A Learning-Based Account of L1 vs. L2 Cluster Repair Differences

Karen Jesney

1. Introduction

Children acquiring a first language and adults acquiring a second language both encounter marked phonological structures that exceed the production abilities of their developing grammars. Although these marked structures may be the same, child L1 learners and adult L2 learners often differ in their preferred means of avoiding them. Onset consonant clusters provide a case in point: while child L1 learners typically repair these structures through deletion (e.g., Dyson & Paden 1983, McLeod et al. 2001, Preisser et al. 1988), adult L2 learners typically make use of epenthesis (e.g., Abrahamsson 1999, Carlisle 1991, 1997, 2006, Eckman 1981, Lin 2001, Weinberger 1987, 1994).

Past accounts of these differences between the learner populations have generally relied on a greater responsiveness to functional communicative concerns among adult learners. Weinberger's (1987, 1994) Recoverability Principle, for example, attributes the increased rate of epenthesis among adult L2 learners to a desire to preserve all of the input segments, thereby facilitating listeners' lexical retrieval. Under this account, the distinct behaviour of L1 and L2 learners is effectively attributed to the age and linguistic development of the learners, rather than to the structure of the grammar or the effects of the L1 on the L2 learning process.

This paper argues for an account based on transfer of covert elements of the L1 grammar to the L2 initial state. In particular, I show that a bias toward different cluster repair strategies emerges automatically when an Optimality-Theoretic (OT; McCarthy & Prince 1995, Prince & Smolensky 1993/2004) grammar is learned using the Gradual Learning Algorithm (GLA; Boersma 1998, Boersma & Hayes 2001). Errors during the L1 learning process give rise to a covert MAX-C >> DEP-V ranking; when this ranking is transferred as part of the initial state for L2 learning, a preference for epenthesis repairs results. These effects are robust across learning simulations, illustrating how the process of L1 acquisition can ultimately affect L2 learning.

The rest of this paper is structured as follows. Section 2 presents the basic OT account of deletion and epenthesis repairs and briefly summarizes some evidence regarding differences in L1 vs. L2 repair preferences. Section 3 introduces the GLA. Section 4 presents two toy languages and shows how different patterns emerge when these are learned as an L1 vs. as an L2. Section 5 discusses some of the predictions that arise from these simulations, and considers how additional factors might influence the results. Section 6 concludes.

2. Onset cluster repair preferences

For children acquiring an L1, onset clusters are inevitably a novel structure. For adults acquiring an L2, the status of onset clusters is more variable. In at least some cases, however, the challenge is similar to that of the L1-acquiring child. A Hawaiian learner of L2 English, for instance, must transition from a grammar that allows no onset clusters at all to a grammar that admits clusters of up to three consonants.

---

* Karen Jesney, University of Southern California, jesney@usc.edu. Thanks to Heather Goad, Joe Pater, Anne-Michelle Tessier, Rachel Walker, Matt Wolf, and the audiences at GALANA 5 and USC, for feedback on various instantiations of this work. Thank you also to the two anonymous reviewers for their helpful comments on this paper.

 Selected Proceedings of the 5th Conference on Generative Approaches to Language Acquisition North America (GALANA 2012), ed. Chia-Ying Chu et al., 10-21. Somerville, MA: Cascadilla Proceedings Project.

Deletion and epenthesis are the most common strategies for avoiding consonant clusters. In Optimality Theory these two repairs are militated against by the constraints MAX-C and DEP-V, respectively. As the tableaux in (1) illustrate, if DEP-V dominates MAX-C, complex onsets are avoided through deletion; if the opposite ranking of the faithfulness constraints holds, complex onsets are avoided through epenthesis.

(1) a. Deletion is the preferred repair for complex onsets when DEP-V >> MAX-C

/blu/	*COMPLEXONSET	DEP-V	MAX-C
blu	*!		
☞ bu			*
bəlu		*!	

b. Epenthesis is the preferred repair for complex onsets when MAX-C >> DEP-V

/blu/	*COMPLEXONSET	MAX-C	DEP-V
blu	*!		
bu		*!	
☞ bəlu			*

Neither L1 nor L2 learners show an absolute preference for one type of repair over the other. The strong generalization emerging from the literature, however, is that L1 learners tend to display the pattern in (1a), while the pattern in (1b) is more common among L2 learners. In a study of L1 English-acquiring children (*N* = 20), Preisser et al. (1988) found that 93% of target clusters were reduced via deletion among children age 1;6-1;9; only 7% of target clusters were either realized as clusters or repaired via epenthesis. This preference for deletion persisted among older children in the study – children age 1;10-2;1 (*N* = 20) showed deletion in 76% of target clusters and children age 2;2-2;5 (*N* = 20) showed deletion in 51% of target clusters. Similar patterns are evident among the twelve Dutch-acquiring children of the CLPF database (Fikkert 1994, Levelt 1994). There, 93.5% of the repairs serving to eliminate clusters in word-initial stressed onsets involve deletion rather than epenthesis. Rates of repair via epenthesis are consistently low in the Dutch data, ranging from 0.99% to 13.69% across children (Jesney 2011).

