Learning to be insensitive to weight in Pintupi

Diana Apoussidou
University of Amsterdam

1 Introduction

It is usually taken for granted that normally developing children acquiring one and the same language end up with one and the same grammar (e.g. Chomsky & Halle 1968:251). The language-acquiring child is supposed to be capable of creating the adult grammar from the information provided in the speech stream, despite the fact that this information may be incomplete in terms of possible ambiguities and gaps in the data they are exposed to (known as the *poverty of the stimulus* problem; e.g. Chomsky 1986:7).

In the computer simulations of acquisition here it is shown that final grammars of virtual learners can differ even though they learned from the same data and have the same output than given in the training data. This is demonstrated by modelling word stress of Pintupi, a language spoken in Western Australia (Hansen & Hansen 1969). The grammatical framework is Optimality Theory (Prince & Smolensky 1993); the learning algorithms of the computer simulations are Error Driven Constraint Demotion (Tesar 1995) and the Gradual Learning Algorithm (Boersma 1997). All virtual learners are exposed to the same training data.

The paper is built up as follows: section 2 outlines the relationship between syllable structures and stress. Section 3 gives two possible Optimality Theoretic analyses of Pintupi word stress. Section 4 is on learnability. The adopted learning process is explained, along with the two learning algorithms. Section 5 gives the remaining ingredients to the computer simulations, the training data and the constraint sets. Sections 6 and 7 show the different grammars of the learners, followed by the discussion of the results and concluding remarks in section 8.

2 Syllable weight[1], stress and trochaic feet

Pintupi was chosen as the target language for the computer simulations because of its very regular and predictable stress pattern. The language has weight-insensitive stress, but displays some properties of weight-sensitivity. One of the questions asked here is the way virtual learners deal with this phenomenon. In the following the close relationship between syllable weight and stress is established. If a language has phonemic vowel length, it is said to have weight (or *quantity*) distinctions, and stress is usually sensitive to the weight of a syllable (2.1). If a language does not have phonemic vowel length, the language does not have weight distinctions, and stress therefore cannot be sensitive to weight. Stress is then assigned by other principles. However, a few languages do make weight distinctions, but assign stress independently from it (2.2). This type of stress system is of interest here.

[1] The terms 'quantity' and 'weight' are used interchangeably throughout the paper.

2.1 Quantity-sensitive stress

A language is called *quantity-sensitive* if it employs phonemic vowel length and assigns stress sensitively to the weight of a syllable. The weight of a syllable is determined by the number of *moras* it contains.[2] A syllable with a short vowel has one mora and is light, and a syllable with a long vowel has two moras and is heavy. Coda consonants can be moraic, i.e. can be linked to a mora, but this is not universal. In some languages a coda consonant is linked to a mora, making the syllable heavy (e.g. Yana; Sapir & Swadesh 1960), but in other cases the coda does not contribute to the weight of a syllable (e.g. Khalkha Mongolian; Walker 1997). In a quantity-sensitive language with trochaic rhythm, feet are ideally bimoraic: they should contain exactly two moras. Such feet are called moraic trochees (Hayes 1991). Moraic trochees can either consist of two light syllables containing one mora each, or of one heavy syllable containing two moras (1).

(1) Moraic trochees

(* .) or (*)
| | |
σ_μ σ_μ $\sigma_{\mu\mu}$

2.2 Quantity-insensitive stress

If a language does not employ any weight distinctions (i.e. it does not have distinctive vowel length) the language is trivially *quantity-insensitive*. Stress cannot be sensitive to weight under these circumstances. Trochees are in this case disyllabic: they contain two syllables, regardless of the syllable structure (2). Such feet are called syllabic trochees (Hayes 1991).

(2) Syllabic trochees

(* .)
| |
σ σ

Next to this kind of stress system are languages that do have weight distinctions, but assign stress independently of the weight of a syllable. Kager (1992) classifies these languages as *truly quantity-insensitive,* stating at the same time that these languages assign stress not completely independently from stress. If these languages have trochaic feet, the feet take the form of *generalized trochees* (Hayes 1991, 1995): feet are preferably disyllabic, else bimoraic (3).

(3) Generalized trochees

(* .) else (*)
| | |
σ σ $\sigma_{\mu\mu}$

These so-called truly quantity-insensitive languages reveal some sort of weight-sensitivity e.g. in secondary stress assignment or in a bimoraic word minimum. For instance in Estonian,

---

[2] Usually it is assumed that only the rhyme of a syllable can contain moraic elements. Reported exceptions to that is e.g. Pirahã (Everett & Everett 1984, Everett 1988), where the onset contributes to the weight of a syllable.

main stress as well as secondary stress is assigned by linking together syllabic trochees (Prince 1980, Kager 1992). However, in words with an odd number of syllables the last syllable is footed (and therefore stressed) if heavy.

Another language that is regarded by Kager (1992) as truly weight-insensitive in this sense is Pintupi (Western Australia; Hansen & Hansen 1969). In Pintupi, syllabic trochees are stringed together like in Estonian, with the difference that final syllables in odd-numbered words are never stressed. The learnability of the Pintupi stress system provides the basis for simulating the acquisition of metrical structure in section 4. The acquisition process is modelled in an Optimality Theoretic framework. Two learning algorithms, the Error Driven Constraint Demotion algorithm (EDCD; Tesar 1995) and the Gradual Learning Algorithm (GLA; Boersma 1997) are compared with each other with respect to whether the correct stress pattern of Pintupi is successfully acquired. The behaviour of the constraints is also examined.

3 Stress in Pintupi: two possible Optimality Theoretic accounts

Pintupi (Hansen & Hansen 1969) has a phonemic vowel length distinction, restricted to the initial syllable of a word. According to Hayes (1991, 1995) and Kager (1992) Pintupi has a bimoraic word minimum. This indicates that Pintupi is a mora-counting language where long vowels are linked to two moras and short vowels are linked to one mora. Stress is not sensitive to the weight of a syllable: primary stress is on the first syllable in a word, secondary stress is on every other following syllable except if that syllable is final in the word.[3] Syllables can have the shape CV, CVC or CVV, where ‘C’ stands for a consonant, ‘V’ for a vowel, and ‘VV’ for a long vowel. Traditionally, a stress pattern like the one in Pintupi is analysed with syllabic trochees iterating from left to right, starting at the left word edge (Hayes 1995). Final syllables in words with an odd number of syllables are unfooted since the language prohibits degenerate feet (feet that contain only one mora, i.e. one light syllable; Prince 1980, Hayes 1995:102).

Some examples are listed in (4) with the corresponding syllable and foot structure. The first column lists the overt forms, here understood as the directly observable auditory forms. These overt forms (displayed in square brackets) include primary stress (‘v́’), secondary stress (‘v̀’), syllable boundaries (‘.’), and vowel length (‘:’). In the second column the overt forms have been interpreted in terms of hidden metrical structure. These surface forms (displayed in slashes) contain foot structure (indicated by parentheses) as well as syllable boundaries (‘.’), the head syllable of the head foot (‘v́’), and the head syllable of a non-head foot (‘v̀’). Long vowels are interpreted as ‘vv’.

