the same as the computationally intractable algorithm, but the worst-case complexity is much more modest: it is exponential in the largest number of [B]s that any word in the support tableau happens to contain.

5. Conclusion

In morphophonemic learning, the underlying representations influence the grammar and the grammar influences the underlying representations. Constraint demotion already ensures that underlying representations will influence the grammar: constraints are ranked on the basis of the input → output maps in a support tableau. The remaining challenge for learning algorithms, then, is how to get information flow going in the other direction as well.

In this article, I have pursued the idea that learners simultaneously consider various hypotheses about underlying representation, rejecting any for which no grammar is possible and preferring the one that allows the most restrictive grammar. The inconsistency-detection property of multi-recursive constraint demotion (Tesar 1997) is crucial in eliminating the hypotheses for which there is no grammar, and the r-measure (Prince and Tesar 2004) is crucial in determining which grammar is the most restrictive.

Morphophonemic learning is a complex problem, and I have addressed only a small piece of it. The cases examined here involve free-ride effects: alternation data show that some [B]s derive from /A/s, and other evidence shows that even nonalternating [B]s must also derive from /A/s. In other words, the grammar maps /A/ to [B], and this information propagates back to the lexicon from nonalternating [B]s. The learner considers lexicons in which nonalternating [B]s are and are not derived from /A/s; if a more restrictive grammar can be found, then the former is chosen.

Morphophonemic learning necessarily intersects with phonological opacity (see the appendix), which presents its own learning challenges. It remains to be seen whether the proposal developed here can be modified and extended to address this notably thornier problem.

Appendix: Further Free-Ride Examples

1. Coalescence in Choctaw

In Choctaw (Lombardi and McCarthy 1991; Nicklas 1974, 1975; Ulrich 1986), preconsonantal vowel-nasal sequences coalesce to form a long nasalized vowel:

(30) Choctaw coalescence

/am-pala/	ãːpala	‘my lamp’	(cf. amissi ‘my deer’)
/am-tabi/	ãːtabi	‘my cane’	
/am-miŋko/	ãːmĩːko	‘my chief’	
/onsi/	õːsi	‘eagle’	

In a closed syllable, coalescence yields a short nasalized vowel (/ta-N-kči/ → [tãkči] ‘to tie’ with a nasal infix). This is an expected consequence of Choctaw’s syllable canons. Word-final nasals coalesce only if they are affixal; the resulting vowel is also short for prosodic reasons. Geminate nasals that are tautomorphemic or derived by assimilation do not coalesce: [onna] ‘dawn’, /ta-l-na-h/ → [tannah] ‘to be woven’. These are typical geminate inalterability effects (Guerssel 1977, 1978; Hayes 1986; Schein and Steriade 1986).

In Choctaw, the vowel inventory has a systematic gap: except in environments where long vowels are prohibited (closed syllables and word-finally), nasalized vowels are always long. In Ulrich’s (1986: 60ff.) analysis, this gap is explained by prohibiting underlying nasal vowels with a MSC and deriving all surface nasal vowels by coalescence of VN sequences. Even nonalternating morpheme-internal nasalized vowels are consistently long (e.g., ‘chief’ and ‘eagle’ in (30)), showing that they too are derived by coalescence. Furthermore, this analysis explains why Choctaw has no nasal vowels followed by a tautomorphemic nasal consonant (Ulrich 1986: 62): geminate inalterability protects a nasal consonant from coalescence when it is followed by a tautomorphemic nasal, so there is no way of deriving a sequence of a nasalized vowel and a tautomorphemic nasal.

In OT with richness of the base, it is not possible to ban nasalized vowels from inputs. The closest alternative is a chain-shift analysis similar to Gnanadesikan's account of Sanskrit: /…aNC…/ → […ãːC…] but /ã(ː)/ → [a(ː)]. In this way, input nasalized vowels are pushed out of the way and all surface nasalized vowels are derived by coalescence. If this analysis is adopted, then it too is an example of free-ride learning closely paralleling Sanskrit. On the basis of alternating morphemes like /am/, the learner discovers the unfaithful coalescent map. The free-ride algorithm permits the learner to extend this map to nonalternating morphemes like [õːsi], thereby accommodating the more restrictive grammar of a language that prohibits nasalized vowels that are short (in non-final open syllables) or that are followed by tautomorphemic nasals.