The situation for L2 learners is more complex, given the varied backgrounds of the individuals in question. Still, higher rates of epenthesis are consistently reported for L2 learners than for L1-acquiring children. In a series of studies, Carlisle (1991, 1997, 1998) reports that Spanish-speaking learners of English make almost exclusive use of epenthesis in repairing /sC(C)/ onset clusters. Similar results have been reported for Portuguese-speaking learners of English (Rauber 2006). Epenthesis has also been shown to be the dominant repair strategy among L2 learners from L1 backgrounds with more restricted onset cluster inventories – e.g., L1 Korean learners of English (Kim 2009, Kwan 2005). Among Mandarin learners of English, epenthesis rates of up to 100% have been reported (Eckman 1981). Lin (2001) found a more moderate level of epenthesis in a study of 20 L1-Mandarin/Taiwanese learners of English; still, across the four tasks in Lin's study, onset clusters were repaired via epenthesis roughly 2.4 times as often as via deletion. This preference for epenthesis was strongest in the minimal-pair task, where 35.13% of clusters were repaired via epenthesis and only 3.38% of clusters were repaired via deletion.

There are, of course, a multitude of factors that influence the repairs selected by both L1 and L2 learners. These include language-internal grammatical factors such as the sonority profile of the cluster (e.g., Abrahamsson 1999, Carlisle 2006, Tropf 1987) and the surrounding phonological context (e.g., Carlisle 1991, 1997, Rauber 2006), and performance factors such as the relative formality of the task (e.g., Lin 2001). The effects of some of these factors are discussed further in §5. The primary finding of this paper, though, is that a bias in favour of higher rates of epenthesis among L2 learners can arise automatically as a result of the learning process itself, without reference to any of these additional factors. This bias emerges in the absence of any direct evidence favouring deletion or epenthesis in either the L1 or the L2. The L2 preference for epenthesis is due to the transfer of a *covert ranking* – i.e., a ranking that arises due to the dynamics of the learning process but that is not needed in order to capture the phonotactic patterns of either the L1 or L2 target language. Additional factors can modulate and refine this bias, but, all else being equal, epenthesis is expected to predominate among L2 learners.

3. Learning with the GLA

The Gradual Learning Algorithm makes use of a version of Optimality Theory where constraints are assigned numeric ranking values arrayed along a linear scale (Boersma 1998, Boersma & Hayes 2001). In most implementations, including those in this paper, a small amount of noise is added to each constraint's value on each iteration of EVAL. The resulting selection values are then transformed into a strict ranking such that constraints with higher values dominate constraints with lower values. Adding noise to the ranking values in this manner results in occasional reversals in the basic priority relations among the constraints, allowing patterns of variation to be effectively captured.

Learning within this model involves adjusting the ranking values of constraints when mismatches between the target language and the output of the learner's current grammar are detected. The procedure is targeted; only the ranking values of constraints that assign different marks to the learner's error and to the target form are adjusted. The precise procedure is given in (2). The terms 'winner' and 'loser' are drawn from the OT learning work of Tesar & Smolensky (1998, 2000), and refer, respectively, to the target form, which will be the winning (optimal) candidate in the target grammar, and the learner's current optimum, which will be a losing candidate in the target grammar.

(2) Gradual Learning Algorithm Procedure

- The learner selects an input form and a corresponding 'winner' output form *W* from the ambient language
- The learner takes the input and feeds it to its grammar to determine which output candidate its current grammar deems optimal.
- If the output candidate selected by the current grammar is identical to the observed output *W*, the grammar remains unchanged.
- If the output candidate selected by the current grammar is not identical to *W* (i.e., there is an error), the learner's output is deemed a 'loser' *L*, and the grammar is adjusted by:
    - adding some small value *r* (the plasticity value) to the ranking value of each winner-favouring constraint, and
    - subtracting the same small value *r* from the ranking value of each loser-favouring constraint.

The GLA procedure can be applied given any set of initial ranking values for the constraints. However, an initial state that assigns higher ranking values to Markedness constraints than to Faithfulness constraints allows the patterns of child language to be better modeled and restrictive final grammars to be more consistently attained (for discussion of MARKEDNESS >> FAITHFULNESS biases in OT learning, see, among others, Demuth 1995, Gnanadesikan 2004, Pater 1997, Smolensky 1996). The L1 learning simulations discussed here all begin with the ranking values of Markedness constraints set at 100 and the ranking values of Faithfulness constraints set at 0. Taken together with the procedure in (2), this bias means that learning gradually decreases the ranking values of Markedness constraints and gradually increases the ranking values of Faithfulness constraints. This process continues until the learner's productions consistently mirror the target grammar's outputs – i.e., until the learner no longer makes errors and learning procedure is said to have converged.

As an initial illustration, we can consider the tableau in (3). With the markedness constraint *COMPLEXONSET dominating MAX-C and DEP-V, the grammar can deem either of the two error forms [bu] or [bəlu] optimal, depending upon the selection values for the two faithfulness constraints once noise is applied.

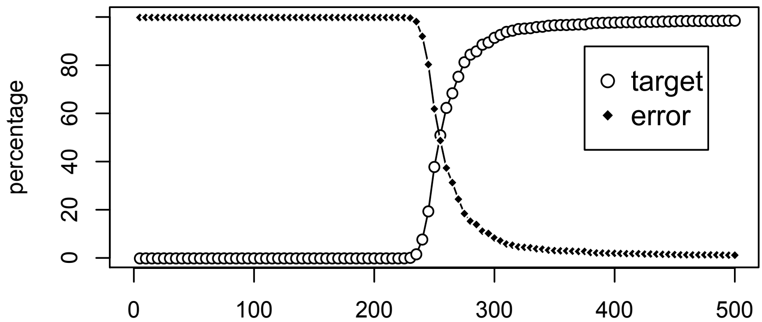
(3) Target and error forms given an initial MARKEDNESS >> FAITHFULNESS ranking

/blu/	*COMPLEXONSET	MAX-C	DEP-V
Target (W) blu	* →		
Error (L) bu		← *	
Error (L) bəlu			← *

Following the procedure in (2), if [bu] is the selected loser, the ranking value of *COMPLEXONSET will decrease and the ranking value of MAX-C will increase. If [bəlu] is the selected loser, the ranking value of *COMPLEXONSET will decrease and the ranking value of DEP-V will increase.

