Comparing Different Optimality-Theoretic Learning Algorithms: The Case of Metrical Phonology

Diana Apoussidou and Paul Boersma

Institute of Phonetic Sciences, University of Amsterdam
Herengracht 338
1016 CG Amsterdam, The Netherlands
d.apoussidou@uva.nl, paul.boersma@uva.nl

Abstract

We fed short overt Latin stress patterns to 100 virtual language learners whose grammars consist of a universal set of 12 Optimality-Theoretic constraints. For 50 learners the learning algorithm was Error-Driven Constraint Demotion (EDCD), for the remaining 50 it was the Gradual Learning Algorithm (GLA). The EDCD learners did not succeed: they ended up in a grammar that could not reproduce the correct stress pattern. The GLA learners did succeed: they came up with an analysis close to one of the analyses proposed in the literature, namely that by Jacobs (2000). These results add to previous findings that the GLA seems to be a more realistic ingredient than EDCD for models of actual language acquisition.

Introduction

In Latin, the positioning of main stress in a word is straightforward (e.g. Allen 1973): stress the penultimate syllable if it is heavy (i.e. if it contains a long vowel or ends in a consonant), else the antepenultimate syllable:

amice 'friend'	a.mí:.ke	L H1 L
rapiditas 'speed'	ra.pí.di.ta:s	L L1 L H
misericordia 'pity'	mi.se.ri.kór.di.a	L L L H1 L L
perfectus 'perfect'	per.fék.tus	H H1 H
incipio 'begin'	in.kí.pi.o:	H L1 L H
domesticus 'domestic'	do.més.ti.kus	L H1 L H
homo 'man'	hó.mo:	L1 H

(':' = long vowel, '.' = syllable boundary, 'é' etc. = stressed vowel, 'H' = heavy, 'L' = light, '1' = main stress)

The positioning of main (primary) stress in Latin is thus *weight-sensitive*, i.e. influenced by whether the syllables involved are heavy or light. Since it is not clear whether Latin also had secondary stress, we will assume in this paper that it did not (but see Apoussidou and Boersma 2003 for similar results with secondary stress). The question addressed in this paper is what a virtual "child" learning Latin would do when she is given overt Latin language data with primary but without secondary stress.

The specific questions addressed in this paper are:

- Will the child come up with a grammar that describes the overt data correctly?
- Are there any differences in performance between different learning algorithms?
- If the child is only trained with short words (up to four syllables), will she be able to generalize her analysis to longer words?

We model the child's metrical phonology as a constraint-based grammar within the framework of Optimality Theory (OT; Prince and Smolensky 1993). For the forms (underlying, surface, and overt) and the constraint set we will follow Tesar and Smolensky (2000), but find it necessary to change one of their constraints. We will consider two learning algorithms: RIP/EDCD (Tesar and Smolensky 2000) and RIP/GLA (Apoussidou and Boersma 2003).

The Training Set of Overt Forms

In their simulations of the metrical phonologies of 124 language types, Tesar and Smolensky (2000) trained their learners with overt forms that consisted of two to seven syllables. For our simulations of Latin, we will only train our learners on forms with two to four syllables, and reserve the longer forms to test the learners' generalization capabilities. All possible sequences of heavy and light syllables will be taken into account. Thus, four patterns of overt disyllables will be fed to the child: [L1 L], [L1 H], [H1 L], [H1 H]. Likewise, there will be eight trisyllabic overt forms: [L1 L L], [L1 L H], [L H1 L], [L H1 H], [H1 L L], [H1 L H], [H H1 L], and [H H1 H]. In the same vein, there are 16 overt forms with four syllables, all following the penultimate/antepenultimate Latin stress rule.

Candidate Surface Structures in Perception

The universal language-learning child does not interpret any overt form as is. Instead, she will assign a phonological surface structure to it, where every stressed syllable is contained in a *foot*, which (in a common simplification

followed by Tesar and Smolensky and by us here) is a hidden structure that consists of one or two syllables, one of which (the *head syllable*) bears main or secondary stress. It is assumed that every word contains exactly one *head foot*, which contains the syllable with main stress, and that a language can have any number of secondary feet. Thus, the child could analyse the overt form [H1 L] in either of two ways. First, she could think that both syllables together build a foot; this leads to the surface form /(H1 L)/, where the parentheses indicate the foot. Second, she could think that the first syllable builds a foot by itself, while the second syllable remains unfooted: /(H1) L/. When hearing [H1 L], therefore, the child will have to choose between /(H1 L)/ and /(H1) L/; no other surface form is possible, e.g., /H1 L/ is impossible because every stressed syllable should be contained in a foot, and /(H1) (L)/ is impossible because every foot must have a stressed syllable.