[3] Auditory cues for primary stress in Pintupi are loudness, often along with higher pitch and greater duration of syllables; the auditory cue for secondary stress is slightly increased loudness (Hansen & Hansen 1969).

(4)	Overt forms	Metrical surface structure	
a.	[tʲáː]	/ (cv́v) /	*mouth*
b.	[mú.ŋu]	/ (cv́.cv) /	*orphan*
c.	[múː.ŋu]	/ (cv́v.cv) /	*fly*
d.	[tʲán.pa]	/ (cv́c.cv) /	*evil spirit*
e.	[ká.pa.li]	/ (cv́.cv) cv /	*mother's mother*
f.	[ŋál.ku.nìn.pa]	/ (cv́c.cv) (cv̀c.cv) /	*eating*
g.	[pú.ɭiŋ.kà.la.tʲu]	/ (cv́.cvc) (cv̀.cv) cv /	*we (sat) on the hill*
h.	[tʲá.mu.lìm.pa.tʲùŋ.ku]	/ (cv́.cv) (cv̀c.cv) (cv̀c.cv) /	*our relation*

Analysing the Pintupi stress pattern in Optimality Theoretic terms[4] requires constraints on foot structure (PARSE, Prince & Smolensky 1993, McCarthy & Prince 1993; FOOTBINARITY, Prince & Smolensky 1993), on foot form (TROCHAIC, Prince & Smolensky 1993; IAMBIC, ref.?), and alignment constraints (ALL-FEET-LEFT/RIGHT, McCarthy & Prince 1993; MAIN-LEFT/RIGHT, ref.?), listed in (5). PARSE is assigned a violation mark once for every syllable not included in a foot, e.g. the final syllables in (4d) and (4f). FOOTBINARITY is assigned a violation mark by feet with only one mora, e.g. if (4d) were */(ká.pa) (lì)/, or in feet with more than two syllables, e.g. if (4d) were */(ká.pa.li)/.[5] TROCHAIC is assigned a violation mark by right-headed binary feet, e.g. */(ka.pá) li/, while IAMBIC is assigned a violation mark by left-headed binary feet (basically in every form in (4)). ALL-FEET-LEFT and -RIGHT are assigned a violation mark by every foot that is not aligned with the respective word edge, e.g. in */ka (páli)/ (ALL-FEET-LEFT) and in the licit /(ká.pa) li/ (ALL-FEET-RIGHT). These constraints are gradient in that they are assigned one violation mark for every syllable that is between a foot and the designated word edge. Moreover they favour forms with as few feet as possible. MAIN-LEFT and -RIGHT are assigned one violation mark for every syllable between the head foot of a word and the respective word edge, e.g. in */(ŋàl.ku) (nín.pa)/ (MAIN-LEFT) and /(ŋál.ku) (nìn.pa)/ (MAIN-RIGHT).

(5) Constraints

ALL-FEET-LEFT/RIGHT (AFL/AFR): The left/right edge of a foot is aligned with the left/right edge of a word.
FOOTBINARITY (FTBIN): Feet are either bimoraic or disyllabic.
IAMBIC: The rightmost syllable in a foot is the head syllable.
MAIN-LEFT/RIGHT (MAIN-L/MAIN-R): The head foot is aligned with the left/right edge of the word.
PARSE: Every syllable is included in a foot.
TROCHAIC: The leftmost syllable in a foot is the head syllable.

[4] The OT analysis given here differs from the one by Kager (1999). Although he too models Pintupi stress and its learnability, he is mainly concerned with the demonstration of how *mark cancellation* and *constraint demotion* work rather than with a fully-fledged analysis of Pintupi stress. He demonstrates constraint demotion by using one fully structured trainig form, and limits his analysis of Pintupi stress to six constraints FTBIN, PARSE-σ, MAIN-L/R and AFL/AFR, leaving out constraints on foot type. My account gives a more comprehensive OT analysis of Pintupi stress as well as a more elaborate approach to Pintupi stress learnability.
[5] Feet of the size of three syllables or more are banned from GEN here.

Let us evaluate the ranking step by step, starting with a trisyllabic word *kápali* 'mother's mother', as in tableau (6). The underlying form (in pipes'| |') as the input to the evaluation is given in the upper left cell. To ensure that feet are built from left to right, AFL has to be ranked above AFR. The reverse ranking would give for instance */ka (páli)/ rather than the required /(kápa) li/ (or, in the case of longer words, */(pú.ɭiŋ) ka (là.tʲu)/ instead of /(pú.ɭiŋ) (kà.la) tʲu/; provided MAIN-L is ranked high). AFL is violated twice in candidate (6d) because the left foot edge of (lì) is two syllables away from the left word edge. Since the language has a strong-weak pattern, TROCHAIC has to outrank IAMBIC; otherwise the grammar would render */(kapá) li/ instead of /(kápa) li/. FTBIN has to be ranked above PARSE, otherwise the grammar would yield */(kápa) (lì)/ instead of /(kápa) li/.[6] The constraints MAIN-L and MAIN-R are not included in the tableau since in trisyllabic forms there is no competition between feet.

(6) *kapali*: TROCHAIC >> IAMBIC; FTBIN >> PARSE; AFL >> AFR

| ka.pa.li |	TROCHAIC	FTBIN	AFL	IAMBIC	PARSE	AFR
☞ a. / (ká.pa) li /				*	*	*
b. / ka (pá.li) /			*!	*	*	
c. / (ka.pá) li /	*!				*	*
d. / (ká.pa) (lì) /		*!	**	*		*

Turning to four-syllable words like *ŋalkuninpa* in (7), we can establish the ranking of MAIN-L above MAIN-R, to ensure that the left foot in a word gets main stress. PARSE has to be ranked above AFL in order to allow more than one foot in the word, since AFL not only causes feet to be left-aligned, but it also favours forms with as few feet as possible (i.e. it is satisfied in forms with no or at most one foot). PARSE ensures the occurrence of secondary stress; otherwise */(ŋál.ku) nin.pa/ would surface instead of /(ŋál.ku) (nìn.pa)/. What one wouldn't expect is that PARSE also has to outrank IAMBIC. If IAMBIC was ranked above PARSE it would kick out the right candidate /(ŋál.ku) (nìn.pa)/ and leave the decision to the lower ranked constraint AFL, which would decide in favour of candidate */(ŋál.ku) nin.pa/.

[6] Note that the dashed lines of the tableau should not be read as crucial ties. Since there is no evidence so far for the ranking between TROCHAIC, FTBIN and AFL on one side and between IAMBIC, PARSE and AFR on the other, no solid lines are drawn yet. The exclamation marks are set after the first violation mark that is sufficient for eliminating the candidate, i.e. the exclamation mark in candidate (6d) could be just as well placed behind the first violation mark of the constraint AFL.