The figure in (4) plots the mean results of ten GLA simulations conducted in Praat (Boersma & Weenink 2012) illustrating the learning path for input /blu/ over 500 learning trials. In this case, with the plasticity set to 0.2 and only a single input form being considered, target-like production begins to overtake the two error forms after approximately 255 learning trials.[1]

(4) Mean percentage of target vs. error mappings of input /blu/ over 500 GLA learning trials



As the learning scenarios become more complex – i.e., as more inputs, outputs, and constraints are added – more detailed predictions about the course of development emerge. As the following section shows, when marked segments are learned alongside marked syllable structures, this leads to asymmetries in the ranking of MAX-C and DEP-V toward the end of L1 learning. Transferred to a second language, these asymmetries provide an account of L1 vs. L2 cluster repair preferences.

4. Emergent differences in L1 vs. L2 repair preferences

4.1. Toy languages

To illustrate the differences in repair preferences that arise in the process of L1 vs. L2 learning, we will consider the two toy languages sketched in (5). The 'Simple Language' has one source of syllable structure markedness – codas – and one source of segmental markedness – voiced obstruents. The 'Cluster Language' is a proper superset of the Simple Language, adding onset clusters as a second source of syllable structure markedness.

(5)

Simple Language		Cluster Language	
[ba]	voiced obstruent	[ba]	voiced obstruent
[pat]	coda	[pat]	coda
[tab]	voiced obstruent & coda	[tab]	voiced obstruent & coda
		[pla]	onset cluster
		[bla]	voiced obstruent & onset cluster

Seven constraints are included in the simulations in order to capture the various sources of markedness and the associated repair possibilities. In addition to the *COMPLEXONSET, MAX-C and DEP-V constraints already discussed, these simulations include NOCODA, which penalizes any syllable ending with a consonant, *VOICEDOBSTRUENT, which penalizes output voiced obstruents, and IDENT[±voice], which penalizes output segments that differ in specification for [±voice] relative to their input correspondents. The constraint ONSET, which penalizes syllables that do not begin with a

[1] All simulations in this paper make use of the "Symmetric All" update rule, which replicates the GLA procedure given in (2). Results were sampled by drawing 100,000 input-output pairs for each input after every five learning trials. The mean results of the ten simulations were computed and plotted using R (R Development Core Team 2012). Additional parameter settings (all simulations): Evaluation noise = 2.0, Plasticity decrement = 1, Number of plasticities = 1, Plasticity spreading = 0, Number of chews = 1.

consonant, is also included; this has the effect of ensuring that the input /ba/ acts primarily as evidence for promotion of IDENT[±voice] rather than as additional evidence for promotion of MAX-C. The table in (6) summarizes the conflicting markedness and faithfulness constraints, along with the specific error mappings and inputs that motivate changes in their ranking values.

(6)

Markedness	**Conflicting Faithfulness**	**Error Mappings**	**Relevant Inputs**
NOCODA	MAX-C DEP-V	CVC → CV CVC → CVCə	/pat/, /tab/
*COMPLEXONSET	MAX-C DEP-V	CCV → CV CCV → CəCV	/pla/, /bla/
*VOICEDOBSTRUENT	MAX-C IDENT[±voice]	b → Ø b → p	/ba/, /tab/, /bla/

Importantly, neither of the target languages includes any alternations that provide evidence for a specific ranking between MAX-C and DEP-V. Mastery of the Simple Language requires that MAX-C and DEP-V dominate NOCODA and *VOICEDOBSTRUENT in order to ensure that inputs /pat/, /tab/ and /ba/ map faithfully. The Cluster Language makes the same requirement, and also demands that MAX-C and DEP-V dominate *COMPLEXONSET. Neither target language provides any overt evidence, however, for a MAX-C >> DEP-V or DEP-V >> MAX-C ranking.

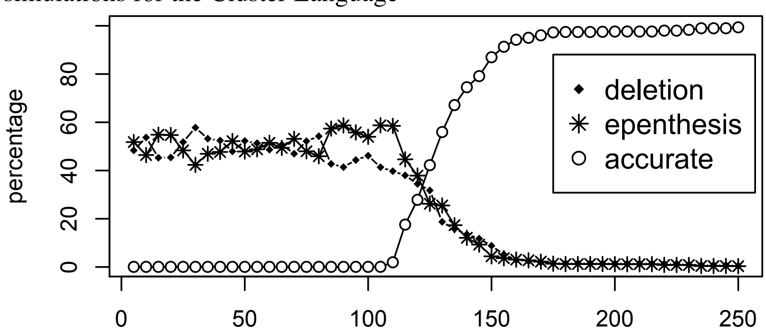
Because we are interested here in the patterns of repairs affecting onset clusters, it is the realization of inputs /pla/ and /bla/ in the Cluster Language that will be the primary focus of our discussion in the rest of this section.

4.2. L1 learning of the Cluster Language

The L1 learning scenario is based on learning the Cluster Language from an initial state like that described in §3 – i.e., an initial ranking value of 100 for all Markedness constraints and an initial ranking value of 0 for all Faithfulness constraints. The learner selects input data from an equal distribution of the five forms of the language – /ba/, /pat/, /tab/, /pla/ and /bla/ – and adjusts the ranking values using a plasticity of 1.0.