In the discussion of Sanskrit in section 2, I mentioned an alternative to the chain-shift analysis: introduce a markedness constraint against short mid vowels. A similar alternative might be proposed for Choctaw: if there is a markedness constraint against short nasalized vowels, could not it explain the same facts without the chain shift and consequently without free-ride learning? The problem with this approach is that it will not account for Ulrich's observation that nasalized vowels are never followed by tautomorphemic nasal consonants. Indeed, it is hard to imagine any reasonable markedness constraint that could give direct expression to this generalization.

2. Coalescence in Rotuman

De Haas (1988) cites Rotuman as an example of non-structure-preserving coalescence. It requires free-ride learning in much the same way as Sanskrit and Choctaw. The primary source on Rotuman is Churchward (1940). For a bibliography of the extensive secondary literature, see McCarthy (2000).

Rotuman distinguishes two «phases», complete and incomplete. The distribution of the two phases is based on prosodic structure; my focus here is on the way that the phase difference is realized in stems. A stem that is in the incomplete phase always ends in a stressed heavy syllable; heavy syllables are otherwise prohibited (except in monosyllables and recent loans). Stem-final /CV_1CV_2/ sequences are turned into heavy syllables in three different ways, depending of the quality of V_1 and V_2. If V_1 is a back rounded vowel and V_2 is front and no lower than V_1, they coalesce to form the front rounded counterpart of V_1:

(31) Coalescence in Rotuman

/hoti/	høt	'to embark'
/mose/	møs	'to sleep'
/futi/	fyt	'to pull'

The underlying forms show up unaltered in the complete phase of these stems: [ˈhoti], [ˈmose], [ˈfuti]. The result of coalescence is a short vowel because syllables are maximally bimoraic.

Front rounded vowels have no other source in Rotuman, and so they are only ever found in final closed syllables of morphemes in the incomplete phase.[19] For example, *[ˈhøti] and *[ˈmøse] are not possible words of this language. As in the chain-shift analyses of Sanskrit and Choctaw, all surface front rounded vowels will be derived by coalescence, while any input front rounded vowel maps unfaithfully to something else, such as its nonround counterpart. The learner must be able to grant a free ride to any front rounded vowels that occur in morphemes that he has only observed in the incomplete phase.

There is no plausible markedness alternative to the chain shift and the free ride. To explain the distribution of front rounded vowels without deriving all of them by coalescence, it would be necessary to posit a markedness constraint banning front rounded vowels from open syllables. Such a constraint enjoys no independent typological support and is purely descriptive.

3. Opacity and allophony in Japanese

Itô and Mester (2001: 280-281; 2003) identify a class of cases, which they dub «allophonic masking», that present problems for input-based theories of opacity such as sympathy theory. The interaction in Japanese between nasalization of voiced velar stops and *rendaku* voicing is a perfect example.

The voiced velar stop in initial position regularly alternates with its nasal counterpart in medial position: [ɡo] ‘the game of go’, [okiŋo] ‘handicap go’. As is well known, medial voiced obstruents block *rendaku* voicing of stem-initial obstruents: [kamikaze] ‘divine wind’, *[kamiɡaze]. In this respect, medial [ŋ] acts just like a voiced obstruent, even if it never alternates with [ɡ]: [sakatoŋe] ‘reverse thorn’, *[sakadoŋe]. (Additionally, [ɡ]s derived by *rendaku* also nasalize: [oriŋami] ‘paper folding’ from /ori-kami/).

In a traditional analysis, Japanese requires a MSC excluding /ŋ/ from underlying representations. The *rendaku* rule is ordered early enough to see the underlying /ɡ/ in /saka-toɡe/, and so voicing of the /t/ is correctly blocked. By a later rule, medial /ɡ/ becomes [ŋ].

OT has no MSCs, leading Itô and Mester to propose a stratal OT analysis that uses an early stratum to approximate the effect of the MSC. Although the rich base offers both /ɡ/ and /ŋ/ in both initial and medial positions, the phonology of the first stratum wipes out this potential distinction, mapping /ŋ/ to [ɡ] in all contexts. *Rendaku* also occurs in the first stratum, where it sees only [ɡ]. In the second stratum, medial [ɡ] that is the output of the first stratum becomes nasalized.