For longer forms there can be more candidate surface forms, e.g. the overt form [L H1 L] has three: /L (H1) L/, /(L H1) L/, and /L (H1 L)/. Note that a foot with three syllables like /(L H1 L)/ is impossible according to the assumptions mentioned above. For each of the 28 possible overt forms in the training set, the choice between the candidates will be determined by an Optimality-Theoretic constraint ranking.

The Constraints

The candidate surface forms are evaluated by Optimality-Theoretic constraints, which in this case are restrictions on the wellformedness of these surface structures. The constraint set that we use is based on the one adopted by Tesar and Smolensky (2000). Their 12 constraints are:

ALL-FEET-LEFT (AFL):
"align each foot with the word, left edge."
ALL-FEET-RIGHT (AFR):
"align each foot with the word, right edge."
MAIN-LEFT (MAIN-L):
"align the head foot with the word, left edge."
MAIN-RIGHT (MAIN-R):
"align the head foot with the word, right edge."
WORD-FOOT-LEFT (WFL):
"align the word with some foot, left edge."
WORD-FOOT-RIGHT (WFR):
"align the word with some foot, right edge."
NONFINAL:
"do not foot the final syllable of the word."
PARSE:
"each syllable must be footed."
FOOTNONFINAL:
"each head syllable must not be final in its foot."
IAMBIC:
"align each foot with its head syllable, right edge."
WEIGHT-TO-STRESS-PRINCIPLE (WSP):
"each heavy syllable must be stressed."
FOOTBIN:
"each foot must be either bimoraic or bisyllabic."

We will discuss the precise meaning of each constraint.

The alignment constraints AFL and AFR make sure that a foot is aligned with one of the edges of a word. Their violation is gradient: AFL is assigned one violation mark for every syllable between the left edge of the word and the left edge of every foot. In the candidate surface form /L (L2 L) (L1 L)/, where '2' stands for secondary stress, AFL is violated four times: once for the first foot, three times for the second foot.

The constraints MAIN-L and MAIN-R do the same as AFL and AFR, but only for the foot that contains the main stress. Thus, the candidate /L (L2 L) (L1 L)/ violates MAIN-L three times, and MAIN-R not at all.

The two WORD-FOOT alignment constraints favour candidates where at least one foot is aligned with the word edge. These constraints are not gradient, but *binary*: they are assigned a single violation mark if there is an unfooted syllable at the edge of the word. Thus, the candidate /L L (L1) (L2 L)/ violates WFL (once), but not WFR.

The constraint NONFINAL expresses extrametricality: it is violated if the last syllable is *parsed* (included) in a foot. This constraint thus prefers /(L1) L/ to /(L1 L)/.

The constraint PARSE favours candidates in which all syllables are parsed into feet. It is assigned one violation mark for each unfooted syllable. Thus, the candidate /L (L1 L) L L/ violates PARSE three times.

The constraint FOOTNONFINAL favours candidates with trochaic (initially stressed) feet like (L1 L), (L2 L), (L1 H), and so on. However, *degenerate feet* consisting of only one syllable, like (L1) and (H2), violate this constraint. The constraint IAMBIC favours candidates with iambic (finally stressed) feet like (L L1), and this constraint is *not* violated in degenerate feet like (L1).

The WEIGHT-TO-STRESS-PRINCIPLE favours candidates that have stress on a heavy syllable. Every unstressed heavy syllable causes a violation. Thus, /(L2 H) H (H1) L/ violates WSP twice (once for the unfooted H, once for the H in the first foot's weak position), whereas /(L H2) (H2) (H1) L/ does not violate WSP.

FOOTBIN is the constraint for foot size: feet should be binary regarding either syllables or *moras*; a light syllable counts as one mora, a heavy syllable as two. In the candidate set under discussion here, this constraint is only assigned a violation mark for each monosyllabic light foot, i.e. (L1) and (L2), whereas feet like (L1 H) and (H H2) do not violate this constraint (remember that candidates with more than two syllables are not generated).