The figure in (7) shows the mean pattern of realization for the onset cluster inputs /pla/ and /bla/ over 250 learning trials based on 10 simulations.

(7) Realization of onset clusters over 250 GLA learning trials – mean results of 10 L1-learning simulations for the Cluster Language



As (7) suggests, deletion and epenthesis repairs are equally represented for inputs with onset clusters in the L1 simulations. During the first one hundred learning trials, before accurate realization of /pl/ or /bl/ becomes possible in any of the simulations, the clusters are repaired via deletion in 49.99% of cases and via epenthesis in 50.01% of cases. This pattern is reflected across all ten simulations; the mean rate of deletion repairs ranges from 45.05% to 57.26% during this time period. For the ten simulations, the rate of deletion vs. epenthesis repairs is essentially at chance – one-sample t-test: $t(9) = 0.01$, $p = .99$. A similar pattern is observed with the second source of syllable-structure markedness –

i.e., codas. Codas are repaired via deletion in 50.68% of cases and via epenthesis in 49.32% of cases – one-sample t-test: $t(9) = 0.38$, $p = .71$. The stages that the L1 learner passes through are summarized in (8) below.

(8) Stages in L1 learning of the Cluster Language[2]

L1 learning (target = Cluster Language)	**Stage 1**	/CCV/ → [CV], [CəCV] /CVC/ → [CV], [CVCə] *equal probability of deletion and epenthesis repairs*
	Stage 2	/CCV/ → [CCV] /CVC/ → [CVC] *faithful mapping of target onset clusters and codas*

A small difference in repair preferences begins to emerge toward the end of Stage 1. As the figure in (7) indicates, epenthesis repairs predominate starting after approximately 85 learning trials. While the rising accuracy rate means that this effect does not reach significance here, it this type of shift that, as we will see in §4.3, is responsible for the epenthesis repair preference observed among L2 learners.

4.3. L2 learning of the Cluster Language

In order to model L2 learning of the Cluster Language, the simulations begin with the learning of a first language that excludes onset clusters – i.e., the Simple Language. The end-state ranking values from the L1 learning stage are then transferred, and serve as the initial state for L2 learning of the Cluster Language. The assumption here, then, is that L2 learning is characterized by full access to the set of constraints and representations provided by UG, and full transfer of the L1 grammar – i.e., the end-state L1 constraint ranking (see Schwartz & Sprouse 1996; in OT learning, see Broselow 2004, Lombardi 2003).

The initial state ranking for L1 learning of the Simple Language is given in (9); this is the same ranking as that employed in §4.2, reflecting the assumed universality of the L1 initial state.

(9) L1 initial state ranking
ONSET, *COMPLEXONSET, *VOICEDOBS, NOCODA >> IDENT[±voice], MAX-C, DEP-V

While the L1 learner of the Simple Language is never exposed to onset clusters, the pattern of repairs for codas is very similar to that seen with L1 learning of the Cluster Language in §4.2. After a period of variability during which deletion and epenthesis are attested in equal measure, epenthesis briefly predominates before accurate realization of the marked structure takes over. At the end of L1 learning of the Simple Language – i.e., after 250 learning trials – the learner consistently maps inputs /ba/, /pat/ and /tab/ faithfully.

From the analyst's perspective, faithful mapping of these inputs provides direct evidence for the rankings in (10). Because the adult language shows no evidence for either deletion or epenthesis alternations, there is no motivation for positing a particular ranking between MAX-C and DEP-V.

(10) Rankings required for L1 mastery of the Simple Language (analyst's perspective)
a. Codas are allowed: MAX-C, DEP-V >> NOCODA
b. Voiced obstruents are allowed: MAX-C, IDENT[±voice] >> *VOICEDOBSTRUENT
c. Onsets are required: ONSET >> *VOICEDOBSTRUENT
d. Onset clusters are not allowed: *COMPLEXONSET >> MAX-C, DEP-V

In fact, the dynamics of the learning algorithm mean that the rankings that exist at the end of the L1-learning stage are more articulated than the adult language suggests. This is illustrated in (11) with the

[2] Different rates of acquisition for onset clusters and codas can be modeled by manipulating the relative frequency of the two types of forms in the input (see, e.g., Boersma & Levelt 2000, Curtin & Zuraw 2002).

final L1 ranking values from three sample simulations. While there is some variation in the precise numerical values, in each case MAX-C dominates DEP-V. Similar results are observed in all ten simulations. The learner establishes a *covert ranking* among the faithfulness constraints – one that is not motivated by the phonotactics of the L1 and that is not evident to the analyst on simple examination of the end-state L1 learner's productions.

(11) Sample constraint ranking values at the end of L1 acquisition of the Simple Language

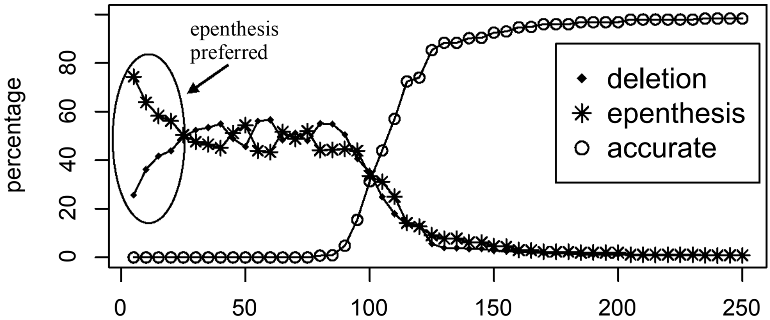
	Simulation 1	Simulation 2	Simulation 3
ONSET	100	100	100
*COMPLEXONSET	100	100	100
IDENT[±voice]	42	40	41
MAX-C	**40**	**39**	**40**
DEP-V	**34**	**34**	**34**
*VOICEDOBSTRUENT	34	31	32
NOCODA	26	27	26

In effect, the initial state transferred to the L2 is that given in (12). The high ranking value of *COMPLEXONSET ensures that the learner will consistently repair onset clusters at the beginning of the L2 learning stage. The MAX-C >> DEP-V ranking means that the preferred repair will be epenthesis.