This example and others like it challenge approaches to opacity that rely on properties of the input, such as sympathy (McCarthy 1999), comparative markedness (McCarthy 2003a), or turbidity (Goldrick 2000; Goldrick and Smolensky 1999). If applied to Japanese, these approaches would refer to the /ɡ/ that underlies

19. An exception: front rounded vowels are also derived by harmony with a front rounded vowel that is itself derived by coalescence: /pulufi/ → [pylyf] ‘to adhere’.

surface [ŋ] to explain why there is no *rendaku* voicing in [sakatoŋe]. For example, one might assume in turbid fashion that the [+voice] feature of underlying /g/ remains present in surface structure, though unpronounced, and its presence is sufficient to block *rendaku*. The problem with this analysis is that there is no guarantee that surface [ŋ] derives from /g/; it might simply derive from /ŋ/, in which case *rendaku* would not be blocked. In fact, the analysis wrongly predicts a contrast between medial [ŋ]s that do and do not block *rendaku*, the former derived from /g/ and the latter from /ŋ/. There is no such contrast, however. The MSC-based analysis and Itô and Mester's stratal approach do not predict this contrast because all surface [g]s and [ŋ]s derive from earlier [g]s.

Gnanadesikan's analysis of Sanskrit suggests a different way of looking at Japanese, however. Suppose that all surface [ŋ]s are derived from underlying /g/s and the grammar maps input /ŋ/ to something other than [g], such as [n]. This is a chain shift: /g/ → [ŋ] (medially) and /ŋ/ → [n] (everywhere). Abstractly, the learning problem presented by this chain shift is the same as we saw with Sankrit.

For concreteness, I will borrow several markedness constraints from Itô and Mester's (2003) analysis.

(32) a. *VgV
Intervocalic *g* is prohibited (a lenition constraint).

b. *g
Voiced velar stops are prohibited (cf. Ohala 1983: 196-197).

c. *ŋ
Velar nasals are prohibited.

We will look at their interaction among themselves and with two faithfulness constraints, IDENT(Nasal) and IDENT(Place).

For the phonotactic learner of Japanese, the support tableau includes observed initial [g]s and medial [ŋ]s, both of which are derived by the identity map.

(33) Support tableau at phonotactic stage

lexicon	cands.	*VgV	*g	*ŋ	ID(Nas)	ID(Place)
/ga/	→ ga		1			
	~ da		L			$_1$ W
	~ ŋa		L	$_1$ W	$_1$ W	
	~ na		L		$_1$ W	$_1$ W
/aŋa/	→ aŋa			1		
	~ aga	$_1$ W	$_1$ W	L	$_1$ W	
	~ ana			L		$_1$ W

On the first pass through BMCD, *VgV is identified as the only rankable markedness constraint. On the next pass, we seek a faithfulness constraint that can be ranked so as to free up one of the remaining markedness constraints for ranking. ID(Place) frees up *ŋ, which accounts for the remaining forms. The residual constraints, *g and ID(Nas) are ranked so as to conform to the markedness-first bias. The resulting ranking is given in (34).

(34) Result of phonotactic learning for Japanese
*VgV » ID(Pl) » *ŋ » *g » ID(Nas)

As the learner grows aware of data from alternations, he learns that there are unfaithful maps /g/ → [ŋ] that occur when stem-initial /g/ becomes medial as a result of the morphology. This unfaithful map is added to the support tableau and BMCD is applied once again.

(35) Support tableau after some morphophonemic learning

lexicon	cands.	*VgV	*g	*ŋ	ID(Nas)	ID(Place)
/ga/	→ ga		1			
	~ da		L			$_1$ W
	~ ŋa		L	$_1$ W	$_1$ W	
	~ na		L		$_1$ W	$_1$ W
/aŋa/	→ aŋa			1		
	~ aga	$_1$ W	$_1$ W	L	$_1$ W	
	~ ana			L		$_1$ W
/a-ga/	→ aŋa			1	1	
	~ aga	$_1$ W	$_1$ W	L	L	
	~ ana			L	$_1$	$_1$ W

BMCD gives this support tableau the same ranking, (34). The commitment to the identity map for nonalternating forms means that [aŋa] must derive from /aŋa/—and the presence of the (ŋ, ŋ) identity map in the support tableau forces the ranking ID(Pl) » *ŋ. The proposed chain shift, where underlying /ŋ/ maps to [n], is apparently not learnable under these assumptions about the input. It is learnable with the free-ride algorithm, however, since its basic character is similar to Sanskrit.

In discussing Sanskrit, I noted that a local solution to the learning problem was possible by adding another markedness constraint. No such move will solve the problem in Japanese, however. Surface [ŋ]s all act like they are derived from underlying /g/s with respect to blocking *rendaku*. Fiddling with the constraints will not somehow magically subvert the commitment to the identity map.