Our simulations will not work with Tesar and Smolensky's constraint set, probably because there *is* no ranking of these constraints that can produce the Latin forms (Apoussidou and Boersma 2003). We therefore replace the trochaicity constraint FOOTNONFINAL with a more common constraint (e.g. Prince and Smolensky 1993, Jacobs 2000), which is defined analogously to IAMBIC, and unlike FOOTNONFINAL does not punish monosyllabic feet such as (H1) or (L2):

TROCHAIC (replaces FOOTNONFINAL):
"align each foot with its head syllable, left edge."

overt: [H1 L]	WFL	WFR	MAIN-L	MAIN-R	PARSE	AFL	AFR	FOOTBIN	IAMBIC	TROCHAIC	WSP	NONFINAL
☞ /(H1 L)/									*			*
/(H1) L/		*!		*	*		*					

Tableau 1. Perception of a disyllabic overt form.

overt: [L H L]	WFL	WFR	MAIN-L	MAIN-R	PARSE	AFL	AFR	FOOTBIN	IAMBIC	TROCHAIC	WSP	NONFINAL
/L (H1) L/	*!	*	*	*	**	*	*					
☞ /(L H1) L/		*		*	*		*			*		
/L (H1 L)/	*!		*		*	*			*			*

Tableau 2. Perception of a trisyllabic overt form.

Evaluation in Perception

We will give some examples of how the choice between the candidates is made in the mapping from overt form to full phonological surface structure. We can call this mapping *perception*, although Tesar and Smolensky (1998, 2000) call it *robust interpretive parsing* (RIP). As an example, imagine a child who has the constraints accidentally ranked in the order shown in tableau 1, which is very different from an order that could accommodate Latin. This *perception tableau* shows that this child will perceive the overt form [H1 L] as /(H1 L)/ (the two candidates were discussed before). The highest-ranked constraint, WFL, cannot decide between the two candidates, because both candidates satisfy it. The second candidate is ruled out by the next constraint, WFR, because the last syllable of the word is not contained in a foot. Constraints below WFR do not contribute to determining the winner of this tableau; their cells are therefore greyed out.

Tableau 2 shows how the same child perceives [L H1 L] (the three candidates were discussed before). This time, the decision is made by WFL, which is violated by both the first and the third candidate, since both forms do not begin with a foot. So far, everything looks fine: the child has found a structural description that suits the given overt form, since in the winning candidates of tableaus 1 and 2, stress is on the same position as in the overt form.

Error Detection by Virtual Production

How can the child learn anything from the forms she has just perceived in tableaus 1 and 2? Tesar and Smolensky's proposal is that the child uses the same constraint ranking to determine what she herself would have said, given the overt forms [H1 L] and [L H1 L]. The first thing she has to do is to determine the underlying forms. This is trivial in this case. Under Tesar and Smolensky's assumption (which is of course not universally valid) that the child will posit no specifications for stress in her lexicon, the overt form [H1 L] can only mean that the underlying form is |H L|. The child then tries to determine what surface form she would have produced herself, given this underlying, stress-free form. The *production tableau* 3, which has the same constraint ranking as the perception tableaus 1 and 2, shows the evaluation of the candidate surface forms in production. There are now six candidates, not only the two with the overt form [H1 L] that we saw above but also two candidates with the overt form [H L1], one with the overt form [H1 L2], and one with the overt form [H2 L1], because the placement of stress in production is not given but has to be determined by the constraint ranking. In this tableau, the two candidates with a single monosyllabic foot are ruled out by high-ranking WFL and WFR. Next, the two candidates with two monosyllabic feet are ruled out by MAIN-L and MAIN-R. The remaining candidates are the

underlying: |H L|	WFL	WFR	MAIN-L	MAIN-R	PARSE	AFL	AFR	FOOTBIN	IAMBIC	TROCHAIC	WSP	NONFINAL
√ /(H1 L)/									*!			*
/(H1) L/		*!		*	*		*					
/(H1) (L2)/				*!		*	*	*				*
/H (L1)/	*!		*		*	*		*			*	*
☞ /(H L1)/										*	*	*
/(H2) (L1)/			*!			*	*	*				*

Tableau 3. Production of a disyllabic underlying form.

ones that contain a disyllabic foot, one of them trochaic, the other one iambic. Since IAMBIC outranks TROCHAIC as well as WSP, the candidate with the iambic foot wins. Now that the child has determined the surface form that she would have produced, as depicted by the pointing finger in tableau 3, she must notice that it is different from the actual form that she has heard the adult say according to tableau 1, which is depicted with a check mark in tableau 3. We can say that the child has detected an *error*.