(7) *ŋalkuninpa*: Main-L >> Main-R; Parse >> AFL >> AFR; Parse >> Iambic

| ŋal.ku.nin.pa |	Trochaic	FtBin	Main-L	Parse	AFL	Main-R	Iambic	AFR
a. / (ŋál.ku) nin.pa /				*!*		**	*	**
b. / ŋal.ku (nín.pa) /			*!*	**	**		*	
c. / ŋal (kú.nin) pa /			*!	**	*	*	*	*
☞ d. / (ŋál.ku) (nìn.pa) /					**	**	**	**
e. / (ŋál.ku) (nìn) pa /		*!		*	**	**		***
f. / (ŋàl.ku) (nín.pa) /			*!*		**		**	**

With this ranking, forms of five or more syllables like *puɭiŋkalat^{j}u* can be evaluated correctly, too (8). Any candidate, i.e. candidates (8b, c, f), that does not align the left edge of the head foot with the left edge of the word is kicked out by Main-L; candidate (8a) is kicked out by Parse, because it has too many unparsed syllables. Candidate (8e) has more violations of AFL, leaving candidate (8d), /(pú.ɭiŋ) (kà.la) t^{j}u/, as the winner.

(8) 5-syllable words: *puɭiŋkalat^{j}u*

| pu.ɭiŋ.ka.la.t^{j}u |	Trochaic	FtBin	Main-L	Parse	AFL	Main-R	Iambic	AFR
a. / (pú.ɭiŋ) ka.la.t^{j}u /				**!*		***	*	***
b. / pu.ɭiŋ.ka (lá.t^{j}u) /			**!*	***	***		*	
c. / pu (ɭíŋ.ka) (là.t^{j}u) /			*	*!	****	**	**	**
☞ d. / (pú.ɭiŋ) (kà.la) t^{j}u /				*	**	***	**	****
e. / (pú.ɭiŋ) ka (là.t^{j}u) /				*	***!	***	**	***
f. / pu (ɭìŋ.ka) (lá.t^{j}u) /			*	*!	****		**	**

So far, we established a suitable partial ranking of the constraints for words with two to five syllables: Trochaic, Parse >> Iambic; Main-L >> Main-R; AFL >> AFR; FtBin >> Parse. However, words with a long initial vowel, e.g. *mi:l^{j}manu* 'whining', cannot be accounted for with this ranking, because it cannot decide between two licit candidates /(mí:l^j.ma) nu/ and */(mí:l^j) (mà.nu)/. The constraint one would regard as competent for the necessary disambiguation, FtBin, cannot decide between these two candidates because both /(mí:l^j.ma) nu/ and */(mí:l^j) (mà.nu)/ satisfy FtBin, as shown in (9). Note that FtBin as a constraint is not formulated like in (3). It does not prefer a disyllabic foot (σσ) over a monosyllabic, yet bimoraic foot ($\sigma_{\mu\mu}$), but is equally satisfied by both kinds of feet. In our ranking, the decision between these two candidates is passed on to Parse, which decides in favour of the wrong candidate */(mí:l^j) (mà.nu)/:

(9) Long initial vowel

| mi:l^j.ma.nu |	Main-L	Trochaic	FtBin	Parse	AFL	Main-R	Iambic	AFR
☹ a. / (mí:l^j.ma) nu /				*!		*	*	*
☞ b. / (mí:l^j) (mà.nu) /					*	**	*	**
/ (mí:l^j) (mà) nu /			*	*	*	**		***

One could think of reversing the ranking between Parse and AFL, but this would lead to problems with forms like *ŋalkuninpa* or *puɭiŋkalat^{j}u*. This means that there is no ranking with these constraints that can account for the Pintupi pattern. The source of the problem might be the constraint Trochaic. In its current definition Trochaic does not decide between the two foot forms (σ́σ) and (σ́). Tesar (1998) reformulated the trochaicity constraint to account for the typological asymmetry in stress systems that trochaic languages may be either quantity-sensitive or quantity-insensitive, but that the majority of iambic languages is quantity-sensitive (Hayes 1995). He redefined Trochaic as FootNonfinal (10):

(10) FootNonfinal (FtNonfin): The foot head is not final in the foot.

In this form FtNonfin punishes monosyllabic feet like (σ́), and favours /(mí:l^j.ma) nu/ over */(mí:l^j) (mà.nu)/. As a by-product, FtNonfin is taking over the function of FtBin, so we will take FtBin out of the tableau.[7] To be able to unfold its full power, FtNonfin has to outrank Parse:

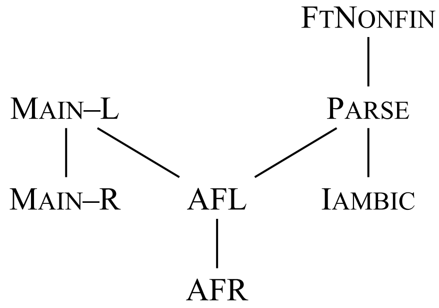
(11) FtNonfin instead of Trochaic

| mi:l^j.ma.nu |	Main-L	FtNonfin	Parse	AFL	Main-R	Iambic	AFR
☞ a. / (mí:l^j.ma) nu /			*		*	*	*
b. / (mí:l^j) (mà.nu) /		*!		*	**	*	**

The crucial ranking is displayed in (12). Main-L is ranked above Main-R so that the left foot within a word gets main stress. Main-L is furthermore ranked above AFL and AFR. AFL is ranked above AFR to ensure that feet are iterated without gaps in longer words. FtNonfin has to outrank Parse to make sure that final syllables remain unfooted in words with an odd number of syllables and to make sure that in words beginning with a long vowel the first two syllables are parsed in the foot (11). FtNonfin and Parse have to be ranked above Iambic; FtNonfin because else iambic feet would surface and Parse because words with only a single foot would surface (7).

[7] It should be noted that FtBin can still play a role in other languages that e.g. have an iambic stress pattern.

(12) A crucial ranking for Pintupi stress with FTNONFIN



Keeping up TROCHAIC

One could keep up an analysis with TROCHAIC by including *CLASH (Kager 1999; this traces back to pre-OT approaches by Liberman 1975, Liberman & Prince 1977, Prince 1983, Hammond 1984, Selkirk 1984).[8] This constraint is commonly employed to prevent stress clashes in two adjacent syllables:

(13) *CLASH: No stressed syllables are adjacent.

If this constraint is included dominating PARSE, the attested candidate becomes optimal (14).

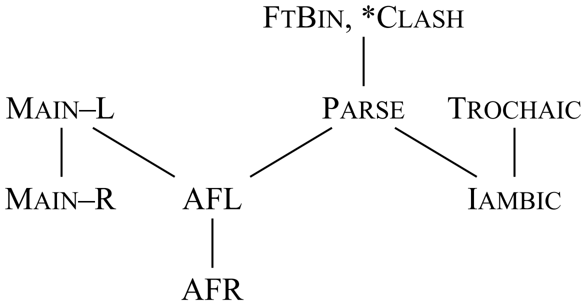
(14) An analysis with *CLASH and TROCHAIC

| mi:l^j.ma.nu |	MAIN-L	TROCHAIC	FTBIN	*CLASH	PARSE	AFL	MAIN-R	IAMBIC	AFR
☞ a. / (mí:l^j.ma) nu /					*		*	*	*
b. / (mí:l^j) (mà.nu) /				*!		*	**	*	**

In the crucial ranking including TROCHAIC (15), *CLASH has to outrank PARSE. For the rest of the analysis, the same crucial rankings apply as in (12). In a constraint set including FTNONFIN, *CLASH would not play a crucial role and could be placed anywhere in the hierarchy.