References

Alderete, John; Tesar, Bruce (2002). «Learning Covert Phonological Interaction: An Analysis of the Problem Posed by the Interaction of Stress and Epenthesis». Report no. RuCCS-TR-72. New Brunswick, NJ: Rutgers University Center for Cognitive Science. [Available on Rutgers Optimality Archive #543, http://roa.rutgers.edu/.]

Angluin, Dana (1980). «Inductive Inference of Formal Languages from Positive Data». *Information and Control* 45: 117-135.

Baker, Carl Lee (1979). «Syntactic Theory and the Projection Problem». *Linguistic Inquiry* 10: 533-581.

Beckman, Jill N. (1998). *Positional Faithfulness*. University of Massachusetts, Amherst, doctoral dissertation. [Available on Rutgers Optimality Archive #234, http://roa.rutgers.edu/. Published as *Positional Faithfulness: An Optimality Theoretic Treatment of Phonological Asymmetries* in 1999 by Garland.]

Bermúdez-Otero, Ricardo (2004). «The Acquisition of Phonological Opacity». University of Newcastle upon Tyne, unpublished manuscript. [Available on Rutgers Optimality Archive #593, http://roa.rutgers.edu/.]

Churchward, Clerk Maxwell (1940). *Rotuman Grammar and Dictionary*. Sydney: Australasia Medical Publishing Co. [Reprint 1978, AMS Press, New York.]

de Haas, Wim (1988). *A Formal Theory of Vowel Coalescence: A Case Study of Ancient Greek*. Dordrecht: Foris.

Dinnsen, Daniel A. (2004). «A Typology of Opacity Effects in Acquisition». Indiana University, Bloomington, IN, unpublished manuscript.

Dinnsen, Daniel A.; Barlow, Jessica A. (1998). «On the Characterization of a Chain Shift in Normal and Delayed Phonological Acquisition». *Journal of Child Language* 25: 61-94.

Dinnsen, Daniel A.; McGarrity, Laura W. (2004). «On the Nature of Alternations in Phonological Acquisition». *Studies in Phonetics, Phonology, and Morphology* 11: 23-42. [Available at http://www.indiana.edu/~sndlrng/DinnsenMcGarrity%2004.pdf.]

Dinnsen, Daniel A.; O'Connor, Kathleen; Gierut, Judith (2001). «An Optimality Theoretic Solution to the Puzzle-Puddle-Pickle Problem». Handout of paper presented at the 75th Annual Meeting of the Linguistic Society of America, Washington, DC.

Dresher, B. Elan (1981). «On the Learnability of Abstract Phonology». In: Baker, Carl Lee; McCarthy, John (eds.). *The Logical Problem of Language Acquisition*. Cambridge, MA: MIT Press, pp. 188-210.

Gnanadesikan, Amalia (1997). *Phonology with Ternary Scales*. University of Massachusetts, Amherst, doctoral dissertation. [Available on Rutgers Optimality Archive #195, http://roa.rutgers.edu/.]

Goldrick, Matthew (2000). «Turbid Output Representations and the Unity of Opacity». In: Hirotani, Masako (ed.). *Proceedings of the North East Linguistics Society 30*. Amherst, MA: Graduate Linguistic Student Association, pp. 231-246.

Goldrick, Matthew; Smolensky, Paul (1999). «Opacity, Turbid Representations, and Output-Based Explanation». Paper presented at the Workshop on the Lexicon in Phonetics and Phonology, University of Alberta, Edmonton.

Guerssel, Mohammed (1977). «Constraints on Phonological Rules». *Language* 3: 267-305.

Guerssel, Mohammed (1978). «A Condition on Assimilation Rules». *Linguistic Analysis* 4: 225-254.

Hayes, Bruce (1986). «Inalterability in CV Phonology». *Language* 62: 321-351.

Hayes, Bruce (1995). *Metrical Stress Theory: Principles and Case Studies*. Chicago: The University of Chicago Press.

Hayes, Bruce (2004). «Phonological Acquisition in Optimality Theory: The Early Stages». In: Kager, René; Pater, Joe; Zonneveld, Wim (eds.). *Fixing Priorities: Constraints in Phonological Acquisition*. Cambridge: Cambridge University Press, pp. 158-203. [Available on Rutgers Optimality Archive #327, http://roa.rutgers.edu/.]

