4 The learnability of grammatical stress in Latin[20]

4.1 Introduction

This section compares the performance of the two reranking strategies Constraint Demotion (CD) and the Gradual Learning Algorithm (GLA) for the metrical stress system of Classical Latin. It turns out that the GLA has a higher success rate than CD when learning from overt forms. This suggests that the GLA may be a better model of acquisition than CD. The results also provide evidence for the discussion in the literature about what the correct linguistic analysis of Latin stress is: if overt forms contain main stress only, the GLA makes the child posit an analysis that makes use of uneven trochees (like the analysis by Jacobs 2000) rather than strictly bimoraic trochees (as suggested by Prince & Smolensky, Mester 1994 and Hayes 1995).

To linguists, learnability theory is about creating formal models of language acquisition, i.e. it investigates what precisely is known by the beginning learner and how precisely the learner proceeds from this initial state to an adult state on the basis of language input. While the universality of constraints could be questioned in general, it is assumed in this book that at least the structural constraints that handle metrical phonology are the same in all languages. As we will see, this opens up the possibility that a descriptively simple metrical system, like that of Latin, turns out to be surprisingly complicated when described in terms of constraints proposed by linguists on the basis of cross-linguistic typology rather than in terms of constraints tailored to the specific language at hand.

---

[20] Sections 4.1 to 4.6, and 4.10 appeared in Diana Apoussidou & Paul Boersma (2003): The Learnability of Latin stress, *IFA-Proceedings* 25:100-148. Sections 4.7 and 4.8 appeared in Diana Apoussidou & Paul Boersma (2004): Comparing different Optimality-Theoretic learning algorithms: the case of metrical phonology. *Papers from the AAAI Spring Symposium*, Technical Report SS-04-05, 1-7. Section 4.9 appeared in Diana Apoussidou & Paul Boersma (2004): Comparing different Optimality-Theoretic learning algorithms for Latin stress. *Proceedings of the 23rd West Coast Conference on Formal Linguistics*, 29-42.

To obtain information about the universal components of the grammar, one can analyse the actual language acquisition process of infants and children. Such an analysis is quite difficult in the case of phonological perception, since we cannot look inside a speaker's head to see what happens during perception, and speakers themselves, children included, have very little conscious access to the perception process, let alone the capability of reliably reporting on it. The analysis is slightly less difficult in the case of the child's language *production*, since in that situation at least part of the output of the grammar can be observed directly. But even when considering produced forms, the researcher encounters hidden structures like metrical feet, which often remain ambiguous.

Another method for identifying universal aspects of the grammar is to try and simulate the acquisition process with the help of a computational learning algorithm. In that way, the universal principles derived from language acquisition data can be tested with respect to their adequacy. To make this work, a learning algorithm needs to be supplied with the universal ingredients of the grammar, which in the case of OT means that the learning algorithm should be supplied with a universal set of constraints. Simulating learnability has a further benefit for linguistics, namely providing evidence for or against existing analyses in the literature. By means of a learning algorithm that is based on OT, existing OT analyses of a language can be tested with respect to their learnability. If it turns out that an analysis proposed in the literature is not learnable with a certain learning algorithm, then either this analysis or this learning algorithm should be rejected (Boersma 2003).

In this chapter, the learnability of the metrical phonology of Latin word stress is tested. Taking a dead language as the test subject is not as awkward as it may look. Nowadays there are no native speakers of Latin that can tell us how it was pronounced originally; thus, no phonetic analysis is available. Still, its prosodic system is at least partly accessible through analyses of written text such as poems or language descriptions of contemporary witnesses. The decision fell on the stress system of Latin as the test subject because it has been studied by linguists at great length. Latin is often taken as the prototypical example in general studies on metrical phonology when it comes to illustrating phenomena like weight-sensitivity and extrametricality (e.g. Allen 1973b; Hayes 1985, 1987; McCarthy & Prince 1986; Prince 1990). Many of the principles found in Latin word stress have fed ideas about universal constraint sets for metrical phonology in general, and have been used to analyse other languages. In turn, Latin has been analysed with constraints whose cross-linguistic validity has been established in analyses of other languages. Since there exist several OT analyses of Latin word stress, they can be compared with each other with respect to their learnability. We will also see whether different sets of data make a difference in learning, to be able to determine the amount of information needed for a successful simulation. In addition, the simulations are run with two OT based learning algorithms that differ with respect to their way of constraint reranking during the acquisition process.

The chapter is structured as follows. Section 4.2 provides a description of Latin word stress, and outlines the various analyses proposed for Latin. Section 4.3 to 4.5 outline the computer simulations of the acquisition of Latin primary stress with the two learning algorithms. Section 4.6 presents the results of the primary stress simulations, showing that in several respects the GLA performs better than CD. Sections 4.7 to 4.9 provide further simulations of primary and secondary stress. Sections 4.10 and 4.11 place the findings in a larger perspective and discuss their implications for learnability theory and OT.

4.2 Latin main stress

As already pointed out in chapter 2, an important constituent for assigning stress to words is the foot. By causing a rhythmic organization of syllables, the foot underlies the metrical patterns of many languages. Feet are usually *binary*, i.e., they group syllables into pairs, resulting in a pattern of (often alternating) weak and strong syllables. But how does a learner find out that the feet in her language are strong-weak sequences (trochaic) rather than weak-strong sequences (iambic) in the absence of other phonetic cues like e.g. iambic lengthening? The way that syllables are grouped within a word will have an effect on how other words in the language are stressed. The problem is that foot structure belongs to the surface representation and is not contained in the overt form that a learner is actually exposed to. Therefore the learner has to find out by herself whether the syllables in her language group together as trochaic or as iambic feet.

As indicated above, feet are usually binary. However, not all languages count syllables only; some count moras as well. A mora is a smaller unit than a syllable, and determines the *weight* of the syllable: syllables with a long vowel or a diphthong contain two moras (they are *heavy*), while syllables with only a short vowel contain only one mora (they are *light*). Depending on the language, syllables that end in a consonant can also count as two moras. Languages in which the number of moras in a syllable influences stress (or other phonological phenomena) are called *weight*- or *quantity-sensitive*. In such languages, heavy syllables tend to be prominent in the output. In OT, this principle is captured in the constraint WEIGHT-TO-STRESS-PRINCIPLE (WSP), as mentioned in (14) and repeated in (40):

(40) WSP: Heavy syllables are stressed.

In quantity-sensitive languages a foot ideally consists of two moras: either two light syllables or one heavy syllable, as illustrated in (4) and (5) of chapter 2. In quantity-insensitive languages, feet ideally consist of two syllables, regardless of their inner construction, as illustrated in (6) of chapter 2. This binarity is expressed in OT as a constraint FTBIN, defined in (41).

(41) FTBIN: Feet are binary on some level of analysis (mora or syllable).

Note that this constraint allows a foot to consist of three or four moras, as long as these moras are contained in a sequence of two syllables (heavy-light, light-heavy, or heavy-heavy).

Stress assignment is not only determined by foot-internal structure, but by the placement of feet within the word (or phrase) as well. Especially for longer words, the question is at what edge of the word the foot (or the feet) will be constructed. Some languages tend to build feet at or from the left edge of a word (i.e. word-initially), others at or from the right edge (word-finally). We have already met two constraints for foot placement in chapter 2 (AFR and AFL). Further relevant constraints for Latin are discussed in the course of this chapter.

In Latin, stress is handled purely by the grammar: the foot structure of a word is predictable from the syllable structure of the word, and the mental lexicon need not contain any information about where in the word the stress is realized. Thus, a learner of Latin does not have to take into account the complexities that would arise if the language had lexically assigned stress as well. This should make it relatively easy for a learner to figure out the ranking of the relevant constraints. One would think.

Basically, Classical Latin has left-prominent feet (trochees), it is weight-sensitive, and the last syllable in a word is extrametrical (i.e., it never receives stress except if it is the only syllable of a word) (Allen 1978). Syllables ending in a short vowel are light (abbreviated here as ‘L’), while syllables with long vowels or diphthongs and syllables that end in a consonant are heavy (‘H’). In words with three or more syllables, the penultimate syllable is stressed if it is heavy. If the penultimate syllable is light, the antepenultimate syllable is stressed, regardless of its weight. In words with only one or two syllables, the leftmost syllable is stressed. Some examples are given in (42). The second column represents overt forms. They include phonetic representations with stress (´) and vowel length (ː), enriched with some hidden phonological structure (periods indicate syllable boundaries) but without foot structure. The third column represents these overt forms without segmental information, i.e. the overt stress patterns (‘1’ for main stress). The issue of secondary stress is ignored at this point.

(42) Weight and stress in Latin

a.	*amice* ‘friend’	[a.míː.ke]	[L H1 L]
b.	*rapiditas* ‘speed’	[ra.pí.di.taːs]	[L L1 L H]
c.	*misericordia* ‘pity’	[mi.se.ri.kór.di.a]	[L L L H1 L L]
d.	*perfectus* ‘perfect’	[per.fék.tus]	[H H1 H]
e.	*incipio* ‘I begin’	[iŋ.kí.pi.oː]	[H L1 L H]
f.	*domesticus* ‘domestic’	[do.més.ti.kus]	[L H1 L H]
g.	*homo* ‘man’	[hó.moː]	[L1 H]

As pointed out above, there is some discussion about the details of Classical Latin stress. The different analyses agree on the minimum size of a trochaic foot (two moras), but not on its maximum size. This discussion especially applies to words with three or more syllables; in words with two syllables such as (40g), *homo*, one is bound to either analyze it with a monomoraic foot with a light syllable plus an extrametrical syllable /(L1) H/ or with a trimoraic foot /(L1 H)/. According to some (e.g. Mester 1994, Prince & Smolensky 1993, Hayes 1995), weight-sensitive feet are strictly bimoraic, while according to others (e.g. Hayes 1981, Jacobs 2000), trochees in Latin can be uneven, i.e. consist of up to three moras. The following section describes the analyses in general, as well as the translation of the problem into OT terms.

4.2.1 Linguistic analyses of Latin stress

The examples in (42) are often analysed as resulting from a combination of extrametricality and right-aligned feet. If we ignore for now the possibility of secondary stress, a recipe by Hayes (1995) will add foot structure to the forms in (42), ending up with the full surface structures /L (H1) L/, /L (L1 L) H/, /L L L (H1) L L/, /H (H1) H/, /H (L1 L) H/, /L (H1) L H/, and /(L1) H/. The recipe goes as follows: first make the last syllable extrametrical, i.e. mark it for not being able to be incorporated into a foot, then create a foot as far to the right as possible; this foot has to be bimoraic, and if the foot is disyllabic stress falls on the first syllable. Hayes' (1995) bimoraic analysis is not uncontroversial. A different approach is to propose that the foot always ends just before the extrametrical syllable (Hayes 1981). The result differs from the bimoraic analysis in two forms in (42): it leads to /L L L (H1 L) L/ (*misericordia*) and /L (H1 L) H/ (*domesticus*). The foot (H1 L) has three moras. Hayes (1995) calls it an *uneven trochee*. Such an analysis satisfies the generalized principle of foot binarity: feet consist either of two moras or of two syllables (as introduced in section 2.3).

The choice between the bimoraic analysis and the uneven trochee analysis cannot be made on the basis of the overt stress patterns alone. Some linguists have voted for a strict bimoraic approach in Latin, on the basis of non-stress evidence like iambic and cretic shortening processes (Mester 1994; Hayes 1987, 1995; Kager 1993; McCarthy & Prince 1986, 1990). According to these authors, an unfooted non-final syllable, like *ti* in /do (més) ti.kus/, is better than an uneven trochee, as in /do (més.ti) kus/ (Hayes 1995:91). Iambic and cretic shortening will not be modelled in the simulated language data fed to the child, because they were optional processes. It will be left to the simulated child to construct either a bimoraic or an uneven trochee analysis (though see §4.10 for a brief discussion of iambic shortening).

Disyllabic words cause several complications. The underlying sequence |L L| is pronounced as the overt form [L1 L]. The question is whether it should be footed as /(L1) L/, violating bimoraicity and foot binarity, or as /(L1 L)/, violating extrametricality. Likewise, |L H| is pronounced [L1 H], and the question is whether it is footed as /(L1) H/, violating foot binarity and bimoraicity, or as /(L1 H)/, violating bimoraicity and extrametricality. These sound like questions about the ranking of constraints, so it is natural to express all these conflicting principles in constraints. The ones often seen in the literature are those in (41) and (43) (a specific constraint for foot bimoraicity will be introduced later).

(43) NONFINAL: The last syllable is not contained in a foot.

There have been several proposals for the foot structure of underlying |L L| words in Latin. Prince (1980) and McCarthy & Prince (1986) argue that the structure must be /(L1 L)/. The argument runs as follows. Latin has a so-called *minimal word* requirement: there are no monosyllabic words in Latin that consist of only a light syllable. This observation can be explained by a combination of two requirements: every word must contain at least one foot, and Latin satisfies a ban on degenerate feet, i.e., the (L1) foot is prohibited completely from Latin surface structure. Apart from ruling out monomoraic words, these requirements also demand that words consisting of two light syllables must incorporate the final syllable into the foot: /(L1 L)/. Expressed in a constraint ranking, this would mean that FTBIN would have to outrank NONFINAL (Prince & Smolensky 1993). Such a ranking also predicts that |L H| is footed as /(L1 H)/, because a form with a degenerate foot /(L1) H/ would violate high-ranking FTBIN. Prince & Smolensky (1993:63) abandon feet with the form (H1 L) because this form is "marked or even absent in trochaic systems" (they refer to Hayes 1987, Prince 1990, and Mester 1994); they formulate this as the constraint *(HL) or RHYTHMICHARMONY.

The foot in Latin thus ideally consists of two moras. The analyses also agree that feet containing four moras, like (H1 H), are forbidden in Latin. Jacobs (2000) accepts the uneven trochee (H1 L), but abandons (L1 H) feet.