[8] One could also think of keeping up an analysis with TROCHAIC by splitting up FTBIN into FOOTBISYLLABIC ("feet are bisyllabic") and FOOTBIMORAIC ("feet are bimoraic"). One would think that feet are exactly bimoraic in weight-sensitive languages, while feet are exactly bisyllabic in weight-insensitive languages. In Pintupi FTBISYLLABIC would then have to outrank FTBIMORAIC, TROCHAIC, and PARSE, in order to ensure that feet are strictly bisyllabic. We will not pursue this idea here because it does not seem to be the task of FTBIN to decide what the exact composition of a foot is; FTBIN only demands that a foot be binary. It seems to be the case that other constraints on rhythm type, such as TROCHAIC or FTNONFINAL and IAMBIC, and on weight, such as WEIGHT-BY-POSITION, *Cμ and WEIGHT-TO-STRESS, decide whether a foot in a particular language counts moras or syllables. We will meet these constraints in (27) and (28). See Apoussidou & Boersma (2003) for an approach to the learnability of Latin stress involving FTBIN and FTBIMORAIC.

(15) A crucial ranking for Pintupi stress with TROCHAIC



By now we have established the analysis for Pintupi word stress with two different constraint sets, one including the constraint FTNONFIN (12) and another with a constraint TROCHAIC (15). It is interesting to see which constraint forms the appropriate restriction on trochaic feet by testing the acquisition process using computer simulations. This is discussed in the following sections.

4 What and how to learn

Having a linguistic analysis for a language is one thing, having a learning path to this analysis is another. If there is no learning path to an analysis of an existing language, something must be wrong with the analysis, because the language is obviously learned by the children of the speakers. Learnability provides in this sense a tool for testing the viability of a linguistic analysis. The learnability of metrical structure is intriguing, since the interpretation of the speech stream can be ambiguous. If a listener hears a trisyllabic form with stress on the middle syllable like [σσ́σ], he or she might interpret this form as having either an iambic foot /(σσ́) σ/ or a trochaic foot /σ (σ́σ)/. Adult speakers of a language ideally know how to interpret the form, based on their knowledge about whether the language is trochaic or iambic, but learners of a language do not know that yet.

The learnability of stress has been modelled before (in OT: Tesar & Smolensky 2000 & former versions; Apoussidou & Boersma 2003, 2004; Kager 1999; some non-OT approaches: Dresher & Kaye 1990, Daelemans et al. 1994). The former OT-based simulations on the acquisition of metrical stress demonstrate the learnability from overt forms, i.e. forms that are marked for stress, but not for foot structure. The learning data consisted of strings of light (‘L’) and heavy (‘H’) syllables, e.g. [L1 L L], [L H1 L], [H H1 L L2].

In the simulations of this paper, the virtual learners[9] know that their language, Pintupi, has weight characteristics in that it has phonemic vowel length and, according to e.g. Hayes (1995), a bimoraic word minimum in that monosyllabic words have a long vowel, e.g. *nʲaː* ‘what’.[10] What they have to find out for themselves is that stress is not sensitive to the weight of a syllable. Besides, they have to figure out whether coda consonants are moraic or not. In Pintupi, stress can appear on syllables with coda consonants, giving room for the

---

9 The simulations are conducted with the Praat program (Boersma & Weenink 1992-2005).

10 Hansen & Hansen (1969), whose data were taken over by Hayes (1995), do not explicitly state that there is a bimoraic word minimum in Pintupi. According to Hansen & Hansen, they found only few instances of monosyllabic words; among them only content words seem to have to meet the bimoraic requirement. They name one example of a monosyllabic function word, *ma* ‘action direction indicator’, that consists of a light syllable with a short vowel.

interpretation that these syllables are heavy and attract stress. But there are also syllables with coda consonants that are unstressed, indicating that these syllables are not heavy since they do not attract stress. It is interesting to see whether virtual learners of Pintupi analyse the language as completely weight-insensitive or as partly weight-sensitive, or even as completely weight-sensitive, because the long vowels in the language occur only word-initially and are always stressed. Another question is whether the virtual learners acquire the correct stress pattern, i.e. whether they will have primary stress on the initial syllable of the forms and secondary stress on every other following syllable, with the exception of final syllables in words with an odd number of syllables. Whether the learners will choose the foot structure in (4) or whether they will assign feet differently is subordinate to the question whether stress is assigned to the correct syllable of the word.

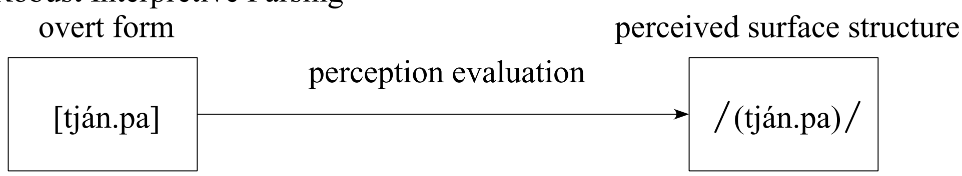
In section 3 two analyses of Pintupi stress were given. They differ in their interpretation of the constraint on trochaicity. To see whether there exists a difference in the learnability of these constraints one half of the virtual learners is equipped with FTNONFIN and the other half is equipped with TROCHAIC. If it turns out that e.g. the learners with FTNONFIN can learn the stress pattern more easily or can learn it at all, and the learners with TROCHAIC fail to learn the pattern it will indicate that TROCHAIC is not a good formulation of the constraint on trochaicity. The same holds for the case where the learners with FTNONFIN fail and the ones with TROCHAIC succeed.

A further issue is how uniform the grammars of the learners are in the final state. If there is the possibility that the learners come up with various ways of assigning foot structure, they will differ in their final constraint rankings. Interesting is also to what extent the learners are able to transfer the learned stress pattern to forms that they have not encountered up to that point. Finally, one half of the learners has Error Driven Constraint Demotion (EDCD; Tesar 1995) as the learning strategy, while the other half has the Gradual Learning Algorithm (GLA; Boersma 1997). What learning algorithm seems to be the better learning strategy? Both learning algorithms are error-driven and make use of the same interpretation method of incoming forms called robust interpretive parsing (RIP; Tesar & Smolensky 1998, 2000), but they differ in the way they rerank the constraints. EDCD allows constraint demotion only, while the GLA admits constraint demotion as well as constraint promotion. The upcoming sections outline the perception mechanism RIP (4.1) and the learning algorithms EDCD (4.2.) and GLA (4.3). The ingredients to the computer simulations are given in section 5, followed by the results in section 6.