Itô, Junko; Mester, Armin (1999). «The Phonological Lexicon». In: Tsujimura, Natsuko (ed.). *The Handbook of Japanese Linguistics*. Oxford: Blackwell, pp. 62-100.

Itô, Junko; Mester, Armin (2001). «Structure Preservation and Stratal Opacity in German». In: Lombardi, Linda (ed.). *Segmental Phonology in Optimality Theory*. Cambridge: Cambridge University Press, pp. 261-295.

Itô, Junko; Mester, Armin (2003). «Lexical and Postlexical Phonology in Optimality Theory: Evidence from Japanese». In: Fanselow, Gisbert; Féry, Caroline (eds.). *Resolving Conflicts in Grammars: Optimality Theory in Syntax, Morphology, and Phonology. Linguistische Berichte, Sonderheft 11*. Hamburg: Helmut Buske, pp. 183-207. [Available at http://people.ucsc.edu/~ito/PAPERS/lexpostlex.pdf.]

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

Kiparsky, Paul (1973). «Phonological Representations». In: Fujimura, Osamu (ed.). *Three Dimensions of Linguistic Theory*. Tokyo: TEC, pp. 3-136.

Lombardi, Linda; McCarthy, John (1991). «Prosodic Circumscription in Choctaw Morphology». *Phonology* 8: 37-71.

McCarthy, John J. (1999). «Sympathy and Phonological Opacity». *Phonology* 16: 331-399.

McCarthy, John J. (2000). «The Prosody of Phase in Rotuman». *Natural Language and Linguistic Theory* 18: 147-197.

McCarthy, John J. (2003a). «Comparative Markedness». *Theoretical Linguistics* 29: 1-51.

McCarthy, John J. (2003b). «Sympathy, Cumulativity, and the Duke-of-York Gambit». In: Féry, Caroline; van de Vijver, Ruben (eds.). *The Syllable in Optimality Theory*. Cambridge: Cambridge University Press, pp. 23-76.

McCarthy, John J. (2004). «The Length of Stem-Final Vowels in Colloquial Arabic». University of Massachusetts, Amherst, unpublished manuscript. [Available on Rutgers Optimality Archive #616, http://roa.rutgers.edu/. To appear in *Perspectives on Arabic Linguistics 18*, ed. by Dilworth B. Parkinson, John Benjamins, Amsterdam & Philadelphia.]

McCarthy, John J.; Prince, Alan (1993). «Generalized Alignment». In: Booij, Geert; Marle, Jaap van (eds.). *Yearbook of Morphology*. Dordrecht: Kluwer, pp. 79-153. [Available on Rutgers Optimality Archive #7, http://roa.rutgers.edu/.]

McCarthy, John J.; Prince, Alan (1995). «Faithfulness and Reduplicative Identity». In: Beckman, Jill; Walsh Dickey, Laura; Urbanczyk, Suzanne (eds.). *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory*. Amherst, MA: Graduate Linguistic Student Association, pp. 249-384. [Available on Rutgers Optimality Archive #60, http://roa.rutgers.edu/.]

McCarthy, John J.; Prince, Alan (1999). «Faithfulness and Identity in Prosodic Morphology». In: Kager, René; van der Hulst, Harry; Zonneveld, Wim (eds.). *The*

Prosody-Morphology Interface. Cambridge: Cambridge University Press, pp. 218-309. [Available on Rutgers Optimality Archive #216, http://roa.rutgers.edu/.]

McCarthy, John J.; Wolf, Matthew (2005). «Less Than Zero: Correspondence and the Null Output». University of Massachusetts, Amherst, unpublished manuscript. [Available on Rutgers Optimality Archive #722, http://roa.rutgers.edu/.]

Moreton, Elliott (2000). «Faithfulness and Potential». University of Massachusetts, Amherst, unpublished manuscript.

Nicklas, Thurston (1974). *The Elements of Choctaw*. University of Michigan, doctoral dissertation.

Nicklas, Thurston (1975). «Choctaw Morphophonemics». In: Crawford, James M. (ed.). *Studies in Southeastern Indian Languages*. Athens, GA: University of Georgia Press, pp. 237-250.

Ohala, John (1983). «The Origin of Sound Patterns in Vocal Tract Constraints». In: MacNeilage, Peter (ed.). *The Production of Speech*. New York: Springer-Verlag, pp. 189-216.

Ota, Mitsuhiko (2004). «The Learnability of the Stratified Phonological Lexicon». *Journal of Japanese Linguistics* 20: 19-40. [Available on Rutgers Optimality Archive #668, http://roa.rutgers.edu/.]