4.2.2 Latin Stress in OT

For the simulations of Latin main stress the same underlying forms, candidate generator, and set of constraints are used as Tesar & Smolensky (2000) did in their simulations of 124 types of languages with metrical stress. To accommodate analyses voting for uneven trochees, as in Jacobs (2000), and for moraic trochees, as in Prince & Smolensky (1993), some constraint sets are investigated that are slight modifications of the Tesar & Smolensky set, to see whether and how Latin stress can be learned. In total, six different constraint sets are considered, but with the same underlying forms and generator. In the following these ingredients are discussed in detail.

4.2.2.1 Underlying forms

With Tesar & Smolensky (2000), underlying forms are considered that consist of two to seven syllables. For the forms with two to five syllables, all possible sequences of heavy and light syllables are taken into account. Thus, the underlying disyllables are |L L|, |L H|, |H L|, and |H H|. Likewise, there are eight trisyllabic underlying forms: |L L L|, |L L H|, |L H L|, |L H H|, |H L L|, |H L H|, |H H L|, and |H H H|. In the same vein, there are 16 forms with four syllables, and 32 with five. For the forms with six or seven syllables, the ones with heavy syllables are ignored (for computational reasons that can be deduced from table (68), column 6), thus leaving only |L L L L L L| and |L L L L L L L|. In total, therefore, there are 62 different underlying forms, the same ones that Tesar & Smolensky used. Unlike Tesar & Smolensky, who taught the learners all 62 possible overt forms in their simulations, the present learners are taught only the 28 forms (i.e. 28 underlying-surface pairs or 28 overt forms) that have maximally four syllables; if a learner then arrives at a grammar appropriate for these 28 forms, we can have a look at how she generalizes this grammar to the 34 forms that consist of five syllables or more.

4.2.2.2 The candidate generator

In production, each of the 62 underlying forms comes with a tableau. For each of the 62 tableaux, the candidate set (GEN) is restricted in the same way as in Tesar & Smolensky (2000). GEN creates surface forms that meet the following criteria: the sequence of syllables is identical to the sequence of syllables in the underlying form with respect to number, weight, and order; every foot contains exactly one primary-stressed or secondary-stressed syllable; every word contains exactly one foot that contains a primary-stressed syllable; every primary-stressed or secondary-stressed syllable is contained in a foot; no foot contains more than two syllables. For each of the four disyllabic underlying forms, there are six candidates for the surface form. For instance, an underlying |H L| has the following candidates: /(H1) L/, /(H1 L)/, /(H1)(L2)/, /H (L1)/, /(H L1)/, /(H2)(L1)/. Each of the 8 underlying trisyllables has 24 candidates, for instance |H L L| has the candidates /(H1) L L/, /(H1 L) L/, /(H1) L (L2)/, /H (L1) L/, /H L (L1)/, /H (L1 L)/, /(H L1) L/, /H (L L1)/, /(H1)(L L2)/, /(H1 L)(L2)/, /(H1)(L2) L/, /(H1)(L2 L)/, /H (L1)(L2)/, /(H L1)(L2)/, /(H2)(L1) L/, /(H2)(L1 L)/, /H (L2)(L1)/, /(H L2)(L1)/, /(H2) L (L1)/, /(H2)(L L1)/, /(H2 L)(L1)/, /(H2)(L1)(L2)/, /(H1)(L2)(L2)/, and /(H2)(L2)(L1)/. The 16 forms with four syllables have 88 candidates each; the 32 forms with five syllables have 300 candidates each. The single form with six syllables has 984 candidates, the form with seven syllables has 3136.[21]

4.2.2.3 The constraints

The basic constraint set used in the simulations is the one adopted by Tesar & Smolensky (2000), as listed in (14) of section 2.6.1. This constraint set takes into account the restrictions enforced by the generator: since the generator does not generate candidates with trisyllabic feet, we need no constraints against trisyllabic feet;[22] and since the generator does not generate candidates without main stress, we need no constraints to enforce that every word should contain at least one stress (such as the constraint LEX≈PR by Prince & Smolensky 1993).

Apart from the set of 12 constraints that Tesar & Smolensky used (the *T&S set*), five slightly different constraint sets are investigated. These sets additionally involve the three constraints listed in (44).

(44) Additional/alternative constraints for Latin stress
TROCHAIC: The leftmost syllable in a foot is the head syllable.
HEADNONFINAL: The head foot is not aligned with the right edge of the word, and the head syllable is not the last syllable in the word.
FOOTBIMORAIC: Each foot must be bimoraic.

In order to replicate the idea behind the uneven-trochee analysis of Jacobs (2000), the *uneven trochee constraint set* is defined. This set of 12

---

[21] The monosyllabic underlying form |H| has only a single output candidate: /(H1)/. This makes it impossible for the learner to learn anything from such a form. This is the reason why monosyllabic forms are not considered in any of the tableaux or simulations.

[22] This means that the only reason for including FTBIN is that it militates against monomoraic feet.

constraints (Jacobs himself used only six) is similar to the T&S set, but FTNONFIN is replaced with the constraint TROCHAIC that we met with earlier in (8) of section 2.5. The formulation mirrors that of IAMBIC. The difference with the T&S set is that degenerate feet like (L2) violate FTNONFIN, but not TROCHAIC. In the uneven-trochee constraint set,[23] the constraints TROCHAIC and IAMBIC conspire to minimize the number of syllables in a foot, since monosyllabic feet violate neither.

The third constraint set is based on the moraic-trochee analysis by Prince & Smolensky (1993). Like the uneven-trochee set, this *moraic-trochee constraint set* contains TROCHAIC rather than FTNONFIN. Moreover, the constraint NONFINAL is replaced with HEADNONFINAL[24], which demands that neither the head syllable of a foot nor the head foot itself are in word-final position. Both of these conditions can assign a violation mark: HEADNONFINAL (HDNONFIN) is violated once in the form /(H2)(L1 L)/, twice in /(H2)(L L1)/, and not at all in /(H1)(L L2)/.

Since FTBIN does not distinguish between disyllabic and bimoraic feet, it might be worthwhile to investigate the workings of an explicit bimoraic analysis. To this end, three more constraint sets are considered, all consisting of 13 constraints. These sets were constructed by adding to the T&S, uneven trochee, and moraic trochee sets (all of which contain 12 constraints) a straightforward constraint FOOTBIMORAIC (as similarly proposed in the pre-OT approach of Kager 1993a and in the OT-approach of Hewitt 1994). Other than FTBIN, this constraint is assigned a violation for every monomoraic foot such as (L1) or (L2), every trimoraic foot such as (H1 L) or (L2 H), and every quadrimoraic foot such as (H H1).

4.2.3 Assessment of Jacobs' OT analysis of Latin stress

Jacobs (2000) prefers the constraint NONFINAL to Prince & Smolensky's (1993) HDNONFIN because the formulation of NONFINAL is much simpler and because HDNONFIN seems to predict unattested ('quarternary') stress patterns. Jacobs' analysis can be translated to the set of constraints and candidates discussed before, by first noting that one of the constraints

---

[23] Jacobs did not actually include IAMBIC in his set. This will turn out to be crucial in §4.2.3. He also excluded most alignment constraints.

[24] Prince & Smolensky called this constraint NONFINAL, but the name is changed in order to make the meaning behind every constraint name unambiguous.

employed by Jacobs, Lx≈Pr (“Every lexical word must correspond to a prosodic word”; Prince & Smolensky 1993) is now part of the candidate generator, so that this constraint can be left out of consideration. For right alignment, Jacobs uses a constraint Last-Foot-Right “align the last foot with the word, right edge” (LFR; which comes close to Edgemost as in Prince & Smolensky 1993). This constraint is assigned one violation mark for every syllable that follows the last foot; it can thus be seen as a gradient version of WFR, and for words with a single foot it has the same number of violations as AFR and Main-R. Jacobs’ article happens to contain a ranking that handles all the forms that he considers.[25] This ranking is Trochaic >> Nonfinal >> FtBin >> LFR >> WSP >> Parse. It correctly predicts the following forms, several of which contain the uneven trochee (H1 L):

(45) Predicted forms

/(L1) L/	*fala*	/(fá) la/	‘siege tower’
/(L1) H/	*fames*	/(fá) meːs/	‘hunger’
/(H1) L/	*fama*	/(fáː) ma/	‘rumor’
/(H1) H/	*fagus*	/(fáː) gus/	‘beech’
/(L1 L) L/	*fabula*	/(fá.bu) la/	‘little bean’
/(L1 L) H/	*fragilis*	/(frá.gi) lis/	‘fragile’
/L (H1) L/	*amice*	/a (míː) ke/	‘friend’
/L (H1) H/	*facultas*	/fa (kúl) taːs/	‘opportunity’
/(H1 L) L/	*fabula*	/(fáː.bu) la/	‘story’
/(H1 L) H/	*flammulae*	/(flám.mu) lai/	‘little flames’

Tableaux (46) to (49) for the disyllabic forms illustrate the ranking of Nonfinal above the four constraints FtBin, LFR, WSP, and Parse. If Nonfinal were ranked below any of these constraints, at least one of these tableaux would have had a different winner. These four tableaux show no

---

[25] For Classical Latin, Jacobs (2003:345) actually proposes the ranking { FtBin, Trochaic } >> Nonfinal >> LFR >> WSP >> Parse. However, this ranking must be incorrect, because it would give final stress in /L (H1)/ in tableau (48), unless PARSE outranks WSP. This latter ranking confusingly occurs in the tableau on Jacobs’ page 342, so that /(L1 H)/ becomes the winning candidate; but the form /L (H1) L/ requires WSP >> Parse in order to beat /(L1 H) L/, as Jacobs notes himself and the reader can see here in tableau (51). At the very end of his article, Jacobs introduces the ranking Trochaic >> Nonfinal >> FtBin >> LFR >> WSP >> Parse for Classical Latin.

evidence for the ranking of Trochaic, nor for the relative rankings of FtBin, LFR, WSP, and Parse with respect to each other.

(46) Extrametricality beats word-foot-right (LFR) and Parse

|H L|	Trochaic	Nonfinal	FtBin	LFR	WSP	Parse
☞ /(H1) L/				*		*
/(H1 L)/		*!				
/H (L1)/		*!	*		*	*
/(H L1)/	*!	*			*	

(47) Extrametricality also beats FtBin

|L L|	Trochaic	Nonfinal	FtBin	LFR	WSP	Parse
☞ /(L1) L/			*	*		*
/(L1 L)/		*!				
/L (L1)/		*!	*			*
/(L L1)/	*!	*				

(48) Extrametricality also beats the WSP

|L H|	Trochaic	Nonfinal	FtBin	LFR	WSP	Parse
☞ /(L1) H/			*	*	*	*
/(L1 H)/		*!			*	
/L (H1)/		*!				*
/(L H1)/	*!	*				

(49) Extrametricality beats LFR, WSP, and Parse

|H H|	Trochaic	Nonfinal	FtBin	LFR	WSP	Parse
☞ /(H1) H/				*	*	*
/(H1 H)/		*!			*	
/H (H1)/		*!			*	*
/(H H1)/	*!	*			*	
/(H1)(H2)/		*!				

The tableaux for the trisyllabic forms show more detailed evidence for rankings. Tableau (50) shows evidence for the existence of TROCHAIC, which prefers /(L1 L) L/ to /(L L1) L/. But there is no evidence for the *ranking* of TROCHAIC; it could just as well be ranked at the bottom, as far as the pair |L L L| - /(L1 L) L/ is concerned. This freedom of ranking of TROCHAIC is caused, of course, by the absence of the counteracting constraint IAMBIC from the tableau. If IAMBIC had been included, the choice of /(L1 L) L/ instead of /(L L1) L/ would have been direct evidence for the ranking TROCHAIC >> IAMBIC.[26]

(50) Evidence for trochaicity

|L L L|	TROCHAIC	NONFINAL	FTBIN	LFR	WSP	PARSE
/(L1) L L/			*!	**		**
☞ /(L1 L) L/				*		*
/L (L1) L/			*!	*		**
/L (L1 L)/		*!				*
/(L L1) L/	*!			*		*

Tableaux for trisyllabic forms that end in a heavy syllable are not shown here, because the high ranking of NONFINAL ensures that L-final and H-final words are always handled in the same way. The next form to consider, then, is |L H L|. Tableau (51) shows direct evidence that PARSE is dominated by WSP as well as by TROCHAIC. If the ranking of WSP and PARSE had been reversed (with high-ranked TROCHAIC), the candidate /(L1 H) L/ would have won. It is apparently worse to have an unstressed heavy syllable than to have an unfooted light syllable. If TROCHAIC had been ranked below WSP and PARSE, the iambic candidate /(L H1) L/ would have won.

---

[26] Note that without the constraint IAMBIC, the last candidate /(L L1) L/ is harmonically bounded: it could not win under any ranking of the constraints.

(51) Weight-to-stress and trochaicity beat PARSE

|L H L|	TROCHAIC	NONFINAL	FTBIN	LFR	WSP	PARSE
/(L1) H L/			*!	**	*	**
/(L1 H) L/				*	*!	*
☞ /L (H1) L/				*		**
/L (H1 L)/		*!				*
/(L H1) L/	*!			*		*

The next underlying form to consider is |H L L|. Tableau (52) shows that the winner yields evidence for the existence of the constraints LFR or PARSE. Without these constraints, the candidate /(H1) L L/ would have been equally harmonic as /(H1 L) L/. Jacobs' constraint set thus favours the uneven trochee analysis /(H1 L) L/ over the bimoraic analysis /(H1) L L/, irrespectively of the ranking of the constraints, since the violations of /(H1) L L/ form a superset of those of the violations of /(H1 L) L/. In order to turn the bimoraic analysis /(H1) L L/ into a winner, we would need the help of an extra constraint that is ranked above LFR and PARSE, perhaps FTBIMORAIC.

(52) Evidence for LFR or PARSE

|H L L|	TROCHAIC	NONFINAL	FTBIN	LFR	WSP	PARSE
/(H1) L L/				**!		**
☞ /(H1 L) L/				*		*
/H (L1) L/			*!	*	*	**
/H (L1 L)/		*!			*	*
/(H L1) L/	*!			*	*	*

Until now, all the forms that Jacobs considers have been discussed. Absent from his article, though, is the underlying form |H H L|. Tableau (53) shows that this form is problematic.