4.1 Robust Interpretive Parsing: hearing is believing

In the learning model adopted here the learner uses the same grammar (i.e. constraint ranking) for perception as for production. The learner, as a listener of a language, tries to make sense of the auditory speech signal in that she applies her grammar to what she hears. Spoken in OT terms the speech signal is the input that undergoes the constraint ranking of the learner/listener.

(16) Robust Interpretive Parsing



The evaluation of this overt form is shown in (17) with an overt form [t^ján.pa]. Imagine that the Universal Grammar of our learner consists of the constraints MAIN-L, MAIN-R, PARSE, AFL, AFR, FTBIN, IAMBIC, TROCHAIC, WSP and NONFIN. Imagine also that the learner/listener has a grammar as in (16). The learner/listener interprets the incoming form [t^ján.pa] by assigning foot structure to it. Since the overt form already contains stress, the possibilities to assign foot structure are limited, and only two candidates are considered for evaluation. Candidate (17)a, /(t^ján.pa)/, has a disyllabic foot while candidate (17)b, /(t^ján) pa/, has a monosyllabic foot. It is assumed here that there cannot be a candidate like /(t^jan.pá)/ since in such a candidate stress would be on a different syllable. While the interpretation of an overt form might be ambiguous for the learner in terms of foot structure (because foot structure is not observable in the input), stress location is unambiguous (because stress is observable in the input).[11] The decision for candidate (17)a, /(t^ján.pa)/, as the optimal one (marked with a '👂' in the tableau) is taken by the learner's current constraint ranking.[12] The highest ranked constraint MAIN-L is satisfied by both candidates, since both align the left foot edge with the left word edge. Candidate (16)b /(t^ján) pa/, though, violates the next ranked constraint Main-R, because it does not align the right foot edge with the right word edge. Thus, candidate (17)a /(t^ján.pa)/ wins as the form that the learner/listener perceives. The learner/listener has interpreted the overt form [t^ján.pa] as having the structure /(t^ján.pa)/.

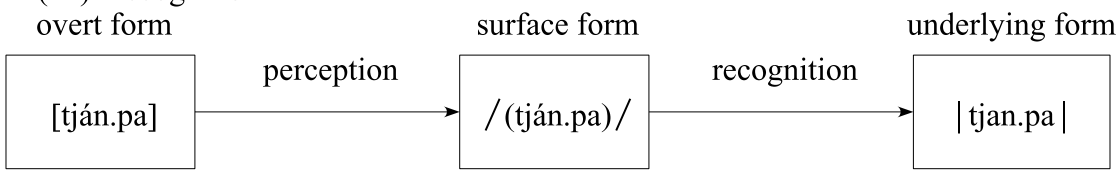
(17) Perception: stress to foot

overt form [t^ján.pa]	MAIN-L	MAIN-R	PARSE	AFL	AFR	FTBIN	IAMBIC	TROCHAIC	WSP	NONFIN
👂 a. / (t^ján.pa) /							*			*
b. / (t^ján) pa /		*!	*		*					

Now that the learner has determined the surface structure /(t^ján.pa)/ of the overt input [t^ján.pa] she can map this perceived form onto a form in her lexicon. In this mapping all metrical structure like feet and stress marks are stripped off the form, leaving behind the segmental and syllabic structure |t^jan.pa|:[13]

[11] The dislocation of stress in perception is possible by violations of corresponding cue constraints. Since cue constraints are not in the scope of this paper, this possibility is excluded.
[12] The constraint ranking is made up for demonstrational reasons.
[13] In real life the acquisition of metrical structure interacts with the acquisition of segmental and syllabic structure and the creation of lexical entries. Once again these issues are left out of consideration since the discussion here is limited to the acquisition of metrical structure alone.

(18) Recognition



From this underlying form the learner is able to evaluate her own production by applying her current grammar, as we see in the following section.

4.2 Virtual production

In 4.1 we established how the learner perceives the overt form [t^ján.pa] as having the surface structure /(t^ján.pa)/. From listening alone the learner will not arrive at the target grammar. She needs to compare her perception to something in order to be able to learn, namely to her production. To do so she does not have to actually vocalize the form she perceives but she will rather produce it *virtually*, i.e. she will compute the form in her head, as soon as she heard it. When the learner is producing this very same form the input to the evaluation is the underlying form |t^jan.pa|, as known from standard OT evaluations. In the production evaluation in (19) of the perceived form /(t^ján.pa)/ (17) there are now more candidates than in the perception evaluation because the learner has to add foot structure *and* stress to the form (note that the candidate set in the perception evaluation is a subset of the one in the production evaluation).[14] Candidate (18a) is the perceived form of (17) and is marked with an ear. This perceived form has trochaic foot structure. However the produced word of the learner has an iambic foot /(t^jan.pá)/, marked by the usual pointing hand. Candidate (18a), /(t^ján.pa)/, the winner of the perception evaluation, is not optimal in production in this case, because it violates the higher-ranked constraint IAMBIC, in contrast to the winning candidate in production (18e), /(t^jan.pá)/.

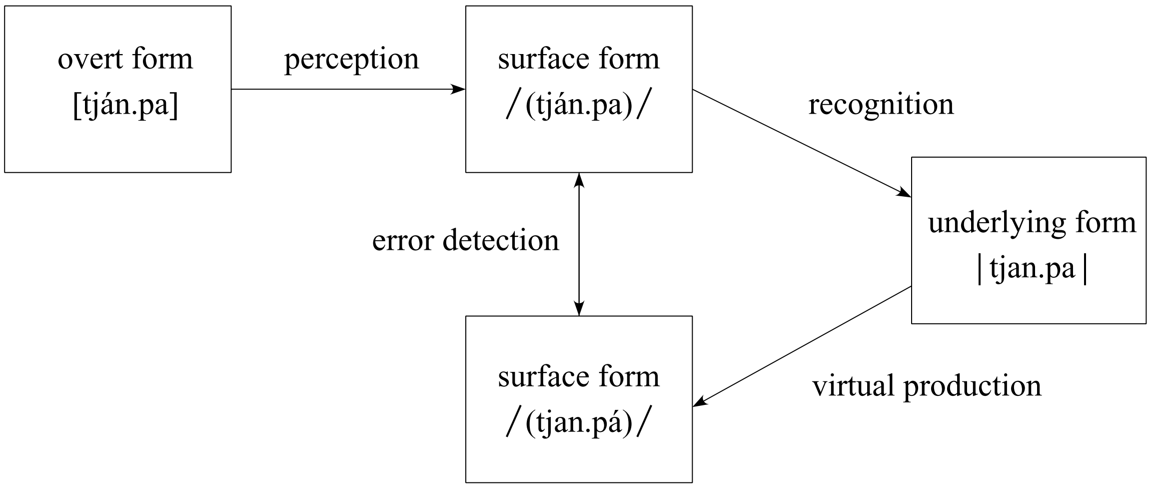
(19) Virtual production: foot to stress

underlying form | t^jan.pa |	MAIN-L	MAIN-R	PARSE	AFL	AFR	FTBIN	IAMBIC	TROCHAIC	WSP[15]	NONFIN
👂 a. / (t^ján.pa) /							*!			*
b. / (t^ján) pa /		*!	*		*					
c. / (t^ján) (pà) /		*!		*	*	*				*
d. / t^jan (pá) /	*!		*	*		*			*	*
☞ e. / (t^jan.pá) /								*	*	*
f. / (t^jàn) (pá) /	*!			*	*	*				*

[14] GEN dictates many more candidates than the ones shown here, but for clarity's sake we will stick to just a few of them.
[15] Let us assume for the moment that Pintupi does have moraic coda consonants. Any unstressed syllable with a coda incurs a violation of WSP then.