Analogous things can be said about the overt form [L H1 L]. The underlying form must be |L H L|. The production is in tableau 4. Apart from the three candidates of tableau 2, we now also have 21 candidates with different overt forms. WFL and WFR rule out the candidates with an initial or final unfooted syllable. They even kick out the candidate that was chosen as optimal by the child in perception (√). MAIN-L and MAIN-R as the next constraints rule out most of the other candidates. The two remaining candidates differ only in that one of them has an iambic foot, while the other has a trochaic foot. Once more, IAMBIC decides: it favours the candidate with the iambic foot /(L H1) (L2)/. Again, the child will detect a mismatch between her perception of the adult form and her own production. What can she do to adjust her grammar in such a way that there is less divergence between future perceived forms and future produced forms?

underlying: |L H L|	WFL	WFR	MAIN-L	MAIN-R	PARSE	AFL	AFR	FOOTBIN	IAMBIC	TROCHAIC	WSP	NONFINAL
/(L1) H L/		*!		**	**		**	*			*	
/(L1 H) L/		*!		*	*		*		*		*	
/(L1) H (L2)/				**!	*	**	**	**			*	*
/(L1) (H L2)/				**!		*	**	*		*	*	*
/(L1 H) (L2)/				*		**	*	*	*!		*	*
/(L1) (H2) L/		*!		**	*	*	***	*				
/(L1) (H2 L)/				**!		*	**	*	*			*
/(L1) (H2) (L2)/				**!		***	***	**				*
/L (H1) L/	*!	*	*	*	**	*	*					
/L (H1 L)/	*!		*		*	*			*			*
√ /(L H1) L/		*!		*	*		*			*		
/L (H1) (L2)/	*!		*	*	*	***	*	*				*
☞ /(L H1) (L2)/				*		**	*	*		*		*
/(L2) (H1) L/		*!	*	*	*	*	***	*				
/(L2) (H1 L)/			*!			*	**	*	*			*
/(L2) (H1) (L2)/			*!	*		***	***	**				*
/L H (L1)/	*!		**		**	**		*			*	*
/L (H L1)/	*!		*		*	*				*	*	*
/L (H2) (L1)/	*!		**		*	***	*	*				*
/(L H2) (L1)/			*!*			**	*	*		*		*
/(L2) H (L1)/			*!*		*	**	**	**			*	*
/(L2) (H L1)/			*!			*	**	*		*	*	*
/(L2 H) (L1)/			*!*			**	*	*	*		*	*
/(L2) (H2) (L1)/			*!*			***	***	**				*

Tableau 4. Production of a trisyllabic underlying form.

Error-Driven Learning

Once the child has detected an error, she can take action by changing the ranking of her constraints. Two methods have been described in the literature. The first is Error-Driven Constraint Demotion (Tesar 1995). This algorithm first looks up the highest-ranked constraint that prefers the 'correct' form (the child's perception of the adult form) to the child's own form. In tableau 4, this is AFL. All the even higher ranked constraints that prefer the child's own form (in tableau 4, these are WFR and PARSE) are then demoted below AFL, i.e. into the *stratum* where AFR is. The immediate result of this is that in this new grammar the originally perceived adult form /(L H1) L/ has become better than the child's own produced form /(L H1) (L2)/. The other algorithm is the Gradual Learning Algorithm (Boersma 1997). This algorithm assumes Stochastic OT (Boersma 1998), a version of Optimality Theory in which all constraints are ranked along a continuous scale. All the constraints that prefer the 'correct' form /(L H1) L/ (namely AFL, FOOTBIN, and NONFINAL) are shifted up along this scale by a small amount, and all the constraints that prefer the child's own form /(L H1) (L2)/ (namely WFR and PARSE) are shifted down by the same amount. The immediate result of this is that it becomes slightly more likely that /(L H1) L/ will be the winner in production in the future, and slightly less likely that /(L H1) (L2)/ will win next time. In tableau 3, completely different constraints will shift: the GLA will cause TROCHAIC and WSP to go up, IAMBIC to go down.

In the simulations below we compare the performances of both algorithms.