(53) A stress clash or a superheavy foot?

|H H L|	Trochaic	Nonfinal	FtBin	LFR	WSP	Parse
/(H1) H L/				**!	*	**
/(H1 H) L/				*	*!	*
/H (H1) L/				*	*!	**
/H (H1 L)/		*!			*	*
/(H H1) L/	*!			*	*	*
☞ /(H1)(H2) L/				*		*
☞ /(H2)(H1) L/				*		*

In tableau (53), two forms with two feet are optimal. To make the form with penultimate main stress win, the constraints for the placement of main stress would have to be included and ranked in the order Main-R >> Main-L.[27] The results for the |H H L| forms generalize to quadrisyllabic and longer forms. Because of the ranking WSP >> Parse, these forms will tend to have secondary-stressed feet around every heavy syllable, and because of the presence of Parse, light syllables will tend to be footed as well if this does not create iambs. Examples are:

(54) Examples for |H H L| forms

/(L2 L) (H1) L/	*manifesta*	[mà.ni.fés.ta]	‘caught in the act’
/(H2 L) (H1) H/	*militaris*	[mìː.li.táː.ris]	‘military’
/L (H2) (L1 L) L/	*amicitia*	[a.mìː.kí.ti.a]	‘friendship’
/(H2)(H2)(H2)(H1) H/	*definitivus*	[dèː.fiː.nìː.tíː.vus]	‘definitive’
/(H2 L) L (H1) L/	*deliciosa*	[dèː.li.ki.óː.sa]	‘spoiled’

The choice between /(L2 L) L (H1 L) L/ ([mì.se.ri.kór.di.a]) and /L (L2 L)(H1 L) L/ ([mi.sè.ri.kór.di.a]) would probably have to be made by constraints such as AFL and AFR.

But it is a question whether secondary-stressed forms should be allowed at all, especially those with *stress clashes* (consecutive stressed syllables) like those in tableau (53). The correct form in tableau (53) should therefore be /H (H1) L/, with a single foot. It is possible to get rid of the two

---

[27] As long as Main-R is ranked above Main-L, it does not matter where in the hierarchy the two constraints are inserted, because they are only violated in the last two candidates, and only decide between those two.

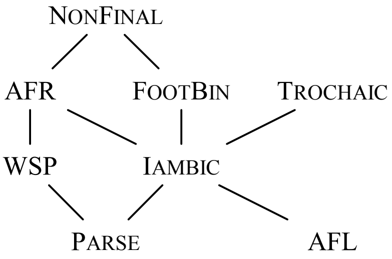
bipedal candidates in (53) by replacing LFR with AFR. This would not change anything in tableaux (46) to (52), but the two last candidates in (53) would get three violations of AFR. If it can be ensured that AFR outranks WSP, the last two candidates in (53) perish. However, the winner will now be the form /(H1 H) L/, with a superheavy foot. This form is observationally incorrect, with its antepenultimate main stress (a speaker of Latin would say [au.díː.re] 'to hear', not [aú.diː.re]); the correct form is /H (H1) L/. However, we can see from (53) that /H (H1) L/ has superset violations when compared with /(H1 H) L/. In order to make /H (H1) L/ more harmonic than /(H1 H) L/, then, an extra constraint would have to be used. An obvious choice is Iambic, and it should be ranked above Parse, as tableau (55) shows.

(55) Jacobs' hierarchy patched up

|H H L|	Trochaic	Nonfinal	FtBin	AFR	WSP	Iambic	Parse
/(H1) H L/				**!	*		**
/(H1 H) L/				*	*	*!	*
☞ /H (H1) L/				*	*		**
/H (H1 L)/		*!			*	*	*
/(H H1) L/	*!			*	*		*
/(H1)(H2) L/				**!*			*
/(H2)(H1) L/				**!*			*

Tableau (55) provides a ranking that will work for all Latin forms. It correctly generalizes to words of more than three syllables, causing all of them to end in /...(H1) X/ or /...(X1 L) X/. The remaining question is where Iambic has to be inserted into the hierarchy. According to (55), it has to outrank Parse. Obviously, it has to be ranked below Trochaic, otherwise /(L L1) L/ would be better than /(L1 L) L/; that would be observationally incorrect, since *iacere* 'to throw' is pronounced [já.ke.re], not [ja.ké.re]. Given the current set of seven constraints, and the low ranking of Parse, Iambic has to be ranked below FtBin, because /(L1 L) L/ has to be better than /L (L1) L/. Finally, AFR has to outrank both WSP and Iambic in order to make /H (L1 L) X/ (e.g. *nobilitas* [noː.bí.li.taːs] 'fame') better than /(H1) L L X/. The complete set of crucial rankings is shown in figure (56). 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; FTBIN could be ranked above AFR or below WSP, as long as it is ranked below NONFINAL and above IAMBIC; and so on.

(56) The crucial ranking for the uneven trochee analysis without secondary stress



Note the *conspiracy* of the constraints TROCHAIC and IAMBIC. Together they have a preference for monosyllabic feet, since such feet violate neither of these constraints.[28] In the ranking at hand, this monosyllabic bias is just enough to rule out /(H1 H) L/, because TROCHAIC and IAMBIC are both ranked above PARSE. The bias is not enough to rule out /(L1 L) L/ and /H (L1 L) L/, because IAMBIC is still ranked below FTBIN and AFR. This combination of requirements on the ranking of IAMBIC brings about a relatively *deep* grammar: the tree in (56) shows that four levels of constraints are needed to describe the Latin stress rule with its relatively simple formulation of “stress the penultimate if it’s heavy, else the antepenultimate”.

4.2.4 Assessment of the Tesar & Smolensky constraint set for Latin stress

The effects of the bias of TROCHAIC and IAMBIC for short feet is not found for Tesar & Smolensky’s combination of FTNONFIN and IAMBIC. This is because these two constraints have complementary violations on the foot level: monosyllabic and iambic feet, i.e. (X1) and (X X1) (where ‘X’ stands for either ‘L’ or ‘H’), violate FTNONFIN, while disyllabic trochaic feet, i.e.

---

[28] This is problematic, since factorial typology predicts languages with monosyllabic feet only (René Kager, p.c.).

(X1 X), violate IAMBIC. Thus, the sum of the number of violations of FTNONFIN and IAMBIC is equal to the number of feet in the word. This means that these two constraints together still have the side effect of minimizing the number of feet in a word, but they are not capable of forcing a specific foot form in the way TROCHAIC and IAMBIC could. The reduced power of FTNONFIN as compared to TROCHAIC turns out to make it impossible for Tesar & Smolensky's set of 12 constraints to handle the facts of Classical Latin stress in the uneven trochee analysis without secondary stress. This will become evident in sections §4.6.1 and 4.6.2.

4.2.5 Assessment of a moraic-trochee analysis of Latin stress

Let us now turn to an analysis of Latin stress with moraic trochees. On grounds of optional processes such as iambic and cretic shortening, Mester (1994) argued that Latin employed moraic trochees, and that words with a final H-L-H sequence must be parsed as /(H1) L H/, and not as /(H1 L) H)/, because in the optional shortening process, the final heavy syllable becomes light, and is parsed as /(H1) (L L)/. If the unshortened form had been parsed as /(H1 L) H)/, the shortened form with two feet /(H1) (L L)/ would violate the *Free Element Condition* (Prince 1985, Steriade 1988, Halle & Kenstowicz 1991), which states that newly built prosodic structure cannot overwrite previously established structure. Prince & Smolensky (1993) gave an OT account for an analysis with moraic trochees in Latin. Next to FTBIN, PARSE, TROCHAIC and WSP, they used an extrametricality constraint NONFINAL and an alignment constraint EDGEMOST. Their NONFINAL constraint will be called HEADNONFINAL (HDNONFINAL) here, in order to distinguish it from the NONFINAL constraint in (14). While NONFINAL requires that the final syllable is not included in a foot, HDNONFINAL requires that no head (be it head syllable or head foot) is final. The difference is that in e.g. a candidate like /(H1) (L2)/, NONFINAL is violated, but HDNONFINAL is not. EDGEMOST is a predecessor of the alignment constraints in (14) and requires that the head syllable of a prosodic word needs to be aligned with the right edge. RHYTHMICHARMONY (or *(HL) in short) rules out candidates with (HL)-feet.

(57) Prince and Smolensky's (1993) additional constraints:
HdNonfinal: No head of PrWd is final in PrWd.
Edgemost (σ', R): The head syllable is situated at the right edge of PrWd.
RhHarmony/*(HL): Final elements in constituents are long.

Prince & Smolensky (1993) arrived at the following ranking:

(58) A ranking for a moraic-trochee analysis

FtBin, Trochaic, *(HL)
|
HdNonFin
|
Edgemost, WSP
|
Parse

With this ranking, feet are bimoraic, but only where possible: a form like [á.moː], with a light and a heavy syllable, is parsed as a disyllabic foot /(L1 H)/, because the competing candidate /(L1) H/ is ruled out by high-ranked FtBin:[29]

(59) Disyllabic feet in disyllabic words

|L H|	FtBin	Trochaic	HdNonfinal	WSP	Edgem	Parse
/(L1) H/	*!			*	*	*
☞ /(L1 H)/			*	*	*	

According to Prince & Smolensky (1993), [H L] words in Latin are parsed into /(H1) L/ feet; (H1 L) feet are ruled out by *(HL):

---

[29] The interested reader is kindly referred to Prince & Smolensky (1993) for a more detailed analysis.

(60) Stress on the penultimate syllable

|H L|	FTBIN	TROCHAIC	HDNONFINAL	WSP	EDGEMOST	PARSE
☞ /(H1) L/					*	*
/(H1 L)/			*!		*	

For [H1 L L] sequences, Prince & Smolensky (1993) suggest that they are parsed as /(H1) (L L)/, and [H H1 L] sequences as /(H) (H1) L/. They do not mention whether the foot preceding or following the stressed syllable carries a secondary stress or not; their representation implies that these feet are stressless.

(61) Stress on the antepenultimate syllable

|H L L|	FTBIN	TROCHAIC	HDNONFINAL	WSP	EDGEMOST	PARSE
/(H1) L L/					**	*!*
/(H1 L) L/					**	*!
/H (L1 L)/			*!	*	*	*
☞ /(H1)(L L)/					**	

To enable a comparison between a moraic-trochee analysis of Latin stress, as proposed by Mester (1994) and Prince & Smolensky (1993), with an uneven-trochee analysis, as e.g. proposed by Jacobs (2000), I will give an account for the moraic trochee analysis with some of the constraints listed in (14). In line with Prince & Smolensky (1993), the approach will include the constraints FTBIN, PARSE, TROCHAIC and WSP. Departing from the constraint set of Prince & Smolensky (1993), the constraint EDGEMOST is replaced with the constraint AFR, as established in (14). The performance of Prince & Smolensky's extrametricality constraint HDNONFINAL will be compared with the constraint NONFINAL as defined in (14). The straightforward constraint *(HL) is discarded. The resulting constraint set is given in (62).

(62) A new constraint set for the moraic-trochee analysis
AFR
FtBin
HdNonfinal
Parse
Trochaic
WSP

I further deviate from the account given in Prince & Smolensky by only considering feet with stress: where there is a stress, there is a foot, and where there is a foot, there is a stress. For this reason, stressless feet that occur to the right of the head foot are not included in the candidate set.

The moraic-trochee analysis is given in the following, adopting a ranking as closest possible to the one given in (58). For a form like [á.moː], this leads to the incorporation of both syllables into a foot, as shown in tableau (63). A candidate /L (H1)/ crucially violates HdNonfinal twice since both the head foot and the head syllable are word-final.

(63) Exhaustive parsing into a ‘wretched’ trochee

|L H|	FtBin	Trochaic	HdNonfinal	WSP	AFR	Parse
/(L1) H/	*!			*	*	*
☞ /(L1 H)/			*	*		
/L (H1)/			*!*			*
/(L H1)/		*!	**			

Since HdNonfinal is fine with secondary stress in final syllables, such analyses must pop up, as shown in tableau (64), unless prevented by a constraints against stress clashes.

(64) Stress clash in a bisyllabic word

|H H|	FtBin	Trochaic	HdNonfin	WSP	AFR	Parse
/(H1) H/				*(!)	*(!)	*
/(H1 H)/			*!	*		
/H (H1)/			*!*	*		*
/(H H1)/		*!	**	*		
☞ /(H1)(H2)/						

Analogously, this analysis predicts forms like /(L1 L) (H2)/ for light-light-heavy sequences.

Sequences ending in |…H L L| would result in trimoraic feet /(H1 L) L/, though. This is shown in tableau (65). The candidate with the bimoraic foot, /(H1) L L/, is ruled out by an additional violation of AFR.

(65) An uneven trochee, again

|H L L|	FtBin	Trochaic	HdNonfin	WSP	AFR	Parse
/(H1) L L/					**!	**
☞ /(H1 L) L/					*	*
/H (L1) L/	*!			*	*	**
/H (L1 L)/			*!	*		*
/(H L1) L/		*!		*	*	*
/(H2)(L1 L)/			*!		**	
/(H1)(L2 L)/			*!		**	

But like Jacobs' analysis, the adopted constraint set and ranking predicts a wrong result for underlying |H H L|, shown in tableau (66): if WSP and AFR are crucially tied so that their violation marks add up and the buck is passed to Parse, or if WSP outranks AFR, /(H1)(H2 L)/ wins.

(66) Main stress on the antepenultimate despite a heavy penult

|H H L|	FtBin	Trochaic	HdNonfin	WSP	AFR	Parse
/(H1) H L/				*(!)	**(!)	**
/(H1 H) L/				*	*	*!
/H (H1) L/				*	*	*!*
/H (H1 L)/			*!	*		*
/(H H1) L/		*!		*	*	*
/(H1)(H2) L/					***!	*
/(H2)(H1) L/					***!	*
☞ /(H1)(H2 L)/					**	
/(H2)(H1 L)/			*!		**	

If, on the other hand, AFR outranks WSP, candidate /(H1 H) L/ wins. As with Jacobs' analysis, the analysis can be saved by assuming that AFR outranks WSP and by inserting Iambic into the hierarchy, as (67) shows.