Now that the learner evaluated what she would produce for the form she perceived, she can compare the two forms and notice a discrepancy (an *error* as Tesar & Smolensky 1998, 2000 call it) between them. This discrepancy between the perceived form and the produced form impels the learner to change her grammar, until her perception and production match.

(20) Error detection



Two learning strategies are pursued in the computer simulations here: Error Driven Constraint Demotion (EDCD; Tesar 1995) and the Gradual Learning Algorithm (GLA; Boersma 1997). They are outlined in sections 4.3 and 4.4.

4.3 EDCD

As mentioned in 4.1 the learner regards the perceived form as the target she wants to match her production to. If she is using EDCD as her learning strategy she will adjust her constraint ranking as follows: she will look up all constraints that prefer her produced form and lower them below the highest ranked constraint that prefers the form she perceives. This will make it more likely that in a future evaluation of this form, the produced form matches the perceived form. In our example at hand, IAMBIC (21) is the constraint preferring the learner's produced form /(t^jan.pá)/, and TROCHAIC is the constraint preferring the learner's perceived form /(t^ján.pa)/. To make it more likely that the perceived form /(t^ján.pa)/ will also be the winner of the production evaluation IAMBIC is demoted below TROCHAIC (21), into the same stratum as WSP.

(21) Error Driven Constraint Demotion

| tʲan.pa |	MAIN-L	MAIN-R	PARSE	AFL	AFR	FTBIN	IAMBIC	TROCHAIC	WSP	NONFINAL
👂 a. / (tʲán.pa) /							*!			*
b. / (tʲán) pa /		*!	*		*					
c. / (tʲán) (pà) /		*!		*	*	*				*
d. / tʲan (pá) /	*!		*	*		*			*	*
☞ e. / (tʲan.pá) /								*	*	*
f. / (tʲàn) (pá) /	*!			*	*	*				*

It is crucial in EDCD that constraints can only be demoted, not promoted, and only in a minimal way. For instance, WSP is not involved in the demoting process of (18), although it also prefers the perceived form. But the constraints eligible for demotion are placed just below the highest ranked constraint that prefers the perceived form which is TROCHAIC in this case.

(22) A revised EDCD-grammar

| tʲan.pa |	MAIN-L	MAIN-R	PARSE	AFL	AFR	FTBIN	TROCHAIC	IAMBIC	WSP	NONFINAL
☞ 👂 a. / (tʲán.pa) /								*		*
b. / (tʲán) pa /		*!	*		*					
c. / (tʲán) (pà) /		*!		*	*	*				*
d. / tʲan (pá) /	*!		*	*		*			*	*
e. / (tʲan.pá) /							*!		*	*
f. / (tʲàn) (pá) /	*!			*	*	*				*

In our example, the perceived and produced form have been brought to agreement now. The ear and the pointing hand are in the same cell, meaning that the produced form is the same as the perceived form: /(tʲán.pa)/. With this new ranking, the learner still perceives this form as /(tʲán.pa)/ if she encounters it again:

(23) Perception with the new ranking

[t^ján.pa]	MAIN-L	MAIN-R	PARSE	AFL	AFR	FTBIN	TROCHAIC	IAMBIC	WSP	NONFINAL
👂 a. / (t^ján.pa) /								*		*
b. / (t^ján) pa /		*!	*		*					

When this happens with all forms that the learner perceives and produces, learning is terminated. However it is possible that the reranking of constraints leads to a different perception of the same form, which in turn potentially leads to a mismatch of the perceived and produced form. The intermediate encounter of other forms might change the constraint ranking as well in a way that perception and production are out of tune again. In general, though, the encounter of different forms should help the learner to come up with a ranking that creates forms matching the adult output.

4.4 The GLA

While EDCD makes use of an ordinal ranking of constraints, the GLA makes use of Stochastic OT (Boersma 1998). Constraints are assigned real numbers (*ranking values*) on the ranking scale as a measurement of the distance between constraints. In each evaluation of a given form, a little bit of noise is added to the ranking value of each constraint, with the consequence that constraints close to each other can swap their order for this specific evaluation. In addition to that, grammar adjustment in the GLA is a bit different. Here all constraints that prefer the perceived form (IAMBIC) and all constraints that prefer the produced form (this time: TROCHAIC and WSP) are looked up. Consider the same grammar as in (18), with the same constraints and the same ranking (24). The constraints preferring the perceived form are shifted upwards, while the constraints preferring the produced form are shifted down the hierarchy. Irrespectively of TROCHAIC and WSP, IAMBIC is lowered on the ranking scale; irrespectively of IAMBIC, TROCHAIC and WSP are shifted up on the ranking scale.

(24) Grammar adjustment with the GLA

| t^jan.pa |	MAIN-L	MAIN-R	PARSE	AFL	AFR	FTBIN	IAMBIC	TROCHAIC	WSP	NONFINAL
👂 a. / (t^ján.pa) /							*!			*
b. / (t^ján) pa /		*!	*		*					
c. / (t^ján) (pà) /		*!		*	*	*				*
d. / t^jan (pá) /	*!		*	*		*			*	*
☞ e. / (t^jan.pá) /								*	*	*
f. / (t^jàn) (pá) /	*!			*	*	*				*

After some learning has taken place, i.e. after the learner encountered more forms like [tʲán.pa], the constraints TROCHAIC and WSP will swap places with IAMBIC, bringing perception and production into agreement (25).[16]

(25) An adjusted GLA-grammar

| tʲan.pa |	MAIN-L	MAIN-R	PARSE	AFL	AFR	FTBIN	TROCHAIC	WSP	IAMBIC	NONFINAL
👂 ☞ a. / (tʲán.pa) /									*	*
b. / (tʲán) pa /		*!	*		*					
c. / (tʲán) (pà) /		*!		*	*	*				*
d. / tʲan (pá) /	*!		*	*		*		*		*
e. / (tʲan.pá) /							*!	*		*
f. / (tʲàn) (pá) /	*!			*	*	*				*

In the GLA constraints can overlap. Together with the stochastic feature this implicates that constraints can swap places in the hierarchy with a certain probability. However the more learning takes place the constraints will move further and further apart, making it less and less likely that this will happen.

In the simulations, the two learning strategies are compared with respect to their overall success in learning from the forms they are confronted with, but also with respect to their success in transfering what they have learned to forms they were not trained on.