(12) L2 initial state – actual ranking established at the end of L1 learning of the Simple Language
ONSET, *COMPLEXONSET >> IDENT[±voice], **MAX-C >> DEP-V** >> *VOICEDOBS >> NOCODA

Taking the transferred L1 ranking values as the starting point, the figure in (13) below shows the mean pattern of realization for the onset cluster inputs /pla/ and /bla/ over the 250 GLA learning trials that comprise the L2 learning phase.

(13) Realization of onset clusters over 250 GLA learning trials – mean results of 10 L2-learning simulations for the Cluster Language (initial state = end state of L1 Simple Language learning)



The initial bias favouring epenthesis repairs is found in all ten simulations, reflecting the consistency with which MAX-C dominates DEP-V at the end of L1 learning. The duration and strength of this preference varies across learning simulations, based both on the precise ranking values transferred at the end of L1 learning and the random selection of input forms and noise values during the L2 learning phase. Still, for all ten learners the rate of epenthesis repairs exceeds the rate of deletion repairs for at least 10 learning trials after L2 learning begins. Across learners there is a significant preference for epenthesis over the first 25 learning trials – one-sample t-test: $t(9) = 3.72$, $p = .004$. As the figure in (13) makes clear, the bias decreases with time; still, over the course of the first eighty-five L2 learning trials, before accurate realizations of /pl/ and /bl/ become possible, clusters are repaired via deletion in 46.91% of cases and via epenthesis in 53.09% of cases. Eight of the ten learners show an epenthesis bias throughout this time period. This effect stands in sharp contrast with the L1 learning scenario in §4.2, where no such preference for epenthesis was found. The learning stages for the full scenario are summarized in (14) – stages 3 through 5 are represented in (13).

(14) Stages in L2 learning of the Cluster Language

L1 learning (target = Simple Language)	**Stage 1**	/CCV/ → ? /CVC/ → [CV], [CVCə] *no exposure to target onset clusters; equal probability of deletion and epenthesis repairs for codas*
	Stage 2	/CCV/ → ? /CVC/ → [CVC] *no exposure to target onset clusters; faithful mapping of target codas*
L2 learning (target = Cluster Language)	**Stage 3**	/CCV/ → [CəCV] /CVC/ → [CVC] *preference for epenthesis repairs with target onset clusters; faithful mapping of target codas*
	Stage 4	/CCV/ → [CV], [CəCV] /CVC/ → [CVC] *equal probability of deletion and epenthesis repairs with target onset clusters; faithful mapping of target codas*
	Stage 5	/CCV/ → [CCV] /CVC/ → [CVC] *faithful mapping of target onset clusters and codas*

4.4. Explaining the L2 epenthesis bias

The covert MAX-C >> DEP-V ranking that is responsible for the epenthesis bias in L2 learning of the Cluster Language is a direct consequence of the earlier L1 learning process. To see this, we can consider the three forms that comprise the Simple Language and the types of errors that these give rise to.[3] Additional output candidates, like /tab/→[tap] and /tab/→[tabə] were included in the simulations, but are not discussed here because they do not satisfy both markedness constraints and therefore are never selected as optimal at the early stages of learning – i.e., they are not errors that are able to inform learning at the critical stage.

(15) a. Target and error forms for input /pat/

/pat/	*VOICEDOBS	NOCODA	MAX-C	DEP-V	IDENT[±voice]
Target (W) pat		* →			
Error (L) pa			← *		
Error (L) patə				← *	

b. Target and error forms for input /ba/

/ba/	*VOICEDOBS	NOCODA	MAX-C	DEP-V	IDENT[±voice]
Target (W) ba	* →				
Error (L) pa					← *

c. Target and error forms for input /tab/

/tab/	*VOICEDOBS	NOCODA	MAX-C	DEP-V	IDENT[±voice]
Target (W) tab	* →	* →			
Error (L) ta			← *		
Error (L) tapə				← *	← *

Each of the faithfulness constraints in (15) is implicated in an equal number of potential errors. MAX-C is implicated in /pat/→[pa] and /tab/→[ta] errors, DEP-V is implicated in /pat/→[patə] and /tab/→[tapə] errors, and IDENT[±voice] is implicated in /ba/→[pa] and /tab/→[tapə] errors. In fact,

[3] ONSET and *COMPLEXONSET are excluded from the tableaux for reasons of space. Both constraints are unviolated in the Simple Language and their ranking values are therefore unaltered during the L1 learning phase.

however, not all of these potential errors are equally probable given the dynamics of the model. This is crucial. The GLA is error driven, and so changes in constraints' values are directly tied to presence of errors that implicate the relevant constraints. To the extent that errors implicating a particular faithfulness constraint are more frequent, that constraint's value will increase at a relatively faster rate.