The Simulations

We fed a large number of overt forms, randomly drawn from the 28 possible forms in our training set, to 50 virtual EDCD learners and 50 virtual GLA learners. For both learning algorithms we used the implementation in the Praat program (www.praat.org). The evaluation model for EDCD was OT with crucial ties, i.e. the violations of constraints that are ranked equally high are added to each other as if these constraints formed a single constraint (in Praat, this can be simulated by setting the evaluation noise to zero). As in Tesar and Smolensky (2000), we allowed the algorithm to chew five times on each piece of language data, with backtracking if the pentuple chews did not succeed in making the (alleged) correct adult form optimal in the learner's grammar. When two forms were equally harmonic, we chose a winner randomly from among them. The evaluation model for the GLA was Stochastic OT with an evaluation noise of 2.0. This noise leads to slightly different rankings of the constraints in each evaluation, i.e., if the ranking value of constraint B along the continuous ranking scale is just a bit lower than that of constraint A, then A will outrank B in most evaluations, but B will in turn outrank A in a minority of cases. Within an evaluation of an overt form, however, the ranking stayed constant: the same ranking values drawn from the Gaussian distributions were used first to interpret the overt form into a perceived surface form and into an underlying form, then to produce the learner's surface form from this underlying form.

For each of the 100 virtual learners, all 12 constraints were initially ranked at a height of 100, whereupon 10,000 language data were drawn randomly with equal probability from the 28 overt forms. When a form caused a mismatch between the child's own produced surface form and her perceived adult form, the EDCD learner had an adjustment model that would demote the ranking of one constraint by a distance of 1 (e.g. to 99 when a constraint is demoted for the first time), and the GLA learner had an adjustment model that would raise some ranking values by 0.1 and lower some ranking values by 0.1; in the case of the GLA learner, this *plasticity* of 0.1 was further randomized by a relative plasticity standard deviation of 0.1.

Results

None of the EDCD learners succeeded in learning the stress pattern of Latin. The ranking after 10,000 data of one showcase EDCD learner is given in table 1.

Constraints	*Ranking values*
FOOTBIN	100.000
NONFINAL	100.000
AFR	99.000
MAIN-R	99.000
PARSE	99.000
WFR	99.000
AFL	98.000
MAIN-L	98.000
WFL	98.000
WSP	–2497.000
TROCHAIC	–2498.000
IAMBIC	–2499.000

Table 1. A failing EDCD learner, after 10,000 data.

At this snapshot in time, this child produces correct forms like /(L1 L) X/ but also incorrect forms like /(H1 H) X/. When hearing the correct overt form [H H1 X], the child will perceive this as /(H H1) X/, given the ranking in table 1. This will lead her to demote TROCHAIC below IAMBIC, i.e. to –2500. But this new grammar will incorrectly produce /(L L1) X/, so that when hearing [L1 L X] the learner will demote IAMBIC below TROCHAIC again. These two constraints will continue to tumble down hopelessly along the ranking scale. They will drag along WSP, because when WSP is ranked above TROCHAIC, the learner can make the error /(L H1)/, so that hearing [L1 H] will lead her to demote WSP below TROCHAIC.

In contrast to the EDCD learners, all 50 GLA learners succeeded (though five of them needed between 10,000 and 200,000 data to converge). Table 2 shows an example.

Constraints	*Ranking values*
NONFINAL	114.290
AFR	108.639
FOOTBIN	104.784
WSP	104.476
TROCHAIC	104.470
IAMBIC	101.302
AFL	100.739
MAIN-R	99.521
MAIN-L	95.039
WFR	85.710
PARSE	82.381
WFL	82.127

Table 2. A successful GLA learner, after 10,000 data.

We will now explain to what forms this ranking leads in production. The top ranking of NONFINAL leads to *final-syllable extrametricality*: all winners have a final unfooted syllable whose weight does not influence foot structure at all. The disyllables therefore become /(L1) X/ and /(H1) X/, where 'X' stands for any final syllable. High-ranked AFR will now make sure that every foot of every word will end after the penultimate syllable. This means that there will only be a single foot in every word, one that ends just before the extrametrical syllable. In forms of more than two syllables, the high ranking of FOOTBIN will make sure that if the penultimate syllable is light, the antepenultimate syllable will be included in the foot. If this antepenultimate syllable is heavy, WSP will make sure that the stress falls on it: /... (H1 L) X/; if it is light, TROCHAIC will make sure that the stress falls on it: /... (L1 L) X/. The situation becomes slightly complicated when we turn to forms ending in |...H X|. Of the three forms /... (L H1) X/, /... (L1 H) X/, and /... L (H1) X/, all of which satisfy FOOTBIN, only the last one satisfies both WSP and TROCHAIC, so it wins. For |...H H X| the relevant candidates are /... (H H1) X/, /... (H1 H) X/, and /... H (H1) X/. All three are equal as far as FOOTBIN and WSP are concerned, and the last two satisfy TROCHAIC. The decision between these two will have to be brought by IAMBIC; see tableau 5.