(67) Moraic-trochee analysis patched up

|H H L|	FtBin	Trochaic	HdNonfin	AFR	WSP	Iambic	Parse
/(H1) H L/				**!	*		**
/(H1 H) L/				*	*	*!	*
☞ /H (H1) L/				*	*		**
/H (H1 L)/			*!		*	*	*
/(H H1) L/		*!		*	*		*
/(H1)(H2) L/				**!*			*
/(H2)(H1) L/				**!*			*
/(H1)(H2 L)/				**!		*	
/(H2)(H1 L)/			*!	**		*	

The Latin stress analyses (termed the *uneven-trochee analysis*, the *moraic-trochee analysis*, and the *T&S analysis*) introduced in this section are tested by computer simulations on three training sets of underlying-surface pairs and two training sets of overt forms, for 10 virtual CD learners and 10 virtual GLA learners. This adds up to a total of 6x5x20 = 600 simulated acquisition processes. Sections 4.3-4.5 describe the constraint sets, the training sets, and the acquisition processes.

4.3 The constraint sets

For the simulations, different constraints sets were set up to accommodate the uneven-trochee analysis and the moraic-trochee analysis, with the slightest possible changes. The basis for all constraint sets was the list of constraints as used in Tesar & Smolensky's (2000) learnability simulations, listed in (14). The uneven-trochee set differed from this set in the constraint on trochaicity: FTNONFIN was replaced with TROCHAIC. The moraic-trochee set differed from the original constraint set in that FTNONFIN was replaced by TROCHAIC and NONFINAL replaced by HDNONFIN. These three constraint sets were further modified by adding a constraint FTBIMORAIC. Table (68) summarizes the six constraint sets of section 4.2.2.3. In order that the reader can perform a simple though perhaps tedious check on the correctness of the evaluator, in the last column a count of the total number of constraint violations is included in the 15344 candidates in the 62 tableaux.

(68) Statistics on the six constraint sets

Constraint set	No. of Const.	Trochaicity constraint	Extrametrical constraint	Bimoraicity constraint	Violations
T&S	12	FTNONFIN	NONFINAL	(none)	370404
uneven trochee	12	TROCHAIC	NONFINAL	(none)	340028
moraic trochee	12	TROCHAIC	HDNONFIN	(none)	335932
T&S + FTBIMOR	13	FTNONFIN	NONFINAL	FTBIMOR	398062
uneven trochee + FTBIMOR	13	TROCHAIC	NONFINAL	FTBIMOR	367686
moraic trochee + FTBIMOR	13	TROCHAIC	HDNONFIN	FTBIMOR	363590

4.4 The training data

As mentioned before, every training set contains 28 different forms: no words with five or more syllables are fed to the listener during acquisition.

Each of the first three training sets consists of 28 pairs of given underlying forms together with the fully specified surface forms. The complete list is in table (69); where two or more analyses predict the same form, some ink has been saved. The 'uneven trochee' set is meant to replicate the uneven trochee analysis. The 'at most bimoraic' and 'at least bimoraic' sets are meant to give an analysis that is even more bimoraic than the moraic-trochee analysis in 4.2.5 (which still includes (HL) feet if no special constraints are added). At this point no explicit analyses with secondary stress are included.

The three analyses in table (69) all share the same overt forms, which can be seen in table (70). These are used in the simulations with overt forms.

An important question when dealing with stress systems is whether the language employs secondary stress. For Latin, this question is not trivial (see §4.6.3 and §4.9). *If* Latin had secondary stress, there are many different possibilities to place it. It could be quantity-sensitive or quantity-insensitive, stress clash could be permitted or not, and so on. Thereby many different sets of overt forms with secondary stress are thinkable. At this point, one secondary-stressed overt data set is included, which is shown in table (70). This set has weight-sensitive secondary stress before the main stress: every H is footed,[30] as is every remaining LL; the ambiguity that this will lead to in cases like |L L L H L| will have to be solved by the learner on the basis of her own generalization from the shorter forms in the training set to forms longer than four syllables.

---

[30] Unlike Jacobs (2003), who assumes *CLASH to prevent two heavy syllables next to each other from being both stressed, every heavy syllable to the left of the main stressed syllable is stressed here.

(69) Three training sets with fully structured surface forms

<table>
<tr><th rowspan="2">underlying forms</th><th colspan="3">surface forms</th></tr>
<tr><th>uneven trochee</th><th>at most bimoraic</th><th>at least bimoraic</th></tr>
<tr><td>|L L|</td><td>/(L1) L/</td><td>/(L1) L/</td><td>/(L1 L)/</td></tr>
<tr><td>|L H|</td><td>/(L1) H/</td><td>/(L1) H/</td><td>/(L1 H)/</td></tr>
<tr><td>|H L|</td><td colspan="3">/(H1) L/</td></tr>
<tr><td>|H H|</td><td colspan="3">/(H1) H/</td></tr>
<tr><td>|L L L|</td><td colspan="3">/(L1 L) L/</td></tr>
<tr><td>|L L H|</td><td colspan="3">/(L1 L) H/</td></tr>
<tr><td>|L H L|</td><td colspan="3">/L (H1) L/</td></tr>
<tr><td>|L H H|</td><td colspan="3">/L (H1) H/</td></tr>
<tr><td>|H L L|</td><td>/(H1 L) L/</td><td colspan="2">/(H1) L L/</td></tr>
<tr><td>|H L H|</td><td>/(H1 L) H/</td><td colspan="2">/(H1) L H/</td></tr>
<tr><td>|H H L|</td><td colspan="3">/H (H1) L/</td></tr>
<tr><td>|H H H|</td><td colspan="3">/H (H1) H/</td></tr>
<tr><td>|L L L L|</td><td colspan="3">/L (L1 L) L/</td></tr>
<tr><td>|L L L H|</td><td colspan="3">/L (L1 L) H/</td></tr>
<tr><td>|L L H L|</td><td colspan="3">/L L (H1) L/</td></tr>
<tr><td>|L L H H|</td><td colspan="3">/L L (H1) H/</td></tr>
<tr><td>|L H L L|</td><td>/L (H1 L) L/</td><td colspan="2">/L (H1) L L/</td></tr>
<tr><td>|L H L H|</td><td>/L (H1 L) H/</td><td colspan="2">/L (H1) L H/</td></tr>
<tr><td>|L H H L|</td><td colspan="3">/L H (H1) L/</td></tr>
<tr><td>|L H H H|</td><td colspan="3">/L H (H1) H/</td></tr>
<tr><td>|H L L L|</td><td colspan="3">/H (L1 L) L/</td></tr>
<tr><td>|H L L H|</td><td colspan="3">/H (L1 L) H/</td></tr>
<tr><td>|H L H L|</td><td colspan="3">/H L (H1) L/</td></tr>
<tr><td>|H L H H|</td><td colspan="3">/H L (H1) L/</td></tr>
<tr><td>|H H L L|</td><td>/H (H1 L) L/</td><td colspan="2">/H (H1) L L/</td></tr>
<tr><td>|H H L H|</td><td>/H (H1 L) H/</td><td colspan="2">/H (H1) L H/</td></tr>
<tr><td>|H H H L|</td><td colspan="3">/H H (H1) L/</td></tr>
<tr><td>|H H H H|</td><td colspan="3">/H H (H1) H/</td></tr>
</table>

(70) Two training sets with overt forms

overt forms	
main stress only	secondary stress
[L1 L]	
[L1 H]	
[H1 L]	
[H1 H]	
[L1 L L]	
[L1 L H]	
[L H1 L]	
[L H1 H]	
[H1 L L]	
[H1 L H]	
[H H1 L]	[H2 H1 L]
[H H1 H]	[H2 H1 H]
[L L1 L L]	
[L L1 L H]	
[L L H1 L]	[L2 L H1 L]
[L L H1 H]	[L2 L H1 H]
[L H1 L L]	
[L H1 L H]	
[L H H1 L]	[L H2 H1 L]
[L H H1 H]	[L H2 H1 H]
[H L1 L L]	[H2 L1 L L]
[H L1 L H]	[H2 L1 L H]
[H L H1 L]	[H2 L H1 L]
[H L H1 H]	[H2 L H1 H]
[H H1 L L]	[H2 H1 L L]
[H H1 L H]	[H2 H1 L H]
[H H H1 L]	[H2 H2 H1 L]
[H H H1 H]	[H2 H2 H1 H]

Five training sets have been established by now, although one could think of several more, both for the underlying-surface pairs and for the overt forms. All thinkable training sets, however, must be identical with respect to where the main stress falls: on the penultimate syllable if this is heavy, and on the antepenultimate otherwise.

4.5 The acquisition processes

The virtual learners of Latin stress learned from tableaux as presented in chapter 3. When learning from overt data as in table (70), learning took place as described in sections 3.2 to 3.6. Applied to Latin, perception looks as in tableau (71).[31] Imagine that at some point in learning, the learner encounters a [L H1 L] sequence; an overt input that consists of words made up of light and heavy syllables, of which one is marked for stress. In tableau (71), the grammar chooses for the left aligned iambic candidate /(L H1) L/ as the perceived surface structure (overt forms are once more excluded).

(71) Perception in Latin

overt: [L H1 L]	AFL	AFR	Trochaic	Iambic
a. /L (H1) L/	*!			
👂 b. /(L H1) L/		*	*	
c. /L (H1 L)/	*!			*

In virtual production, shown in (72), a left-aligned trochee is chosen by the constraint ranking. The winning candidate in production is indicated by '👄', and the winning candidate in perception is indicated by ' 👂 '.

(72) Production and error-detection in Latin

underlying: |L H L|	AFL	AFR	Trochaic	Iambic
a. /(L1) H L/		**!		
👄 b. /(L1 H) L/		*		*
c. /L (H1) L/	*!	*		
d. /L (H1 L)/	*!			*
👂 e. /(L H1) L/		*	*!	
f. /L H (L1)/	*!*			
g. /L (H L1)/	*!		*	

Thereby an error was detected and the constraints will be reranked depending on the reranking strategy.

---

[31] This tableau does not show the starting point of learning; it rather applies the perception process demonstrated in section 3.2 to Latin.

When learning from surface structures as given in table (69), learning proceeded as in section 3.7. In this case, only the computation of surface forms in production had to be learned.

For the two reranking strategies the implementations in the Praat programme (Boersma & Weenink 1992-2006) are used. The evaluation model for CD 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 & Smolensky (2000), the algorithm was allowed to chew five times on every piece of language data, with backtracking if the quintuple chews did not succeed in making the (alleged) correct adult form optimal in the learner's grammar. A slight difference with Tesar & Smolensky's evaluation model was that when two forms were equally harmonic, a winner was chosen randomly from among them, whereas Tesar & Smolensky somewhat less realistically chose the form that occurred first in the tableau (p.c. between Bruce Tesar and Paul Boersma).

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 at each evaluation. 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 surface form and an underlying form, then to produce the learner's surface form from the interpreted underlying form.

For each of the 600 virtual learners, the 12 constraints (or 13 constraints, respectively) were initially ranked at a height of 100. After this, language data were drawn randomly with equal probability from the 28 underlying-surface pairs or from the 28 overt forms. All learners therefore heard the forms in different orders and with (very slightly) different frequencies. When a pair or a form caused a mismatch between the learner's own produced surface form and (her guess of) the correct adult form, the CD 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 the rankings of some constraints by 0.1 and lower the rankings of some others 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.

A CD learner was allowed to listen to maximally 1 000 pairs of underlying-surface forms or 1 000 overt forms. After every 100 pairs or forms, however, it was checked whether the learner had already arrived at a grammar in which all 28 pairs or forms were *singly grammatical*. An underlying-surface pair is singly grammatical if the surface form is the only optimal candidate for the underlying form, i.e. if it is optimal in its tableau and no other candidate in the same tableau is equally harmonic. An overt form is singly grammatical if for all the tableaux in which it occurs (in the current case this is always a single tableau), this overt form is shared by all optimal candidates. For instance, the overt form [H H1 L L] can be considered singly grammatical if the optimal candidate in the tableau for |H H L L| is /(H H1) L L/, /H (H1) L L/, or /H (H1 L) L/. If all 28 pairs or forms are singly grammatical, it is certain that the learner will not be capable of any more learning with these forms because encountering new forms will not lead to error detection anymore. When this occurred, learning was considered successful and the simulation was stopped (i.e., no more forms were fed to the learner). A CD learner usually either successfully acquired the language within the first or second round of 100 pairs or forms, or she did not acquire the language even after 1 000 pairs or forms; in the latter case it is certain that the learner will never succeed, as is exemplified in the discussions on tables (78) and (85).

GLA learners (who take much smaller steps than CD learners) were allowed to listen to maximally 40 000 underlying-surface pairs or 40 000 overt forms. After every 1 000 pairs or forms, it was checked whether the learner had arrived at a grammar in which all of the pairs or forms were singly grammatical. If so, the simulation was stopped. When deciding whether a pair or form was singly grammatical, the evaluation noise was set to zero and the optimal candidates in the 28 relevant tableaux were computed, then proceeding as above. Although the learner would still be likely in this situation to make several mistakes if the evaluation noise had the usual value of 2.0, it was decided that learning had succeeded, because it was certain that the learner's constraints were already ranked in the correct order and that future learning would reduce the error rate but not change the crucial rankings.

4.6 Results

Table (73) shows the results for the 600 learners. In each cell, the result is indicated as x/y, where *x* is the number of CD learners that succeeded and *y* is the number of GLA learners that succeeded. When none of the 10 learners succeeded, this is indicated by "–"; when all 10 learners succeeded, this is indicated by "√".