In the case of the Simple Language, errors involving MAX-C are relatively more frequent than errors involving DEP-V; this ultimately leads to a covert MAX-C >> DEP-V ranking. The reason for this lies in (15c), where the input /tab/ contains both a coda and a voiced obstruent. The learner has two options for repairing this form: deletion, mapping /tab/ to [ta], or epenthesis *plus* a change in voicing specification, mapping /tab/ to [tapə]. In practice, the deletion mapping, implicating MAX-C, is more frequent than the epenthesis plus feature change mapping, which implicates DEP-V and IDENT[±voice]. This difference in error probability is a straightforward consequence of the number of total rankings among the faithfulness constraints that give rise to the two possible errors. As (16) shows, four of the six rankings prefer the deletion mapping; if the application of noise leads the learner to select any of these four rankings, deletion results.

(16) Errors resulting from different rankings of the Faithfulness constraints

Ranking	Error	Effect
1. MAX-C >> DEP-V >> IDENT[±voice]	/tab/→[tapə]	increase value of DEP-V and IDENT[±voice]
2. MAX-C >> IDENT[±voice] >> DEP-V	/tab/→[tapə]	
3. DEP-V >> MAX-C >> IDENT[±voice]	/tab/→[ta]	increase value of MAX-C
4. DEP-V >> IDENT[±voice] >> MAX-C	/tab/→[ta]	
5. IDENT[±voice] >> MAX-C >> DEP-V	/tab/→[ta]	
6. IDENT[±voice] >> DEP-V >> MAX-C	/tab/→[ta]	

In order for /tab/→[ta] to be selected by the learner, it is sufficient for either DEP-V *or* IDENT[±voice] to dominate MAX-C on a given iteration of EVAL. On the other hand, in order for /tab/→[tapə] to be selected by the learner, MAX-C must dominate both DEP-V *and* IDENT[±voice]. As a result, deletion errors with /tab/ are twice as probable as epenthesis + feature change errors. All else being equal, an error that involves only a single faithfulness violation is more probable than an error that involves multiple faithfulness violations.[4]

The ultimate consequence of this asymmetry is a MAX-C >> DEP-V ranking at the end of L1 learning, which, when transferred to the L2, leads to the observed epenthesis bias. This ranking is covert. The Simple Language does not require that MAX-C dominate DEP-V; it only requires that both of these constraints dominate the conflicting markedness constraints. Nonetheless, the complexity of the L1 learning task – and particularly the fact that sources of segmental markedness are being learned simultaneously and in interaction with sources of syllable-structure markedness – means that not all errors are equally probable and more articulated rankings of the faithfulness constraints emerge during acquisition.

5. Discussion

5.1. The special role of MAX-C

The patterns discussed in §4 rely on differences in the probability of specific errors over the course of learning. The particular properties of MAX-C play an important role here, as segmental deletion can resolve a wide range of segmental and syllable-structure issues – e.g., deletion of a single segment can eliminate both the voiced obstruent and the coda in an input like /tab/.

Within the context of the GLA, this power of MAX-C has two effects. First, it predicts that certain types of forms will favour deletion repairs to a greater extent than others. This is true in the

[4] The probability of epenthesis + feature change errors is further reduced by the fact high-ranking ONSET means that only one type of error is ever committed with input /ba/, and that error consistently causes promotion of IDENT[±voice]. Within the dynamics of the system, this further reduces the probability of MAX-C dominating both DEP-V and IDENT[±voice].

simulations of L1 acquisition of the Simple Language, where /tab/, in particular, shows a preference for deletion over epenthesis. The empirical prediction is simple: children should display a greater tendency toward deletion with forms likes /tab/ where sources of segmental and syllable-structure markedness overlap, than with forms like /pat/ where there is only a single source of markedness.

Second, because the development of a covert MAX-C >> DEP-V ranking is dependent upon the existence of forms like /tab/ that favour deletion repairs, L1s that do not include such forms should not lead to this type of bias. Languages with only CV input syllables, for example, will generally give rise to few errors that trigger promotion of either MAX-C or DEP-V during phonotactic learning. As a result, a learner of such an L1 will begin L2 learning from an initial state that does not include a covert MAX-C >> DEP-V ranking. This learner is predicted to display a much weaker bias toward epenthesis than a learner coming from an L1 with a richer inventory of syllable types. Conversely, learners whose L1 has a particularly rich syllable-structure inventory are expected to display a relatively stronger epenthesis bias.[5]

5.2. Strengthening the bias

The effects discussed in §4 are subtle. While these simulations demonstrate an increased preference for epenthesis in L2 acquisition, the bias found here is considerably weaker than one might expect given the differences of between L1 and L2 learners described in the literature. The question, then, is how the emergent biases might be strengthened in order to better reflect the empirical patterns.

Modeling an actual preference for deletion in L1 acquisition is one important component of this project. The small inventories of the toy languages discussed in §4 are in part to blame for the non-appearance of a clear deletion bias in the L1 case. As more diverse sources of syllable-structure and segmental markedness are added to a language, the range of forms where deletion is the preferred repair increases. This translates into higher overall rates of deletion in the learner's productions. The inclusion of other types of markedness constraints can further perpetuate this effect, and even categorically rule out epenthesis in certain cases. For instance, the inclusion of a constraint militating against weak initial syllables will disfavour epenthesis into onset clusters quite effectively (Fikkert 1994). Similarly, the inclusion of constraints disfavouring high-sonority segments in syllable-initial position will rule out epenthesis into clusters like /fl/, where the addition of a vowel creates a new violation. As the preference for deletion increases in the L1, so too does the number of errors motivating promotion of MAX-C. This leads to a relatively greater difference in the ranking values of MAX-C and DEP-V at the end of L1 learning, which leads to a stronger epenthesis bias in the L2.