5 Simulating the acquisition of stress

The ingredients to the computer simulations are the training data (5.1), the constraints (5.2) and the learning algorithms EDCD and GLA, which were discussed in 4.2 and 4.3. The training data were the same for all learners, except that each learner encountered the data in a different order. The learners started out with an initial ranking where all constraints were equally ranked (at 100 on the ranking scale). The EDCD learners had a plasticity of 1, meaning that constraints were reranked by 1 (e.g. first to 99 on the ranking scale, then to 98 etc.). They learned with zero evaluation noise and were fed with 1 000 forms. The GLA learners had decreasing plasticity, starting out by 1, with four times a decrement of 0.1. This means that in the beginning, the GLA learners took ranking steps as big as the EDCD learners (rather large ones), but decelerated their learning pace in the course of time. Their evaluation noise was set to 2.0, and they were fed with 40 000 training forms.

[16] This holds under the simplified assumption that the learner hears the same form [tján.pa] over and over again without interference of other forms.

5.1 The learning data

The learning data set consists of 17 word-like forms, listed in (26). These forms are two to four syllables long and are made up of syllables and stress marks. Monosyllabic forms are excluded from the set since there is only one possibility to stress a monosyllabic word, therefore the virtual child cannot learn much about stress placement from it. The length of words is limited to at most four syllables because of two reasons: to account for the claim that child-directed speech often contains simplified utterances (Phillips 1973) and to see what the learners will do when they are asked to produce forms that they have not been trained on. The learning set covers only a selection of all possible combinations of syllables in Pintupi, imitating an impoverished input.[17] All forms are overt, i.e. they contain stress marks, but neither foot nor moraic structure. Because we focus on the acquisition of stress here, the overt forms furthermore contained syllable boundaries.[18] Each vowel in a syllable is inherently moraic, coda consonants are not marked as such.

(26) The training data

2-syllable forms	3-syllable forms	4-syllable forms
[cv́.cv]	[cv́.cv.cv]	[cv́.cv.cv̀.cv]
[cv́.cvc]	[cv́.cv.cvc]	[cv́.cvc.cv̀.cv]
[cv́v.cv]	[cv́.cvc.cv]	[cv́.cvc.cv̀c.cvc]
[cv́c.cv]	[cv́c.cv.cv]	[cv́c.cv.cv̀.cv]
[cv́c.cvc]	[cv́c.cv.cvc]	[cv́c.cv.cv̀c.cv]
	[cv́c.cvc.cv]	[cv́c.cvc.cv̀c.cv]

The EDCD learners heard 1 000 tokens of these forms, so that they heard each of the 17 types about 59 times. The data were presented in random order. The GLA learners heard 40 000 tokens, so they heard each of the 17 types about 2353 times. The reason why the EDCD learners were trained on 1 000 forms and the GLA learners were trained on 40 000 is due to the different plasticities of the learners. The GLA learners have decreasing plasticity, i.e. they decelerate their learning pace, and therefore need more data to learn from, since constraints shift only small amounts on the ranking scale.

5.2 The constraints

If the claims about UG are true, i.e. if constraints are innate, as the general interpretation is (e.g. Tesar & Smolensky 1993:1), then language learners have a larger set of constraints at their disposal than just the ones established in section 3. The consequence for a theory of learning is that the constraints come for free, and that learners only have to find out the ranking of their particular language. Therefore the virtual learners here are equipped with more constraints than they would need to analyse the Pintupi stress system. The constraint set

[17] Only forms that were found as examples in the Hansen & Hansen (1969) paper were selected for the learning data set.
[18] In fact, children also have to learn to set syllable boundaries as well as what kind of syllable structure their language allows, e.g. whether their language allows for codas and/or consonant clusters etc. This issue is beyond the scope of this paper and would have to be modelled elsewhere.

used in the computer simulations here is an extended version of the Tesar & Smolensky (2000) constraint set. Their original constraint set includes twelve constraints on metrical phenomena, shown in (27), and is supposed to cover a large range of stress systems. Some of the constraints have already been explained in section 3, but are listed again in (27) for the sake of completeness.

(27) Original constraint set

AFL/AFR: The left/right edge of a foot is aligned with the left/right edge of a word.
FTBIN: Feet are binary on the mora or syllable level.
FTNONFIN: The foot head is not final in the foot.
IAMBIC: The rightmost syllable in a foot is the head syllable.
MAIN-L/MAIN-R: The leftmost/rightmost foot in a word is the head foot.
NONFIN: The final syllable is not included in a foot.
PARSE: Every syllable is included in a foot.
WORD-FOOT-LEFT/RIGHT (WFL/WFR): The left/right word edge is aligned with a foot.
WEIGHT-TO-STRESS PRINCIPLE (WSP): Heavy syllables are stressed.

NONFIN (inferred from Liberman & Prince 1977; Hayes 1980) is the constraint on extrametricality and is violated by any final syllable included in a foot, whether the syllable is stressed or not. WFL/WFR are alignment constraints, but are not gradient: WFL/WFR are violated for unfooted word edges irrespective of how many syllables between the foot and the word edge are not incorporated into a foot. WSP (Prince 1990) is violated by any unstressed heavy syllable.

To be able to model the learning data that contain syllable structure instead of syllables already marked for light or heavy, some constraints (28) are added to the original constraint set of Tesar and Smolensky:

(28) Additional constraints

*C_μ: Coda consonants are not moraic.
*CLASH: Stressed syllables are not adjacent.
*LAPSE: There are not more than two consecutive unstressed syllables.
WEIGHT-BY-POSITION (WBP): Coda consonants are moraic.

*C_μ (Broselow, Chen & Huffman 1997) militates against moraic coda consonants. Adjacent stress clashes violate a constraint *CLASH (cf. Prince 1983, OT-versions by e.g. Kager 1994; Green & Kenstowicz 1995; Kenstowicz 1995), while *LAPSE (Elenbaas & Kager 1999) is violated in forms that have more than three consecutive unstressed syllables. WEIGHT-BY-POSITION (cf. Hayes 1989) is violated by any coda consonant that does not have a mora assigned to it.

With these additional constraints, learners are enabled to decide whether coda consonants are moraic or not. To see whether there is a difference between learning with FTNONFIN or with TROCHAIC, half of the learners are equipped with the 16 constraints mentioned in (27) and (28), and half of them are equipped with all the constraints except that TROCHAIC is replaced by FOOTNONFINAL. In the initial state all constraints are ranked equal.

This results in four different types of learners. Since each learner encounters the data in a different order, variation in the results is anticipated (see also Apoussidou & Boersma 2003, 2004). To cover differences in the learning results, 50 learners of each learning type were created, resulting in a total of 200 virtual learners:

(29) Learning types:

	TROCHAIC set	FTNONFIN set
EDCD learning strategy	50 learners	50 learners
GLA learning strategy	50 learners	50 learners

Apart from the rankings established in (12) and (15) it is expected that the constraints *LAPSE and WFL/WFR show little effect in the outcome, since an analysis of Pintupi stress is not depending on them. NONFINAL should not show much of an influence, either, and should be ranked below PARSE, but since the data do not give explicit evidence against extrametricality, there might be some effects. A linguist would analyse Pintupi as not having moraic codas and therefore expect *Cμ being ranked above WBP. Even if it turns out that the learners analyse codas as being moraic, WSP should not show an effect because stress assignment is definitely weight-insensitive. *CLASH and FTBIN should only play a role in learners that have the TROCHAIC constraint set.