(73) Simulation results for 600 learners, in the form "CD/GLA"

	Learning from pairs of underlying and surface forms			Learning from overt forms	
Constraint set	uneven trochees	at most bimoraic	at least bimoraic	main stress only	secondary stress
T&S	–/–	–/–	–/–	–/–	–/√
uneven trochee	√/√	–/–	–/–	–/√	–/√
moraic trochee	√/√	–/–	–/–	1/–	–/–
T&S + FtBim	√/√	√/√	–/–	–/–	–/√
uneven trochee + FtBim	√/√	√/√	–/–	–/7	–/√
moraic trochee + FtBim	√/√	√/√	√/√	9/–	–/–

4.6.1 Informed learning of primary stress in Latin

Table (73) shows that CD and GLA were equally successful in learning from pairs of underlying and surface forms: every cell in the first three columns either contains "√/√" (all 10 CD learners and all 10 GLA learners succeeded) or "–/–" (all 20 learners failed). This is not surprising. CD is a generally applicable OT learning algorithm that when supplied with fully specified underlying-surface pairs, is guaranteed to find a ranking that can generate those forms, if there *is* such a ranking. Thus, from the first "–" in every cell with "–/–" it appears that there is no ranking at all that can generate the 28 underlying-surface pairs at hand with the constraint set at hand. This necessarily means that the GLA will not be able to find an appropriate ranking either (as confirmed by the second "–" in all these cells). From the first "√" in every cell with "√/√" it appears that there *is* a ranking, and the second "√" in these cells tells us that the GLA has also been able to find it. Since there are no cells with "√/–" in the first three columns, we can conclude for the case of Latin stress that in all the cases in which CD works, the GLA works as well.

As expected from the ranking we found in §4.2.3, the uneven trochee analysis was learnable with the uneven trochee constraint set. Two of the CD learners arrived at the ranking in (74).

(74) Idealized results for uneven-trochee CD learning of the uneven-trochee analysis

Constraints	Ranking values
NONFINAL, TROCHAIC	100
AFR, FTBIN, MAIN-R, WFR	99
IAMBIC, WSP	98
AFL, MAIN-L, PARSE, WFL	97

This ranking is exactly what can be predicted from the crucial rankings in figure (56). CD is an algorithm that is claimed to rank every constraint maximally high. When comparing the ranking of the learner in (74)with the ranking in (56), we can see that NONFINAL and TROCHAIC are undominated, so their ranking stays at the original 100 in the simulations. The constraint AFR and FTBIN are outranked only by undominated constraints, so they end up at a height of 99. Each of the constraints WSP and IAMBIC is dominated by a constraint from the second level, so they end up at 98. The deepest constraint in (56) is PARSE; it must end up at 97. The remaining five constraints end up as high as they can without altering any of the optimal candidates: MAIN-R can end up at 99 because it is assigned the same number of violations in all winning candidates as AFR. WFR has to go below NONFINAL, with which it has complementary violations.

But the constraints are not always ranked maximally high. One learner ends up with a ranking similar to (74), but with WFR ranked at 98; and another learner ends up with both WFR and MAIN-R ranked at 98. While this makes no difference in the output forms, the maximally-high-ranking claim of CD is violated here, probably because of the existence of solutions with crucial ties, for which we will now see some more dramatic examples.

Four CD learners ended up with what is probably the minimum number of strata: they collapsed the AFR - FTBIN - MAIN-R - WFR stratum with the IAMBIC - WSP stratum (at 99), and had the four bottom constraints (AFL - MAIN-L - PARSE - WFL) end up at 98. At first sight this violates the crucial rankings established in figure (56). Crucial ties save the analysis, as shown in tableau (75): the three violations of AFR in /(H2)(H1) L/ outnumber the single violations of WSP and AFR in /H (H1) L/.

(75) Crucial ties invalidate crucial rankings

|H H L|	TROCHAIC	NONFINAL	FTBIN	AFR	WSP	IAMBIC	PARSE
/(H1) H L/				**	*(!)		**
/(H1 H) L/				*	*	*(!)	*
☞ /H (H1) L/				*	*		**
/H (H1 L)/		*!			*		*
/(H H1) L/	*!			*	*		*
/(H1)(H2) L/				***!			*
/(H2)(H1) L/				***!			*

What's more, even FTBIN can be ranked equally high as IAMBIC, at least if crucial ties are allowed, as tableau (76) shows.

(76) A crucial tie between FTBIN and IAMBIC at work

|L L L|	NONFIN	TROCHAIC	AFR	FTBIN	IAMBIC	WSP	PARSE
/(L1) L L/			**(!)	*(!)			**
☞ /(L1 L) L/			*		*		*
/L (L1) L/			*	*			**!
/L (L1 L)/	*!						*
/(L L1) L/		*!	*				*

The concept of the crucial tie, inherited from the early days of OT, may not be worth pursuing. After all, how can two violations of the doubly gradient constraint AFR (which counts feet as well as distance) are worse than a single violation of the singly gradient constraint WSP (which counts syllables)? Under a more realistic interpretation of tied constraints, namely that by Anttila (1997), an equal ranking of AFR and WSP in tableau (55) means that both /H (H1) L/ (/au (díː) re/) and /(H2)(H1) L/ (/(aù)(díː) re/) would win in 50% of the cases, and an equal ranking of IAMBIC and FTBIN in tableau (76) would mean that both /(L1 L) L/ (/(já.ke) re/) and /L (L1) L/ (the overtly incorrect /ja (ké) re/) would win in 50% of the cases. This optionality could be introduced in the simulations by taking a tiny evaluation noise, say 0.000001, for the CD simulations performed with Praat. All 10 CD learners would end up in the ranking in (74).

It remains to be said what the remaining two CD learners did. Like the four crucial tie learners just discussed, they had three strata, but one of them had WSP in the bottom stratum (at 98), and the other had MAIN-R, WFR, and WSP in that stratum. Apparently, both of these had managed to learn the language, but again by relying on the crucial tie principle.

The ranking in tableau (74) and those discussed after (74) produce all 28 forms in the first column of table (69). The rankings also correctly generalize to the 34 longer forms that the learner has never heard: they predict, for instance, /H L (H1 L) H/ (/in.di (gén.ti) a/ 'want') and /L L L L (L1 L) L/ (a form that we have not encountered so far).

The next step is to see how the GLA learners have performed. They cannot be bothered by crucial ties, because with a non-zero evaluation noise the probability that two constraints are ranked equally high at evaluation time is practically zero. If two constraints are ranked at nearly the same height, the distribution of outputs of the grammar will be very similar to the Anttila interpretation of a pair of tied constraints. All GLA learners end up with the ranking in table (77), although the precise ranking values differ among the learners, and half of the learners have a reversed ranking for the two bottom-ranked constraints WFL and PARSE.

(77) A typical uneven-trochee GLA learner of the uneven-trochee analysis

Constraints	Ranking values
NONFINAL	110.027
TROCHAIC	105.725
AFR	105.057
FTBIN	104.664
IAMBIC	100.539
WSP	99.984
MAIN-R	99.826
AFL	97.967
MAIN-L	94.105
WFR	89.973
WFL	89.702
PARSE	88.618

The crucial rankings of figure (56) can be found as large ranking distances in table (77). WSP and IAMBIC have stayed where they began, around 100. The three constraints that crucially outrank these two in (56) have been pushed up to about 105. The single constraint that crucially outranks two of the constraints around 105 has been pushed up to a height of 110. The constraint crucially dominated by WSP and IAMBIC (i.e. PARSE) has fallen a double distance, to the region near 90. This deep falling of weak constraints is typical of what the GLA does in general; in this case it is not a result of a domination by MAIN-L or so.

So far, CD and GLA learners have performed equally well, although the CD learners have practiced fancy behaviour by ingeniously inventing analyses with crucial ties whereas the GLA learners have boringly but reliably mimicked the expected ranking of figure (56).

For the T&S constraint set, table (73) shows that there exists no ranking that produces the forms associated with the uneven trochee analysis. According to table (68), this can only be due to a difference between TROCHAIC and FTNONFIN. Indeed we saw in §4.2.4 that the combination of FTNONFIN and IAMBIC is not capable of performing the conspiracy that led the combination of TROCHAIC and IAMBIC to force a winner with a monosyllabic foot in tableau (55). If TROCHAIC is replaced with FTNONFIN in tableau (55) or figure (56), we see that /(H1 H) L/ becomes the winner, because /H (H1) L/ now violates FTNONFIN, which is higher ranked than IAMBIC, which remains the only constraint in (55) and (56) that favours /H (H1) L/ over /(H1 H) L/. But it is still instructive to see how CD and GLA learners perform with the T&S set. Table (78) shows where one CD learner was after the simulations had to stop, i.e. after 1 000 language data.

(78) A T&S CD learner of the uneven-trochee analysis

Constraints	Ranking values
NONFINAL	100
AFR, FTBIN, MAIN-R, WFR, WSP	99
WFL	-102
IAMBIC	-109
AFL, FTNONFIN , MAIN-L, PARSE	-110

This learner has not ended up in a stable grammar. If she encounters more language data, the six constraints at the bottom will continue tumbling down the hierarchy. All ten CD learners have these six constraints ranked in different orders, but all in the vicinity of -110, which will be around -320 after 2 000 language data. At this snapshot in time, the learner of table (78) has iambic forms like /(L L1) L/ (/(ja.ké) re/). When being told that the form should be /(L1 L) L/ (/(já.ke) re/), she will demote IAMBIC to -111. Unfortunately, this will in turn lead her to generate a trochaic /(H1 H) L/ (/(aú.diː) re/). When being told that this should have been /H (H1) L/ (/(au.díː) re/), she will demote WFL, AFL, FTNONFIN, MAIN-L, and PARSE to -112, because all of these constraints prefer /(H1 H) L/ to /H (H1) L/ (and are higher ranked than IAMBIC, the highest constraint that prefers /H (H1) L/). This will go on forever. To measure how well these learners behave as speakers of Latin, their *error rates* were computed in the following way. A 1 000 underlying-surface pairs were randomly drawn, chosen with equal probability from the 28 underlying-surface pairs that have been used in training (therefore, each pair was chosen approximately 36 times on average), and then the learner's surface form for the given underlying form was computed. Each learner's form was then compared with the given adult surface form, and the learner was considered correct if the surface forms were identical. If a learner had e.g. 600 forms correct, her error rate was 40%. Eight of the CD learners turned out to have error rates of approximately 65%, the remaining two had error rates of about 44%.

The GLA learners also fail with the T&S set, but in a different way from the CD learners. The GLA learners all end up in a stable grammar. Table (79) shows the result for one learner.

(79) A T&S/GLA learner, fed with the uneven-trochee analysis

Constraints	Ranking values
NONFINAL	156.752
AFR	150.041
WSP	144.622
IAMBIC	139.944
FTNONFIN	139.618
FTBIN	95.926
MAIN-R	70.001
WFR	43.248
WFL	-401.795
AFL	-1078.115
MAIN-L	-1095.503
PARSE	-1204.395

IAMBIC and FTNONFIN are ranked very close together. Half of the 10 GLA learners have the same ranking as in (79), half have IAMBIC and FTNONFIN reversed. The distance between these two constraints is always small, so that if the learner has evaluation noise during her productions, she will have the ranking IAMBIC >> FTNONFIN approximately half of the time, and FTNONFIN >> IAMBIC the other half of the time. The error rates computed with an evaluation noise of 1.0 (smaller than the noise during training) are between 48% and 58%; the typical errors are that the learners show variation between /(L1 L) L/ and /(L L1) L/ and between /(H1 H) L/ and /(H H1) L/.

Table (73) shows that none of the three constraint sets without FTBIMORAIC is capable of learning a truly bimoraic analysis, like the 'at most bimoraic' and 'at least bimoraic' analyses of table (69). This is not so surprising. We have already seen in §4.2.5 that without constraints that favour strictly bimoraic feet, like *(HL) or FTBIMORAIC, one cannot expect the grammars to be able to learn bimoraic data. Still, the simulations with the moraic-trochee constraint set were successful in learning the uneven trochee analysis. Table (80) shows the resulting grammar for a CD learner.

(80) The generic result for moraic-trochee CD learning of the uneven-trochee analysis

Constraints	Ranking values
HDNONFIN, TROCHAIC	100
AFR, FTBIN, MAIN-R	99
IAMBIC	98
AFL, MAIN-L, WFL	97
PARSE, WFR, WSP	96

This grammar was reached by six of the ten CD learners. This grammar works for both versions of non-stochastic OT: that with crucial ties and that with variationist (*Anttila*) ties. The remaining CD learners had a grammar with a depth of 4, in which IAMBIC was one stratum higher, at 99 (now that the third stratum had been vacated, the ranking of the six constraints dominated by IAMBIC was of course 1 higher as well); this grammar relied on a crucial tie between IAMBIC and FTBIN, as above in the case of the T&S constraint set. Table (81) shows the result for a GLA learner. In (81), no clear layering has yet been established. This could be due to the fact that learning was designed to stop when the error rate was 0% if the evaluation noise was set to zero. With an evaluation of 2.0, i.e. the same as during learning, the error rate for the learner in (81) is still 30%. This means that in the grammar state of (81), the learner will detect mismatches for 30% of the incoming data and therefore take another learning step in 30% of the cases. These learning steps will continue to increase the separation between the constraints in (81).

(81) A moraic-trochee GLA learner of the uneven-trochee analysis

Constraints	Ranking values
HDNONFIN	110.744
TROCHAIC	106.141
FTBIN	105.867
AFR	105.058
IAMBIC	103.180
MAIN-R	102.464
AFL	100.627
WSP	97.805
MAIN -L	88.622
WFL	85.479
PARSE	78.739
WFR	78.037

In order to see whether the crucial rankings had been established in (81), the error rate was computed for an evaluation noise of 1.0. It was 7%; this number tells us something about how the learner will behave after making twice as many learning steps as she has made before reaching the state in (81).

In all cases the resulting ranking for the moraic-trochee constraint set is rather different from Prince & Smolensky's (1993) proposal, which was discussed in §4.2.5. As predicted, a high ranking of AFR rules out secondary stress before the main stress, and AFL >> WSP rules out secondary stress *after* the main stress.