An alternative approach builds on the fact that the simulations in §4 assumed that learning proceeds at the same rate throughout the course of both L1 and L2 learning. In fact, it is likely that the rate of phonological learning is considerably slower in the L2 than in the L1. This can be straightforwardly accommodated within the GLA by decreasing the plasticity during the L2 learning phase. This has the effect of increasing the duration of the initial stage where the effects of the transferred MAX-C >> DEP-V ranking are most evident – i.e., Stage 3 in (14) – and of slowing the overall rate of progression toward a target-like L2 grammar. Fossilization during this phase will entrench the epenthesis bias over the long term.

5.3. Modulating the bias

Beyond the dynamics discussed in this paper, analyses of L1 and L2 cluster acquisition have shown that a number of additional phonological factors can influence learners' repairs. Carlisle (2006), for instance, found that Spanish-speaking learners of English are more likely to repair onset /sC/ clusters with worse sonority profiles, and that repair is more likely when the preceding word ends with a consonant rather than a vowel. In both cases epenthesis was the only repair employed, suggesting that the primary effect here is on the rate of acquisition. These types of findings can be readily

---

[5] These predictions rely on two assumptions. First, the L1s of the learners being compared must have similar degrees of segmental complexity. Second, the L1s of the learners should not include any alternations that provide overt evidence for either a MAX-C >> DEP-V or DEP-V >> MAX-C ranking.

accommodated within the OT model through the addition of constraints that distinguish between the different types of clusters and contexts. The sonority facts, for instance, can be captured through the inclusion of Sonority Sequencing constraints that are either stringently defined or arranged in a fixed hierarchy (for one approach, see Baertsch 2002). The formulation of these constraints ensures that, across languages, clusters with more harmonic sonority profiles are preferred to those with less harmonic sonority profiles. The existence of these constraints in the L1, and their subsequent transfer to the L2, predicts that learners will be more likely to repair falling-sonority onset clusters like /st/ than rising-sonority onset clusters like /sl/ and /sn/.

A second important modulating effect on the bias discussed in this paper is transfer of *overt* rankings – i.e., rankings motivated by alternations in the L1. If, unlike the Simple Language discussed in §4, the L1 includes alternations like /pat/→[pa], the learner will promote DEP-V above MAX-C (and NOCODA). Transferred to the L2, this DEP-V >> MAX-C ranking will ensure that deletion, rather than epenthesis, dominates as the preferred cluster repair strategy. Conversely, if the L1 includes alternations like /pat/→[patə], the learner will have additional evidence for promotion of MAX-C above DEP-V, strengthening the overall epenthesis preference. Rankings established among other constraints during L1 acquisition can also shape the L2 learner's treatment of clusters. For example, Broselow et al. (1998) attribute L1-Mandarin learners' increased preference for epenthesis with monosyllabic (vs. longer) inputs to the fact that epenthesis into monosyllables allows for satisfaction of the transferred high-ranking L1 constraint WDBIN ("Words should consist of two syllables" – Broselow et al. 1998:272). These types of effects are fully expected given the model developed here.

6. Conclusion

Within the GLA, error patterns during L1 learning affect not just the immediate productions of the child, but also the final constraint ranking that is established. In many cases, this constraint ranking may be more articulated than is strictly necessary in order to capture the patterns of the L1, and so its effects are largely unobserved. With the transfer of the L1 grammar to the L2, however, these covert rankings have new opportunities to make themselves known. This paper has demonstrated how this mechanism – and specifically the establishment of a covert MAX-C >> DEP-V ranking – can give rise to an epenthesis bias in L2 learning. This pattern emerges automatically as a simple consequence of constraint interaction and error-driven learning. Future research will determine to what extent other patterns can be attributed to the transfer of covert rankings, unifying the treatment of cluster repair preferences and other L1 vs. L2 asymmetries.

References

Abrahamsson, Niclas. 1999. Vowel epenthesis of /sC(C)/ onsets in Spanish/Swedish interphonology: a longitudinal case study. *Language Learning* 49(3): 473-508.

Baertsch, Karen. 2002. *An Optimality Theoretic Approach to Syllable Structure: the Split Margin Hierarchy*. PhD dissertation. Indiana University.

Boersma, Paul. 1998. *Functional Phonology: Formalizing the Interaction Between Articulatory and Perceptual Drives*. The Hague: Holland Academic Graphics.

Boersma, Paul & Bruce Hayes. 2001. Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32: 45-82.

Boersma, Paul & Clara Levelt. 2000. Gradual constraint-ranking learning algorithm predicts acquisition order. In E.V. Clark (ed.), *Proceedings of the 30th Child Language Research Forum,* 229–237. Stanford, CA: CSLI.

Boersma, Paul & David Weenink. 2012. *PRAAT: Doing Phonetics by Computer v.5.2.42*. [www.praat.org].

Broselow, Ellen. 2004. Unmarked structures and emergent rankings in second language phonology. *International Journal of Bilingualism* 8(1): 51-65.

Broselow, Ellen, Su-I Chen & Chilin Wang. 1998. The emergence of the unmarked in second language phonology. *Studies in Second Language Acquisition* 20: 261-280.

Carlisle, Robert S. 1991. The influence of environment on vowel epenthesis in Spanish/English interphonology. *Issues and Developments in English and Applied Linguistics* 3: 15-23.

Carlisle, Robert S. 1997. The modification of onsets in a markedness relationship: testing the Interlanguage Structural Conformity Hypothesis. *Language Learning* 47(2): 327-361.