Figure 1 shows which of the rankings in tableau 5 are crucial (ignoring the four less interesting and low-ranked alignment constraints WFL, WFR, MAIN-L, and MAIN-R). The rankings not marked by lines in this figure are not fixed. Thus, TROCHAIC could be ranked anywhere between the very top and a position below WSP, as long as it outranks IAMBIC; FOOTBIN could be ranked above AFR or below WSP, as long as it is ranked below NONFINAL and above IAMBIC; and so on.

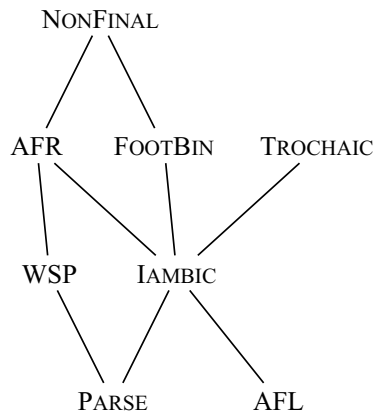

Figure 1. The crucial rankings for main-stress-only Latin.

The resulting ranking in figure 1 is close to a ranking that can be found in Jacobs (2000): TROCHAIC >> NONFINAL >> FOOTBIN >> LFR >> WSP >> PARSE, where LFR is a less-gradient version of AFR (it counts the number of syllables from the *last* foot to the end of the word). The crucial difference is our insertion of IAMBIC into the hierarchy. This is required to account for the |... H H X| forms, which Jacobs did not consider in his analysis. Figure 1, ultimately arrived at as a result of computer simulations, thus presents a correction to the analyses found in the literature (as Apoussidou & Boersma 2003 show, the analysis by Prince & Smolensky 1993 with a different constraint set suffers from the same lacuna as Jacobs' analysis). We have to admit that it is counter-intuitive that in a language with exclusively trochaic feet, it is the constraint IAMBIC that has to fix the analysis.

|H H L|	NONFINAL	AFR	FOOTBIN	WSP	TROCHAIC	IAMBIC	AFL	MAIN-R	MAIN-L	WFR	PARSE	WFL
/(H1) H L/		**!		*				**		*	**	
/(H1 H) L/		*		*		*!		*		*	*	
☞ /H (H1) L/		*		*			*	*	*	*	**	*
/H (H1 L)/	*!			*			*		*		*	*
/(H H1) L/		*		*	*!			*		*	*	
/(H1) (H2) L/		**!*					*	**		*	*	
/(H2) (H1) L/		**!*					*	*	*	*	*	

Tableau 5. A constraint hierarchy that works for all Latin forms.

Generalization to unheard forms

We have not trained our learners with forms of five syllables, but we can nevertheless run the 32 possible forms with five syllables through their respective tableaus and see what happens. All forms were handled correctly, e.g. /L L (L1 L) H/, /H L L (H1) L/, /L H (H1 L) L/, and /H H H (H1) H/. The forms with six and seven syllables are /L L L (L1 L) L/ and /L L L L (L1 L) L/. Thus, the generalization to longer forms has succeeded.