It appears from table (73) that including the FTBIMORAIC constraint improves learnability from underlying-surface pairs. It is not surprising that if the uneven-trochee set and the moraic-trochee set succeeded in learning the uneven trochee analysis, this analysis is still learnable if a constraint is added to these constraint sets. But the addition of FTBIMORAIC seems to be just enough for the T&S set to achieve successful acquisition. The cause of this is that FTBIMORAIC is capable of ruling out (H H) feet but not (L L) feet: FTNONFIN can now outrank IAMBIC in order to produce /(L1 L) L/ rather than /(L L1) L/, without fear of producing /(H1 H) L/, because this form is ruled out by FTBIMORAIC. Otherwise, uneven trochees remain, as in /(H1 L) L/.

In general, the uneven trochee analysis seems to require fewer constraints (twelve) than the bimoraic analyses. The uneven-trochee and the moraic-trochee constraint set seem to be more successful than the T&S constraint set. But the differences between the constraint sets are small, especially regarding the success of the moraic-trochee+FTBIMORAIC set with the 'at least bimoraic' analysis. The next section shows whether there are any differences between the constraint sets when learning from overt forms only. The current section has at least shown that there were *some* combinations of constraint sets and analyses that were capable of learning the Latin stress system, so that we can now turn with confidence to the more realistic simulations, those for learning from overt forms, where hidden structures like feet are not explicitly provided to the learner but where she will have to construct them by herself.

4.6.2 Learning hidden structure and primary stress in Latin

Now we are going to have a look at the results of the simulations with more realistic primary language data. Table (73) shows that the T&S constraint set is not capable of learning a ranking for primary-stress-only overt data. This is not surprising, since the three primary-stress-only analyses (i.e. sets of given underlying-surface pairs) are not learnable with the T&S set either. Of course the learners could have invented a fourth analysis, perhaps one that includes /(L1 L)/ and /(L1) H/ or so, but they did not, so it is possible that there exists no analysis at all for primary-stress-only Latin with the T&S set.

The simulations with the uneven-trochee constraint set are more interesting: CD fails with this constraint set, the GLA succeeds. The first question now is: what analysis did the GLA learners come up with? The answer is that all learners came up with the same analysis, namely the uneven trochee analysis, i.e. for each of the 28 underlying forms in (69) they would produce the corresponding surface form in the 'uneven trochee' column (these surface forms were computed by running the 28 underlying forms through the learner's final grammar with an evaluation noise of zero). These learners ended up with the ranking in (82), sometimes with a different permutation of the very closely ranked constraints FTBIN, WSP, and TROCHAIC, or of IAMBIC, AFL, and MAIN-R.

(82) A typical result for uneven-trochee GLA learning from overt forms: creation of the uneven-trochee analysis

Constraints	Ranking values
NONFINAL	114.290
AFR	108.639
FTBIN	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

The ranking looks very different from that in (77). Still, (82) satisfies all of the crucial rankings. It will come to no surprise that these learners also correctly generalize the uneven-trochee analysis to forms of more than four syllables. The learners were rather slow in constructing the uneven trochee analysis by themselves. Whereas in the case of the underlying-surface pairs of table (69) all GLA learners had succeeded after the first 1 000 data, the learners of the overt forms needed 3 to 35 rounds of 1 000 data to arrive at an appropriate ranking. But they all succeeded.

Table (73) showed that for the three sets of 12 constraints, 29 out of 30 CD learners of primary-stressed overt forms fail. There is only one CD learner who happens to acquire an appropriate 12-constraint grammar; this learner uses the moraic-trochee set and invents an analysis that has not been considered so far, combining the two ‘at least bimoraic’ forms /(L1 L)/ and /(L1 H)/ of table (69) with the uneven trochees /...(H1 L) X/. This does not sound as a success for CD, since if only 10% of the children had been capable of learning Latin, this language would have perished much faster than it did.

When the constraint sets are enriched with FTBIMORAIC, the performance of CD improves. With the moraic-trochee+FTBIMORAIC set, nine learners managed to construct a functioning analysis. Seven of these came up with the uneven trochee analysis with the ‘at least bimoraic’ form

/(L1 X)/ mentioned before. The rankings of these learners slightly varied, as before. Table (83) shows an example.

(83) A typical result for moraic-trochee+FTBIMORAIC CD learning from overt forms: creation of the at-least-bimoraic uneven-trochee analysis, with empty strata

Constraints	Ranking values
FTBIN, HDNONFIN	100
AFR, FTBIMORAIC, MAIN-R	99
AFL, MAIN-L, WFL	98
PARSE, WFR	97
TROCHAIC, WSP	94
IAMBIC	93

A conspicuous property of seven of the resulting rankings is that they contained *empty strata*. In table (83), which is an average case, strata 95 and 96 are empty. Such empty strata can never occur when CD learns from fully specified underlying-surface pairs because constraints are demoted minimally, but they can when CD learns from overt forms only. Another conspicuous property of the seven rankings is that none of them is correct under the variational interpretation of tied constraints. To see whether there exists such a ranking at all, a simulation would have to be run in which the moraic-trochee+FTBIMORAIC constraint set learns an explicitly given at-least-bimoraic uneven trochee analysis. If so, and if we want to see whether CD can also learn it from overt forms, the simulations that led to table (83) will have to be rerun with a tiny evaluation noise.

Two of the CD learners constructed an at-least-bimoraic analysis. Both relied on crucial ties. Table (84) shows one of the rankings.

(84) Another result for moraic-trochee+FTBIMORAIC CD learning from overt forms: creation of the at-least-bimoraic analysis

Constraints	Ranking values
FTBIMORAIC, FTBIN, HDNONFIN	100
AFR, MAIN-R	99
AFL, MAIN-L, WFL	98
PARSE, TROCHAIC, WFR, WSP	97
IAMBIC	96

The ranking of FTBIMORAIC above AFR and MAIN-R causes the preference for /(H1) L L/ over /(H1 L) L/. We can compare this to the ranking in (83), where the crucial tie between these three constraints favours the uneven trochee /(H1 L) L/ over the bimoraic /(H1) L L/: FTBIMORAIC casts a single vote in favour of /(H1) L L/ whereas AFR and MAIN-R gang up with two votes in favour of /(H1 L) L/. The tenth moraic-trochee+FTBIMORAIC CD learner did not succeed in learning Latin. Her ranking after 1 000 data is given in (85).

(85) The single failure for moraic-trochee+FTBIMORAIC CD learning from overt forms

Constraints	Ranking values
FTBIN	100
HDNONFINAL	99
AFR, MAIN-R	98
AFL, FTBIMORAIC, MAIN-L, WFL	97
PARSE, WFR, WSP	96
IAMBIC	-104
TROCHAIC	-105

This learner has experienced IAMBIC and TROCHAIC tumbling down the hierarchy, alternatingly making the by now usual mistakes of /(H1 H) L/ and /(L L1) L/. To see whether this learner would learn the language later, she was taught 10 000 extra overt forms. This had no other effect than demoting IAMBIC and TROCHAIC down to -2238 and -2239. It appears that this learner, in contrast with the tenth GLA learner of the uneven-trochee set discussed above, has really got trapped in a sequence of grammars that she can never get out of (a 'non-globally-optimal limit cycle').[32] This may mean that the '9' in table (73) indicates that primary-stressed-only Latin is not learnable by the whole generation of learners if they entertain the moraic-trochee+FTBIMORAIC constraint set. Whether this situation means that this combination of constraint set, training set, and learning algorithm can be ruled out as a proposal for how Latin children acquired their language, or whether it is just a predictor of sound change, depends on the exact fraction

---

[32] This 'non-globally-optimal limit cycle' means that the learner ended up in something like a one-way dead-end street: no matter how much more language data this learner encounters, she will never be able to find a way back, and a way to a correct grammar.

of failures. The best guess at this point is 10%, but this number could be estimated more accurately after a future simulation of, say, 1 000 learners. The 10 GLA learners, by the way, were consistent in not learning with the moraic-trochee+FTBIMORAIC constraint set at all.

The remaining interesting figure for the main-stress-only forms in table (73) is the ‘7’ for the GLA learners with the uneven-trochee+FTBIMORAIC constraint set. Apparently, adding the FTBIMORAIC constraint to the set made the language *less* learnable from overt forms, compared to the uneven-trochee constraint set without FTBIMORAIC. The seven successful learners ended up with rankings that follow the stratification in (86). Interestingly, although FTBIMORAIC is ranked above FTBIN, this grammar renders uneven trochees such as /(H1 L) L/ and /(H1 L) H/, because FTBIMORAIC is dominated by AFR.

(86) A typical success for uneven-trochee+FTBIMORAIC GLA learning from overt forms

Constraints	Ranking values
NONFINAL	120.563
AFR	113.294
WSP	106.877
FTBIMORAIC	105.181
AFL	103.154
TROCHAIC	103.016
FTBIN	102.790
MAIN-R	102.439
IAMBIC	100.746
MAIN-L	97.615
PARSE	81.554
WFR	79.437
WFL	73.763

The remaining three learners were not lucky. Even after 50 000 data, they stuck with grammars like in (87). TROCHAIC and IAMBIC are always ranked very closely. Grammar (87) is of the type that we have seen several times before: since TROCHAIC and IAMBIC are very closely ranked, these learners end up producing one of the two mistakes /(H1 H) L/ or /(L L1) L/. The cause of the problem is that these learners have moved AFL too high up, and not managed to raise FTBIMORAIC above it. If FTBIMORAIC is ranked higher than AFL, it is capable of ruling out /(H1 H) L/, so that IAMBIC is freed from the task of ruling out/(H1 H) L/; this allows IAMBIC to fall below TROCHAIC, so that the learner also stops producing /(L L1) L/ errors. Apparently, adding a constraint does not necessarily improve learnability from overt forms.

(87) A typical failure for uneven-trochee+FTBIMORAIC GLA learning from overt forms

Constraints	Ranking values
NONFINAL	124.255
AFR	115.923
AFL	107.371
WSP	107.226
MAIN-R	103.226
FTBIN	101.215
MAIN-L	100.711
TROCHAIC	98.868
IAMBIC	98.633
FTBIMORAIC	95.829
PARSE	83.892
WFR	75.745
WFL	71.173

4.6.3 Learning hidden structure including secondary stress in Latin

Table (73) showed that CD is not capable of learning from overt forms with the secondary stresses listed in the last column of table (70), with any constraint set. By contrast, the GLA is successful with the T&S and uneven-trochee constraint sets, regardless of whether FTBIMORAIC is included or not. This looks better than the performance with the primary-stress-only forms, which could mean that additional information such as secondary stress does support learning.

Apart from the striking difference between the learning algorithms, the most conspicuous result in table (73) is that the T&S constraint set is successful for the first time. The 10 GLA learners created grammars very similar to the one in table (88).

(88) The result for T&S GLA for secondary-stressed overt forms

Constraints	Ranking values
NONFINAL	108.705
WSP	104.865
MAIN-R	102.437
FTBIN	101.430
WFL	100.773
FTNONFINAL	99.888
AFL	99.852
PARSE	99.273
IAMBIC	97.686
MAIN-L	95.353
AFR	91.682
WFR	91.295

For the 28 overt forms in table (70), this learner constructs an analysis with rather exhaustively parsed syllables and both iambic and uneven trochaic feet, with a preference for trochees:

(89) Mixed foot structure

/(L1) X/
/(H1) X/

/(L1 L) X/
/(L H1) X/
/(H1 L) X/
/(H2)(H1) X/

/L (L1 L) X/
/(L2 L)(H1) X/
/L (H1 L) X/
/(L H2)(H1) X/
/(H2)(L1 L) X/
/(H2 L)(H1) X/
/(H2)(H1 L) X/
/(H2)(H2)(H1) X/

The learner generalizes this exhaustivity to the 32 forms with five syllables, some of them containing both iambs and trochees:

(90) Forms with five syllables

/(L2 L)(L1 L) X/
/(L2 L)(L H1) X/
/(L2 L)(H1 L) X/
/(L2 L)(H2)(H1) X/
/(L H2)(L1 L) X/
/(L H2)(L H1) X/
/(L H2)(H1 L) X/
/(L H2)(H2)(H1) X/
/(H2 L)(L1 L) X/
/(H2 L)(H1 L) X/
/(H2 L)(H2)(H1) X/
/(H2)(H2)(L1 L) X/
/(H2)(H2 L)(H1) X/
/(H2)(H2)(H1 L) X/
/(H2)(H2)(H2)(H1) X/
/(H2 L)(L H1) X/

Importantly, the learner analyses /(H2 L)(L H1) X/ and not /(H2)(L2 L)(H1) X/. Note that in these cases, the learner has created her own patterns of overt forms, e.g. [L2 L L H1 X], which were not in the training set. This means that the learner will produce reasonably good pronunciations for five-syllable forms, even if she has never heard them before; for instance, if the learner is familiar with the nominative /ra (pí.di) taːs/ 'speed', she will come up with the form /(rà.pi) (di.táː) te/ for the ablative singular even if she has never heard that form. For the longest forms consisting of light syllables only, the analyses have a single left-aligned foot that contains a secondary stress: /(L2 L) L (L1 L) L/ and /(L2 L) L L (L1 L) L/. The exhaustivity noted above thus reduces (only in the case of light syllables) to a right-aligned main foot and a left-aligned secondary foot, which is caused by a high ranking of WFL, and a ranking of AFL above PARSE. Three other learners have exactly the same language, and three learners have a slightly different ranking that leads to exactly the same forms as above except that the form with seven syllables scans as /(L2 L)(L2 L)(L1 L) L/. This even more exhaustive parsing of syllables is caused by the ranking PARSE >> AFL. Actually, the speaker in table (88), with her close ranking of AFL and PARSE, can be expected to waver between the two forms with seven syllables. This variation (both between speakers and within speakers) seems to be similar to what real speakers of English, German or Dutch do with longer words (it could even depend on speaking rate, i.e., you could rank PARSE a bit lower when speaking fast). The remaining three GLA learners have /L (H2 L)(H1) X/ instead of /(L H2)(L H1) X/ (with variation in the seven-syllable form) caused by a ranking of FTNONFIN over WFL and PARSE. It is highly questionable whether the difference between those two forms is audible, so large-scale interspeaker variation for such hidden structures within the speech community should come to no surprise.