Carlisle, Robert S. 1998. The acquisition of onsets in a markedness relationship. *Studies in Second Language Acquisition* 20: 245-260.

Carlisle, Robert S. 2006. The sonority cycle and the acquisition of complex onsets. In B. Baptista & A. Watkins (eds.), *English with a Latin Beat: Studies in Portuguese/Spanish-English Interphonology*, 105-137. Amsterdam: John Benjamins.

Curtin, Suzanne & Kie Zuraw. 2002. Explaining constraint demotion in a developing system. In Anna H.-J. Do, Laura Domínguez & Aimee Johansen (eds.), *BUCLD 26: Proceedings of the 26th Annual Boston University Conference on Language Development.* Somerville, MA: Cascadilla Press.

Demuth, Katherine. 1995. Markedness and the development of prosodic structure. In Jill Beckman (ed.), *Proceedings of the 25th Meeting of the North East Linguistic Society.* Amherst, MA: GLSA. [ROA-50].

Dyson, Alice T. & Elaine P. Paden. 1983. Some phonological acquisition strategies used by two-year-olds. *Journal of Childhood Communication Disorders* 7(1): 6-18.

Eckman, Fred. 1981. On the naturalness of interlanguage phonological rules. *Language Learning* 31: 195-216.

Fikkert, Paula. 1994. *On the Acquisition of Prosodic Structure*. PhD dissertation. University of Amsterdam.

Gnanadesikan, Amalia. 2004. Markedness and faithfulness constraints in child phonology. In René Kager, Joe Pater & Wim Zonneveld (eds.), *Fixing Priorities: Constraints in Phonological Acquisition,* 73-108. Cambridge: Cambridge University Press.

Jesney, Karen. 2011. *Cumulative Constraint Acquisition in Phonological Acquisition and Typology*. PhD dissertation. University of Massachusetts Amherst.

Kim, Jong-mi. 2009. Insertion preferred to deletion in learner speech: a study of Korean English. *Jungang Journal of English Language and Literature* 9: 101-123.

Kwan, Bo-Young. 2005. The patterns of insertion in IL phonology: the P-map account. *Studies in Phonetics, Phonology and Morphology* 11(2): 197-225.

Levelt, Clara C. 1994. *On the Acquisition of Place*. PhD dissertation. Leiden University.

Lin, Yuh-Huey. 2001. Syllable simplification strategies: a stylistic perspective. *Language Learning* 51(4): 681-718.

Lombardi, Linda. 2003. Second language data and constraints on Manner: explaining substitutions for the English interdentals. *Second Language Research* 19(3): 225-250.

McCarthy, John J. & Alan Prince. 1995. Faithfulness and reduplicative identity. *Papers in Optimality Theory: University of Massachusetts Occasional Papers* 18: 249-384. Amherst, MA: GLSA.

McLeod, Sharynne, Jan van Doorn & Vicki A. Reed. 2001. Normal acquisition of consonant clusters. *American Journal of Speech-Language Pathology* 10: 99-110.

Pater, Joe. 1997. Minimal violation in phonological development. *Language Acquisition* 6(3): 201-253.

Preisser, Debra A., Barbara W. Hodson & Elaine P. Paden. 1988. Developmental phonology: 18-29 months. *Journal of Speech and Hearing Disorders* 53: 125-130.

Prince, Alan & Paul Smolensky. 1993/2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Technical Report, Rutgers University and University of Colorado Boulder, 1993. Revised version published by Blackwell, 2004.

R Development Core Team. 2012. R: *A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. [www.R-project.org].

Rauber, Andréia Schurt. 2006. Production of English initial /s/-clusters by speakers of Brazilian Portuguese and Argentine Spanish. In B. Baptista & A. Watkins (eds.), *English with a Latin Beat: Studies in Portuguese/Spanish-English Interphonology*, 155-167. Amsterdam: John Benjamins.

Schwartz, Bonnie & Rex Sprouse. 1996. L2 cognitive states and the full transfer / access model. *Second Language Research* 12: 4-72.

Smolensky, Paul. 1996. On the comprehension / production dilemma in child language. *Linguistic Inquiry* 21: 720-731.

Tesar, Bruce & Paul Smolensky. 1998. Learnability in Optimality Theory. *Linguistic Inquiry* 29(2): 229-268.

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

Tropf, Herbert S. 1987. Sonority as a factor in second language phonology. In J. Leather & A. James (eds.), *Sound Patterns in Second Language Acquisition*, 173-191. Dordrecht: Foris.

Weinberger, Steven H. 1987. The influence of linguistic context on syllable simplification. In Georgette Ioup & Steven H. Weinberger (eds.), *Interlanguage Phonology: The Acquisition of a Second Language Sound System,* 401-417. Rowley, MA: Newbury House.

Weinberger, Steven H. 1994. Functional and phonetic constraints on second language phonology. In Mehmet Yavaş (ed.), *First and Second Language Phonology,* 283-302. San Diego, CA: Singular.

Selected Proceedings of the 5th Conference on Generative Approaches to Language Acquisition North America (GALANA 2012)

edited by Chia-Ying Chu,
Caitlin E. Coughlin, Beatriz Lopez Prego,
Utako Minai, and Annie Tremblay

Cascadilla Proceedings Project   Somerville, MA   2014

Copyright information

ISBN 978-1-57473-463-8 library binding

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Jesney, Karen. 2014. A Learning-Based Account of L1 vs. L2 Cluster Repair Differences. In *Selected Proceedings of the 5th Conference on Generative Approaches to Language Acquisition North America (GALANA 2012)*, ed. Chia-Ying Chu et al., 10-21. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #3062.