Conclusion

It has turned out to be possible to learn Latin stress with the limited set of constraints that many Optimality-Theoretic phonologists nowadays tend to regard as universal (i.e. cross-linguistically valid) as a result of years of typological research on many different stress systems (e.g. Hayes 1995). We tested our virtual learners on two *online* learning algorithms, whose only memory of past events is indirectly and concisely stored in the ranking values of the constraints: Error-Driven Constraint Demotion (EDCD) and the Gradual Learning Algorithm (GLA). Only the GLA learners turned out to succeed, while the EDCD learners failed. This is good news, since the GLA has been shown earlier to be a more realistic ingredient of human language acquisition than EDCD: like real children, GLA learners learn gradually rather than abruptly, thus showing realistic gradual learning curves and realistic effects of the distributions of forms in the language data (Boersma and Levelt 2000; Curtin and Zuraw 2001); GLA learning is robust against modest levels of errors in the language data (Boersma 1998); the GLA is capable of handling continuous input data, like auditory cues in L1 and L2 perception (Escudero and Boersma 2003; to appear); and the GLA has been able to model language change induced by bidirectional language acquisition (Jäger 2003). Nevertheless, neither EDCD nor the GLA are capable of learning every metrical system predicted by *factorial typology*, i.e. every metrical system that results from a permutation of the rankings of our twelve constraints (Boersma, to appear). Both learning algorithms fail for some rankings, but the rankings for which the two fail are different. If a learning algorithm fails precisely for those rankings that do not correspond to any existing language, this should be regarded as positive evidence for the appropriateness of such a learning algorithm for the description of real language acquisition. For the case discussed in this paper, the results provide direct evidence against the appropriateness of EDCD and are compatible with a possible appropriateness of the GLA. More languages and, especially, gaps in factorial typology (i.e. expected but non-existent languages) need to be investigated before we can conclude that any Optimality-Theoretic learning algorithm provides the appropriate model for the acquisition of language.

Acknowledgments

This work was supported by an NWO grant awarded to P. Boersma for the Vernieuwingsimpuls project 'Adequacy and acquisition of functional constraint grammars'.

References

Allen, W. S. 1973. *Vox Latina. The Pronunciation of Classical Latin*. Second edition. Cambridge University Press.

Apoussidou, D. and Boersma, P. 2003. The Learnability of Latin Stress. In Proceedings of the Institute of Phonetic Sciences 25: 101–148. University of Amsterdam.

Boersma, P. 1997. How we learn Variation, Optionality, and Probability. Proceedings of the Institute of Phonetic Sciences 21: 43–58. University of Amsterdam.

Boersma, P. 1998. Functional Phonology: Formalizing the Interactions between Articulatory and Perceptual Drives. Ph.D. diss., Institute of Phonetic Sciences, University of Amsterdam.

Boersma, P. Review of Tesar and Smolensky 2000: Learnability in Optimality Theory. To appear in *Phonology*.

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

Boersma, P. and Levelt, C. 2000. Gradual Constraint-Ranking Learning Algorithm Predicts Acquisition Order. In E. Clark ed., *Proceedings of the Thirtieth Annual Child Language Research Forum,* 229–237. Stanford, Calif.: CSLI.

Curtin, S. and Zuraw, K. 2001. Explaining Constraint Demotion in a Developing System. In B. Skarabela, S. Fish, and A. H.-J. Do eds., *Proceedings of the 26th Annual Boston University Conference on Language Development*, 118–129. Somerville, Mass.: Cascadilla Press.

Escudero, P. and Boersma, P. 2003. Modelling the Perceptual Development of Phonological Contrasts with Optimality Theory and the Gradual Learning Algorithm. In S. Arunachalam, E. Kaiser, and A. Williams eds., Proceedings of the 25th Annual Penn Linguistics Colloquium (Penn Working Papers in Linguistics 8.1), 71–85.

Escudero, P. and Boersma, P. Bridging the Gap between L2 Speech Perception Research and Phonological Theory. To appear in *Studies in Second Language Acquisition*.

Hayes, B. 1995. *Metrical Stress Theory*. University of Chicago Press.

Jacobs, H. 2000. The Revenge of the Uneven Trochee: Latin Main Stress, Metrical Constituency, Stress-Related Phenomena and OT. In A. Lahiri ed., *Analogy, Levelling, Markedness*, 333–352. Berlin: Mouton de Gruyter.

Jäger, G. 2003. Learning Constraint Sub-Hierarchies. The Bidirectional Gradual Learning Algorithm. In R. Blutner and H. Zeevat eds., *Pragmatics in OT*. Basingstoke: Palgrave McMillan.

Prince, A. and Smolensky, P. 1993. *Optimality Theory: Constraint Interaction in Generative Grammar*. Ms., Rutgers University, New Jersey.

Tesar, B. 1995. *Computational Optimality Theory*. Ph.D. diss., University of Colorado.

Tesar, B. and Smolensky, P. 1998. Learnability in Optimality Theory. *Linguistic Inquiry* 29:229–268.

Tesar, B. and Smolensky, P. 2000. *Learnability in Optimality Theory*. MIT Press.