Pater, Joe (2004). «Course Handout». University of Massachusetts, Amherst, unpublished manuscript.

Pater, Joe; Tessier, Anne-Michelle (2003). «Phonotactic Knowledge and the Acquisition of Alternations». In: Solé, Maria-Josep; Recasens, Daniel; Romero, Joaquín (eds.). *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona: Universitat Autònoma de Barcelona, pp. 1177-1180. [Available at http://people.umass.edu/pater/pater-tessier.pdf.]

Prince, Alan (2002). «Arguing Optimality». Rutgers University, New Brunswick, NJ, unpublished manuscript. [Available on Rutgers Optimality Archive #562, http://roa.rutgers.edu/.]

Prince, Alan; Smolensky, Paul (2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden & Oxford: Blackwell. [Revision of 1993 technical report no. RuCCS-TR-2, Rutgers University Center for Cognitive Science, New Brunswick, NJ. Available on Rutgers Optimality Archive #537, http://roa.rutgers.edu/.]

Prince, Alan; Tesar, Bruce (2004). «Learning Phonotactic Distributions». In: Kager, René; Pater, Joe; Zonneveld, Wim (eds.). *Fixing Priorities: Constraints in Phonological Acquisition*. Cambridge: Cambridge University Press, pp. 245-291. [Available on Rutgers Optimality Archive #353, http://roa.rutgers.edu/.]

Schane, Sanford (1987). «The Resolution of Hiatus». In: Bosch, Anna; Need, Barbara; Schiller, Eric (eds.). *Papers from the 23rd Annual Regional Meeting of the Chicago Linguistic Society, Vol. 2: Parasession on Autosegmental and Metrical Phonology*. Chicago: Chicago Linguistic Society, pp. 279-290.

Schein, Barry; Steriade, Donca (1986). «On Geminates». *Linguistic Inquiry* 17: 691-744.

Tesar, Bruce (1997). «Multi-Recursive Constraint Demotion». Rutgers University, New Brunswick, NJ, unpublished manuscript. [Available on Rutgers Optimality Archive #197, http://roa.rutgers.edu/.]

Tesar, Bruce (1998). «Error-Driven Learning in Optimality Theory Via the Efficient Computation of Optimal Forms». In: Barbosa, Pilar; Fox, Danny; Hagstrom, Paul;

McGinnis, Martha; Pesetsky, David (eds.). *Is the Best Good Enough? Optimality and Competition in Syntax*. Cambridge, MA: MIT Press, pp. 421-435.

Tesar, Bruce; Alderete, John; Horwood, Graham; Merchant, Nazarré; Nishitani, Koichi; Prince, Alan (2003). «Surgery in Language Learning». In: Garding, Gina; Tsujimura, Mimu (eds.). *Proceedings of the 22nd West Coast Conference on Formal Linguistics*. Somerville, MA: Cascadilla Press, pp. 477-490.

Tesar, Bruce; Prince, Alan (forthcoming). «Using Phonotactics to Learn Phonological Alternations». In: Cihlar, Jonathan E.; Franklin, Amy L.; Kaiser, David W.; Kimbara, Irene (eds.). *Papers from the 39th Regional Meeting of the Chicago Linguistics Society, Vol. 2: The Panels*. Chicago: Chicago Linguistic Society. [Available on Rutgers Optimality Archive #620, http://roa.rutgers.edu/.]

Tesar, Bruce; Smolensky, Paul (1994). «The Learnability of Optimality Theory». In: Aranovich, Raul; Byrne, William; Preuss, Susanne; Senturia, Martha (eds.). *Proceedings of the Thirteenth West Coast Conference on Formal Linguistics*. Stanford, CA: CSLI Publications, pp. 122-137.

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

Tessier, Anne-Michelle (2004). «Looking for OO-Faithfulness in Phonological Acquisition». University of Massachusetts, Amherst, unpublished manuscript.

Ulrich, Charles (1986). *Choctaw Morphophonology*. UCLA, doctoral dissertation.

Whitney, William Dwight (1889). *Sanskrit Grammar*. Cambridge, MA: Harvard University Press.

Winter, Edgar (1972). *Free Ride*. Epic Records.

Zwicky, Arnold M. (1970). «The Free-Ride Principle and Two Rules of Complete Assimilation in English». In: Campbell, M. A. *et al.* (eds.). *Papers from the Sixth Regional Meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society, pp. 579-588.