The results with the uneven-trochee constraint set are quite different. Table (91) shows the final ranking of one learner:

(91) One uneven-trochee GLA learner for secondary-stressed overt forms

Constraints	Ranking values
NONFINAL	108.761
FTBIN	104.053
WSP	103.321
MAIN-R	102.504
TROCHAIC	102.322
PARSE	99.965
AFL	99.818
WFL	97.991
IAMBIC	97.735
MAIN-L	95.671
AFR	94.134
WFR	91.239

This learner avoids iambic forms: she has /L (H1) X/, /L (H2)(H1) X/, /(L2 L) L (H1) X/, /L (H2 L)(H1) X/, and, this time, no other choice than the exhaustive form /(H2)(L2 L)(H1) X/. Since PARSE outranks AFL, the seven-syllable form is /(L2 L)(L2 L)(L1 L) L/. All nine other uneven-trochee GLA learners have AFL >> PARSE, and therefore the forms /(H2 L) L (H1) H/ and /(L2 L) L L (L1 L) L/. Adding FTBIMORAIC to the uneven-trochee constraint set can result in the ranking in (92). The learner in (92) has come up with an analysis that has uneven trochees for main stress (caused by MAIN-R >> FTBIMORAIC), but avoids uneven trochees for secondary stress (caused by the ranking FTBIMORAIC above PARSE and AFR): both phenomena can be seen in /(H2) L (H1 L) L/ . This learner also has /(H2)(L2 L)(H1) X/ and /(L2 L)(L2 L)(L1 L) L/.

Eight other GLA learners arrive in the same language as the learner in (92), except that three of them have a reverse ranking of PARSE and AFL, resulting in forms such as /(H2) L L (H1) X/ and /(L2 L) L L (L1 L) L/.

(92) uneven-trochee+FTBIM/GLA learning from secondary-stressed overt forms

Constraints	Ranking values
NONFINAL	108.926
WSP	103.858
MAIN-R	103.641
FTBIN	103.238
FTBIMORAIC	102.983
PARSE	101.836
AFL	100.866
TROCHAIC	100.785
WFL	97.937
MAIN-L	97.039
IAMBIC	96.771
AFR	94.656
WFR	91.074

The remaining learner, shown in (93), happened to come up with a real bimoraic analysis that avoids all uneven trochees, e.g. /(H2) L (H1) L L/.

(93) uneven-trochee+FTBIM/GLA learning with secondary-stressed overt forms: creation of the at-most-bimoraic analysis

Constraints	Ranking values
NONFINAL	108.711
WSP	103.607
FTBIMORAIC	102.917
MAIN-R	102.891
FTBIN	102.450
TROCHAIC	101.201
PARSE	100.534
AFL	100.430
WFL	97.738
IAMBIC	96.920
MAIN-L	96.906
AFR	94.262
WFR	91.289

The 10 GLA learners with the T&S constraint set and FTBIMORAIC behaved similarly: eight created the bimoraic analysis with the exhaustive forms /(H2)(L2 L)(H1) X/ and /(L2 L)(L2 L)(L1 L) L/, one a bimoraic analysis

with medially unfooted light syllables, i.e. /(H2) L L (H1) X/ and /(L2 L) L L (L1 L) L/, and one allowed uneven trochees in main feet only.

4.6.4 Conclusions

First it has to be said that learning Latin from overt data turns out to be possible. However, it also brings about some instances of the expected failures of CD and GLA (as mentioned in section 3.6), since the overt forms are often ambiguous with respect to their structural analysis. In fact, the only combination of constraint set and algorithm that was capable of learning from primary-stress-only overt forms for all 10 learners was the uneven-trochee set with the GLA. A combination that got close to this performance was the moraic-trochee+FTBIMORAIC set with CD, where nine out of ten learners detected a correct ranking. In order to reliably prove that the former combination is better than the latter, it would be necessary to show that it is nearly 100% correct, for instance by teaching 1 000 learners with the uneven-trochee/GLA combination and computing the percentage correct. This could take two weeks of computer time.

Since both CD and GLA make use of the same interpreting mechanism (Robust Interpretive Parsing), any crucial differences in performance between the two have to be attributed to the different kinds of reranking strategy (demotion-only vs. demotion-and-promotion, and one-shot learning vs. graduality).

Again, learning Latin from overt data turns out to be possible, at least with the GLA algorithm. Whether this means that CD should be ruled out as a candidate for describing Latin with secondary stress remains to be seen, since different secondary stress patterns than tested so far are thinkable, as shown in section 4.9.

The learners came up with ten different analyses for the overt data with secondary-stressed forms, with a total of 109 different surface forms for the 62 underlying forms. For the forms with at most four syllables, the overt forms associated with these ten analyses were (and need to be) identical. Differences between the analyses showed up only in a couple of overt forms with five syllables (namely [H2 L L H1 X] versus [H2 L2 L H1 X]) and in a form with seven syllables (namely [L2 L L L L1 L L] versus [L2 L L2 L L1 L L]), and this difference occurred with all five pairs of analyses that we have seen (i.e. ‘iambic & trochaic’, ‘less iambic’, ‘trochaic’, ‘bimoraic in secondary feet’, ‘bimoraic everywhere’), only depending on the relative ranking of AFL and PARSE, which were always closely ranked. Attested Latin forms with more than four syllables, if weight-sensitively secondary-stressed as here, would therefore give us no information about whether Latin learners used the T&S set or the uneven-trochee set, with or without FTBIMORAIC, and which of the five analysis types they created. Whether other patterns of secondary stress give us such information is investigated in §4.9.

4.7 More learners, different results?

The results in the previous section showed that the GLA learners equipped with the uneven-trochee constraint set (the one with TROCHAIC as the trochaicity constraint) could learn the Latin stress pattern from overt forms with only primary stress, whereas the CD learners could not learn with this constraint combination, but performed best when learning with the moraic-trochee constraint set including FTBIMORAIC. Since the learners were trained on the same word forms, but encountered each learning item in a different order, variation in the outcome is possible. To get a better understanding on whether the CD learners can *consistently* not learn the Latin stress pattern with the uneven-trochee constraint set as opposed to the GLA learners a 100 more learners (50 CD and 50 GLA learners) were run. Again the CD learners did not succeed: they ended up in a grammar that could not reproduce the correct stress pattern. And again the GLA learners did succeed: they came up with an analysis close to Jacobs’ (2000) analysis.

Once more, the learners were trained on all possible sequences of heavy and light syllables. Thus, four patterns of overt disyllables were fed to the virtual child. Likewise, there are eight trisyllabic overt forms, and in the same vein 16 overt forms with four syllables, all following the penultimate/antepenultimate Latin stress rule. The overt forms displayed in (94) are the same as in table (70):

(94) 28 primary-stress-only overt forms

Disyllables	Trisyllables	Quadrisyllables	
[L1 L]	[L1 L L]	[L L1 L L]	[H L1 L L]
[L1 H]	[L1 L L]	[L L1 L H]	[H L1 L L]
[H1 L]	[L H1 L]	[L L H1 L]	[H L H1 L]
[H1 H]	[L H1 H]	[L L H1 H]	[H L H1 H]
	[H1 L L]	[L H1 L L]	[H H1 L L]
	[H1 L H]	[L H1 L H]	[H H1 L H]
	[H H1 L]	[L H H1 L]	[H H H1 L]
	[H H1 H]	[L H H1 H]	[H H H1 H]

For each of the 100 virtual learners (50 CD learners and 50 GLA learners, once more created in the Praat programme; Boersma & Weenink 1992-2006), 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. The evaluation model for CD was OT with crucial ties. As in Tesar and Smolensky (2000), the algorithm was allowed 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, a winner was chosen randomly from among them. The evaluation model for the GLA was Stochastic OT with an evaluation noise of 2.0. The CD learners took learning steps of 1.0, and the GLA learners had decreasing learning steps, starting with 0.1; this plasticity of 0.1 was further randomized by a relative plasticity standard deviation of 0.1.

4.8 More learners, the same results

None of the CD learners succeeded in learning the stress pattern of Latin. The ranking after 10 000 data of one showcase CD learner is given in (95).

(95) A failing CD learner, after 10 000 data

Constraints	Ranking values
FTBIN, NONFINAL	100
AFR, MAIN-R, PARSE, WFR	99
AFL, MAIN-L, WFL	98
WSP	–2497
TROCHAIC	–2498
IAMBIC	–2499

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 (95). 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 CD learners, all 50 GLA learners succeeded (though five of them needed between 10 000 and 200 000 data to converge). Table (96) shows an example.

(96) A successful GLA learner, after 10 000 data

Constraints	Ranking values
NONFINAL	114.290
AFR	108.639
FTBIN	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

It is now discussed 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 FTBIN 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 it is stressed: /...(H1 L) X/; if the antepenultimate syllable is light, it is TROCHAIC that will make sure that it is stressed: /...(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 FTBIN, 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 FTBIN and WSP are concerned, and the last two satisfy TROCHAIC. The decision between these two will have to be brought by IAMBIC, shown in tableau (97).

(97) A constraint hierarchy that works for all Latin forms

|H H L|	NONFINAL	AFR	FTBIN	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/		**!*					*	*	*	*	*	

Figure (56) showed which of the rankings in (97) are crucial (ignoring the four less interesting and low-ranked 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; FTBIN could be ranked above AFR or below WSP, as long as it is ranked below NONFINAL and above IAMBIC; and so on.

The ranking in (56) is close to Jacobs' (2000) ranking: TROCHAIC >> NONFINAL >> FTBIN >> 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 the 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. Admittedly, it is counter-intuitive that in a language with exclusively trochaic feet, the constraint IAMBIC has to fix the analysis.

As far as longer forms are concerned: the learners have not been trained with forms of five syllables, but we can nevertheless run the 32 possible forms with five syllables through their respective tableaux and see what happens. All forms are handled correctly, for instance /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.

4.9 Secondary stress in Latin?

Another group of virtual learners was trained on several other kinds of overt Latin stress patterns: a case with main stress only, three cases with overtly available secondary stress, and a case in which the learners are free to invent their own secondary stress patterns. Several of these cases turn out to be learnable with the GLA, none with CD. The simulations in sections 4.61 and 4.6.2 dealt with primary stress only. Since it is controversial whether Latin had secondary stress, and if so, what it exactly looked like, several different data sets are designed with secondary stress. The virtual learners were then tested whether they were able to learn from these data sets, provided with the basic metrical constraint sets. For each simulation, a number of virtual CD and GLA learners was created, with constraint sets that contained either TROCHAIC or FTNONFIN. As before, the constraints started out with the same ranking heights (100.000). The training data were two to four syllables long and drawn randomly with equal probability from 28 possible overt forms. The CD learners were fed with 1 000 data pieces, while the GLA learners were fed with 10 000 up to 40 000 data pieces (because the GLA learners take smaller reranking steps). All of the simulations were carried out with the Praat programme (Boersma & Weenink 1992-200).

4.9.1 Very weight-sensitive secondary stress

One option would be to have weight-sensitive secondary stress by stressing every heavy syllable and every other light syllable before the main-stressed one. The resulting data set that the learners are fed with is given in table (98). Disyllables were also used in the simulations, but they are suppressed in the table since they do not differ from those in table (94).

(98) Very weight-sensitive secondary stress

Trisyllables	Quadrisyllables	
[L1 L L]	[L L1 L L]	[H2 L1 L L]
[L1 L L]	[L L1 L H]	[H2 L1 L L]
[L H1 L]	[L2 L H1 L]	[H2 L H1 L]
[L H1 H]	[L2 L H1 H]	[H2 L H1 H]
[H1 L L]	[L H1 L L]	[H2 H1 L L]
[H1 L H]	[L H1 L H]	[H2 H1 L H]
[H2 H1 L]	[L H2 H1 L]	[H2 H2 H1 L]
[H2 H1 H]	[L H2 H1 H]	[H2 H2 H1 H]

CD learners training with these 28 overt forms failed: with TROCHAIC, they produce initially stressed forms like */(L1 L) L L/, and with FTNONFINAL, they produce forms like */(H2)(L1 H) L/. Again, the GLA learners training with the same primary language data succeeded, independently from the constraint set they were using. GLA learners with TROCHAIC produced forms like /(L H1) L/, and /(L H2)(H1) H/ as in /(vo.lùp)(táː) teːs/. The GLA learners with FTNONFIN produced forms such as /L (H2)(H1) H/ as in /vo (lùp)(táː) teːs/, /L (H2)(L1 L) L/ as in /a (mìː)(kí.ti) a/, and /(H2)(H2)(H2)(H1) H/ as in /(dèː)(fìː)(nìː)(tíː) vus/.

4.9.2 Weight-insensitive secondary stress in Latin?

Another possibility to assign secondary stress is to build a weight-insensitive disyllabic foot at the left edge of the word, resulting in overt forms as in (99) (the 20 forms without secondary stress are suppressed).

(99) Left-aligned binary weight-insensitive secondary stress quadrisyllables

[L2 L H1 L] [L2 H H1 L] [H2 L H1 L] [H2 H H1 L]
[L2 L H1 H] [L2 H H1 H] [H2 L H1 H] [H2 H H1 H]

All learners that learn from this data set fail, simply because there *is* no ranking that can describe the data (at least no ranking with the constraints involved here).

The same happens with a training set that has left-aligned, binary weight-sensitive secondary stress, as in (100): there is no OT analysis with the constraint sets involved here that could describe this pattern, so again, all learners fail (data without secondary stress, like [L H1 L H] and [L H1 L], are again suppressed, although they would make the weight-sensitivity more explicit).

(100) Left-aligned binary weight-sensitive secondary stress

Trisyllables	Quadrisyllables			
[H2 H1 L]	[L2 L H1 L]	[L2 H H1 H]	[H2 L H1 L]	[H2 H1 L H]
[H2 H1 H]	[L2 L H1 H]	[H2 L1 L L]	[H2 L H1 H]	[H2 H H1 L]
	[L2 H H1 L]	[H2 L1 L L]	[H2 H1 L L]	[H2 H H1 H]

If learning from data that contain left-aligned weight-insensitive secondary stress that is not binary, as in (101), all learners fail, again because there is no constraint ranking that could produce the data.

(101) Left-aligned weight-insensitive secondary stress

Trisyllables	Quadrisyllables			
[L2 H1 L]	[L2 L1 L L]	[L2 H1 L L]	[H2 L1 L L]	[H2 H1 L L]
[L2 H1 H]	[L2 L1 L H]	[L2 H1 L H]	[H2 L1 L L]	[H2 H1 L H]
[H2 H1 L]	[L2 L H1 L]	[L2 H H1 L]	[H2 L H1 L]	[H2 H H1 L]
[H2 H1 H]	[L2 L H1 H]	[L2 H H1 H]	[H2 L H1 H]	[H2 H H1 H]

4.9.3 Freely assignable secondary stress

A further possibility to assign secondary stress is to let the learners invent it. This is done in the final simulation. The idea is that even if there appears only one audible stress in a word, the surface structure could be made up with several feet that are simply not articulated (see Halle & Vergnaud's 1987 *conflation*, and Hayes' 1995 reformulation of it). The consequence is that although children hear only primary stress, they could construct more than one foot in a word. Given this, GEN would then provide an additional candidate for a form like [H1 L]: /(H1) (L2)/.[33] Alternatively, secondary stress in the input could be ignored by a learner so that [H1 L2] could be perceived as /(H1) L/. Both strategies constitute a violation of faithfulness between overt form and surface form for secondary stress.[34] As usual, the choice between the candidates is determined by the ranking. The input to the simulations with freely assignable stress was therefore the same as in table (94): overt forms with primary-stress only. But this time the learners were allowed to invent secondary stress, i.e. their GEN contained not only forms with main stress, but also forms with main *and* secondary stress. Examples for the resulting constraint rankings are listed in tables (102) and (103):

(102) A CD learner

Constraints	Ranking values
FTBIN	100
NONFINAL	99
AFR, MAIN-R	98
AFL, MAIN-L, WFL	97
PARSE, WFR, WSP	96
IAMBIC	-104
TROCHAIC	-105

(103) A GLA learner

Constraints	Ranking values
NONFINAL	116.962
MAIN-R	110.198
WSP	106.139
PARSE	105.612
AFL	104.276
MAIN-L	100.623
TROCHAIC	99.743
WFL	99.185
IAMBIC	97.045
FTBIN	87.208
WFR	83.038
AFR	80.461

---

[33] This comes close to the forms in GEN of Prince & Smolensky (1993), where forms with feet to the right of the head foot were permitted: /(H1)(L L)/. In their analysis, these feet apparently did not carry secondary stress.

[34] This violation of faithfulness is not modelled with constraints in the simulations, but by adding candidates in GEN.

The CD learners were not able to produce main stress correctly. The GLA learners assigned main stress correctly, and furthermore created secondary stress in some forms: e.g. /(L2)(L1 L) X/ as for /(fà)(kí.li) ter/, and /(L2)(H1 L) X/ as for /(sù)(pér.bi) ter/, and furthermore forms like /(H2)(H1) X/, /(L2)(L H1) X/, /(L H2)(H1) X/, /(H2)(L1 L) X/, /(H2)(H1 L) X/, /(H2)(H2)(H1) X/, and /(H2)(L H1) X/.[35]

Their generalizations to longer forms were weird but correct: /(L2 L)(L H1) X/ as for /(rà.pi)(di.táː)tem/, and /(L2)(L2 L)(H1 L) X/ as for /(rà)(pì.di)(táː.ti)bus/. They are weird because the secondary stress assigned to the left of the main stress is influenced by what happens to the right of the main stressed syllable: if it is heavy and penultimate, an iambic foot is built as in /(L2 L)(L H1) X/; if it is heavy and antepenultimate, a trochaic foot is built, as in /(L2)(L2 L)(H1 L) X/. However, this does not look like a natural pattern.

4.9.4 Summary

Summary of successes. The successful simulations included three very different patterns with secondary stress, which could be learned by GLA learners only. A word like *voluptates* was analysed in the first simulation as /vo.lup (tá:) te:s/, in the second simulation as /vo (lùp)(tá:) te:s/ (with FTNONFIN), and as /(vo.lùp)(tá:) te:s/ (with TROCHAIC; the same in simulation 6). CD learners never converged upon a grammar that rendered the stress patterns in question. The immediate cause for this lies in the behaviour of the constraints for trochaicity and iambicity, since the CD strategy moves them to the bottom of the hierarchy early, while the GLA keeps them ranked in the middle.

Summary of failures. What is missing in the results of the simulations are analyses with strictly bimoraic feet such as /(H1) L <L>/, as proposed by Mester (1994) on the basis of segmental changes such as iambic and cretic shortening. The constraint sets involved here are not capable of producing this pattern, regardless of the input. Furthermore, although forms like

---

[35] This pattern (secondary stress on the initial syllable) comes close to the one proposed in the literature by e.g. Allen (1973:86, and reference therein), except that stresses in consecutive syllables is unlikely: Allen (1973) states that secondary stress only occurred if there was more than one syllable preceding the mainly stressed one.

[L2 H H1 H] were given in simulations 4.9.2 and allowed in the simulation in 4.9.3, no learner came up with the analysis of secondary stress actually proposed in the literature (Allen 1973, 1978), which contains wretched trochees such as /(L2 H)(H1) H/ as for */(vò.lup)(táː)teːs/*. It is likely that neither constraint set was suitable for such an analysis; the addition of a constraint like *CLASH could improve the results.

4.10 Discussion

In this rather long section on the learnability of Latin stress, the performance of two learning algorithms, six constraint sets, three analyses, and several kinds of overt forms have been investigated. This section reports some results on all these issues and indicates how several more constraint sets, analyses, and kinds of overt forms should be investigated in the future.

Analyses. The present investigation started by giving a couple of analyses that are capable of handling the positioning of Latin main stress correctly. The uneven-trochee analyses (§4.2.3) derived from Jacobs' (2000) analysis, which was augmented with IAMBIC in order to handle |H H L|. The moraic-trochee analysis (§4.2.5) derived from Prince & Smolensky's (1993) analysis, which was augmented with the ranking of AFR >> WSP and with the constraint IAMBIC, again in order to handle |H H L|. The uneven trochee analysis was better learnable than either of the two bimoraic analyses. However, from the simulations with overt forms a fourth analysis transpired that had not been considered before: an analysis with uneven trochees, as in Jacobs (2000), but with at-least-bimoraic feet, so that the light-initial disyllables become /(L1 L)/ and /(L1 H)/.[36] This fourth analysis may well lie at the basis of the process of *iambic shortening* in Pre-Classical Latin (underlying |L H|, e.g. the concatenation of the verb stem |am-| 'love' with the first singular ending |-oː|, becomes /(L1 L)/, e.g. /(á.mo)/ 'I love'), which many authors discuss (Kager 1993a, Prince & Smolensky 1993, Mester 1994, Jacobs 2000). Future research will have to take this analysis into account.

---

[36] These two forms actually occurred in Jacobs' original analysis for Classical Latin, but as shown in §4.2.3, these forms require (with the uneven-trochee constraint set) a ranking of FTBIN >> NONFINAL and of PARSE >> WSP, the latter of which fails to handle |L H L| correctly.

Learning algorithms. CD and the GLA performed strikingly differently in the simulations with overt forms. As summarized in table (73), the GLA succeeded with five combinations of constraint sets and kinds of overt forms. CD performed a bit better on the two constraint sets that contain the perhaps implausible constraint HDNONFIN, although each of these four groups of 10 CD learners had at least one learner who did not acquire Latin (see the rows 'moraic trochee' and 'moraic-trochee+FTBIMORAIC' in table (73), in combination with the columns 'main stress only' and 'secondary stress').

Constraint sets. The simulations seem to reveal that some of the proposed constraint sets are more adequate than others. For instance, TROCHAIC seems to be a better formalization for a trochaic foot pattern than FTNONFIN, which seems to be too restrictive. Also, NONFINAL seems to be a more effective formalization of extrametricality than HDNONFIN. But no constraint set can be ruled out completely yet. As usual in OT, the legitimacy of a constraint set ultimately has to be proven in combination with systems of other languages than the specific language under study.

Overt forms. Learning from forms with a certain type of secondary stress turned out to be easier than learning from forms with primary stress only. There is disagreement in the literature about whether Latin had secondary stresses, and therefore feet, before the primary-stressed foot, and if it had, where these secondary stresses were: they could have been weight-sensitive (Allen 1978) or not (Jacobs 1989). Allen (1978) states that if there was secondary stress, it was on the initial syllable, except if it was pretonic. Therefore it is unlikely that every heavy syllable was stressed in Latin as it was in the present series of forms, the last column in table (70). This series already led to ten different analyses, and other secondary stress patterns will lead to many more. A possible solution to this problem is to *let the learner decide*, as was done in chapter 4.9.3: the learner encountered only overt forms with primary-stressed syllables, but was allowed to invent a full foot structure with secondary-stressed syllables.

Frequency. In the present simulations, the learners were fed every type of underlying form equally often. The typical mistakes of the virtual learners were superheavy trochees in /(H1 H) X/ and iambs in /(L L1) X/. In most cases, the mistakes were caused by a close ranking of TROCHAIC (or FTNONFIN) and IAMBIC without a compensatory ranking somewhere else in the hierarchy. If these learners encountered more |H H X| than |L L X| forms, they would probably end up with a ranking of TROCHAIC slightly above IAMBIC. It is not unlikely that such a ranking would have helped the learner to avoid non-global optima, but this is left for later investigation.

Sound change. There were cases in which a small percentage of the learners did not succeed in acquiring the provided Latin stress pattern, while the great majority of the same type of learners did succeed. Such cases can be predictors of acquisition-induced sound change. It is possible, for instance, that not all constraints are innate, but that they are instead constructed by the learner. In that case, some learners may well entertain constraint sets that have here been shown to lead to unlearnability. The typical mistakes were trochaicity in /(H1 H) X/ and iambicity in /(L L1) X/. The disadvantage of taking a dead language to study acquisition can thus turn into an advantage, since we know a lot about what happened in the daughter languages. With some luck, later investigations may also be able to model the historical change from initial stress in Pre-Classical Latin to weight-sensitive right-aligned stress in Classical Latin.

More realistic models of metrical acquisition. It has simplifyingly been assumed that the learner's productions contained the same number of syllables as their underlying forms and the adult forms. However, it is likely that Latin children were similar to Dutch children (Fikkert 1994) and English children (Gnanadesikan 1995) in that they started out by truncating longer words, e.g. by turning trisyllabic words into disyllabic words consisting of a single foot, and that segmental structure interfered. Such a situation would have strong implications for all of the steps in the modelling of acquisition. For instance, this could mean that learners start out by acquiring everything there is to know about short words, before they go on to consider longer words. Dresher (1999), for instance, provides a non-OT metrical acquisition model that takes into account selective attention to specific structures.

Conspiracies between constraints. As seen in §4.2.2.3, TROCHAIC and IAMBIC conspire to minimize foot size. The alternative are FTNONFIN and IAMBIC. These two have fewer side effects since they have complementary violations on the foot level. This means that doing OT with the pair

FTNONFIN - IAMBIC is close to having a *parameter* “foot direction” in the grammar that is set to one of the values *nonfinal* or *iambic* (but not entirely, because these constraints still conspire to minimize the number of feet).

4.11 Conclusions

Latin stress turned out to be learnable with a limited set of constraints that many OT 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. The virtual learners were tested on two *on-line* learning algorithms, whose only memory of past events is indirectly and concisely stored in the ranking values of the constraints: Constraint Demotion (CD) and the Gradual Learning Algorithm (GLA). The GLA learners turned out to be successful in more constraint set/analysis combinations than the CD learners. It can be argued that the GLA is a more realistic ingredient of human language acquisition than CD, which has been shown earlier: 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 2004); and, last but not least, the GLA has been able to model language change induced by bidirectional language acquisition (Jäger 2003). Nevertheless, neither CD 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 the twelve constraints provided in (14) (Boersma 2003). 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 more evidence for an 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 OT learning algorithm provides the appropriate model for the acquisition of language.

Looking at the failure of the learners to come up with strictly bimoraic analyses, it appears that something has to be done about FTBIN. It should be split up into separate constraints that refer to moras on one hand and syllables on the other, as e.g. proposed by Hewitt (1994).

Also, if perception precedes lexical access, foot structure has to be assigned before word boundaries are. This order is problematic because some of the used constraints imply a dependence of foot assignment on word boundaries. Consider the overt form [áː.bra.ka.dáː.bra], to which the learner has to assign two feet and a word boundary. Under an analysis with uneven trochees the following problem emerges: if the word boundary is as in *aː.bra#ka.daː.bra*, the footing would have to be *(áː)bra#ka(dáː)bra*. If it is *aː.bra.ka#daː.bra*, footing would have to be *(áː.bra)ka#(dáː)bra*. This makes the strictly bimoraic analysis more likely, since this bimoraic analysis would predict identical footing in *(áː)bra#ka(dáː)bra* and *(áː)bra.ka#(dáː)bra*, so that feet can be assigned independently from (e.g. before) word boundaries.

A last point is that the learners were given too much information about syllable weight. Children have to e.g. learn the heaviness of CVC syllables. In some languages, CVC is light (e.g. final, monomoraic CVC-feet in Chuukese; Davis 1999, Muller 1999, Kennedy 2003), while in others it is heavy (e.g. in Latin).

In sum, it all smells like we need a more emergentist modelling of representations and constraints, meaning that much less should be given to the learner than was done in Tesar & Smolensky's (2000) and the present simulations.

The next chapter models weight-insensitive stress in Pintupi, where the learnability of coda moraicity is tested.