6 Results

The results were taken from the learnes after their training on 1 000 (for the EDCD learners) and 40 000 forms (for the GLA learners) respectively. All 200 learners produced the same overt output (30). Learning can be considered successful because stress is correctly on the first syllable in all forms and on the third syllable (secondary stress) in forms with four syllables. We see in (30) that the learners did not only produce the forms correctly that they have been trained on, but also the forms up to four syllables which are logically possible in Pintupi. The forms that occurred in the training data are printed in bold.

(30) The overt output of all 200 virtual learners

disyllables	trisyllables	quadrisyllables	
[cv́.cv]	**[cv́.cv.cv]**	**[cv́.cv.cv̀.cv]**	[cv́v.cvc.cv̀.cv]
[cv́.cvc]	**[cv́.cv.cvc]**	[cv́.cv.cv̀.cvc]	[cv́v.cvc.cv̀.cvc]
[cv́v.cv]	**[cv́.cvc.cv]**	[cv́.cv.cv̀c.cv]	[cv́v.cvc.cv̀c.cv]
[cv́v.cvc]	[cv́.cvc.cvc]	[cv́.cv.cv̀c.cvc]	[cv́v.cvc.cv̀c.cvc]
[cv́c.cv]	[cv́v.cv.cv]	**[cv́.cvc.cv̀.cv]**	**[cv́c.cv.cv̀.cv]**
[cv́c.cv]	[cv́v.cv.cvc]	[cv́.cvc.cv̀.cvc]	[cv́c.cv.cv̀.cvc]
	[cv́v.cvc.cv]	[cv́.cvc.cv̀c.cv]	**[cv́c.cv.cv̀c.cv]**
	[cv́v.cvc.cvc]	**[cv́.cvc.cv̀c.cvc]**	[cv́c.cv.cv̀c.cvc]
	[cv́c.cv.cv]	[cv́v.cv.cv̀.cv]	[cv́c.cvc.cv̀.cv]
	[cv́c.cv.cvc]	[cv́v.cv.cv̀.cvc]	[cv́c.cvc.cv̀.cvc]
	[cv́c.cvc.cv]	[cv́v.cv.cv̀c.cv]	**[cv́c.cvc.cv̀c.cv]**
	[cv́c.cvc.cvc]	[cv́v.cv.cv̀c.cvc]	[cv́c.cvc.cv̀c.cvc]

However, the learners came up with five different ways of assigning a foot structure that results in the correct overt stress pattern. An overview is given in (31). The numbers refer to the number of learners that came up with the analysis at hand. 103 learners (shown in (27)a) came up with the foot structure proposed by linguists ("linguist's favourite"). 5 learners analysed coda consonants consistently as moraic (27b). 24 learners analysed only codas in stressed syllables as moraic (27c). 40 learners analysed final syllables as being extrametrical (27d), and 28 learners analysed codas as being moraic and final syllables as being extrametrical (27e). The different analyses are discussed in sections 6.1-6.5.

(31) Distribution of analyses

Analysis	Examples	GLA learners		EDCD learners		Total
		FtNonf	Trochaic	FtNonf	Trochaic	
a. Linguist's favourite:	/ (cv́c.cv) cvc / / (cv́c.cv) (cv̀c.cv) /	23	3	30	47	103
b. Moraic codas:	/ (cv́c$_{\mu}$.cv) cvc$_{\mu}$ / / (cv́c$_{\mu}$.cv) (cv̀c$_{\mu}$.cv) /	5				5
c. Moraic codas in stressed syllables:	/ (cv́c$_{\mu}$.cvc) cvc / / (cv́c$_{\mu}$.cvc) (cv̀c$_{\mu}$.cv) /	1		20	3	24
d. Extrametricality:	/ (cv́c.cv) cvc / / (cv́c.cv) (cv̀c) cv /	11	29			40
e. Extrametricality & moraic codas:[19]	/ (cv́c$_{\mu}$.cv) cvc$_{\mu}$ / / (cv́c$_{\mu}$.cv) (cv̀c$_{\mu}$) cv /	10	18			28

The GLA/FtNonfin learners invented five different ways to realize the Pintupi stress pattern. The learners with the GLA/Trochaic combination came up with three of the analyses that the GLA/FtNonfin learners invented, while the EDCD learners came up with two analyses.[20]

6.1 The linguist's favourite

Back in section 3, we established the foot structure linguists would assign to the Pintupi stress pattern: disyllabic, left-headed feet assigned iteratively from left to right, unfooted final syllables in odd-numbered words, and non-moraic coda consonants. 103 learners came up with this pattern, as shown in (32):

[19] One learner analysed final syllables as extrametrical and codas as moraic, but was not consistent in that within one form, some codas were moraic and some were not, independently of whether the codas occurred in stressed syllables. This learner is subsumed under category (21e).

[20] For the determination of the GLA learners' outputs the evaluation noise was set to zero, meaning that the ranking they displayed on check-up after learning was frozen in as the final ranking.

(32) Examples for linguist’s favourite foot structure

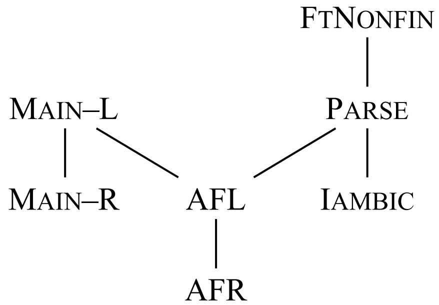
overt forms	surface forms
[cv́.cvc]	/ (cv́.cvc) /
[cv́v.cv]	/ (cv́v.cv) /
[cv́c.cv.cvc]	/ (cv́c.cv) cvc /
[cv́v.cv.cv̀.cvc]	/ (cv́v.cv) (cv̀.cvc) /

The grammars, i.e. constraint rankings, that the learners came up with to produce this pattern are discussed next. In 6.1.1 the rankings are shown of the FtNonfin learners, and in 6.1.2 the rankings of the Trochaic learners.

6.1.1 FtNonfin learners

23 GLA/FtNonfin learners and 30 EDCD/FtNonfin learners came up with a constraint ranking that assigned feet in the way described above. In order to be able to compare the rankings of the learners with that established in section 3, the ranking of (12) is repeated here as (33):

(33) A crucial ranking for an analysis of FtNonfin learners



GLA learners

A final constraint ranking typical for a GLA/FtNonfin learner is shown in (34). Important is whether the crucial rankings are maintained: Main-L outranks Main-R and AFL; FtNonfin outranks Parse and Iambic. Parse in turn outranks Iambic and AFL, and AFL dominates AFR. *Cμ is ranked above WBP, therefore coda consonants surface as non-moraic.

Let us have a look at the ranking of the constraints that a linguist would not invoke for an analysis of Pintupi, but that a language learner came up with. *Lapse ends up high in the hierarchy for both GLA- and EDCD learners in (34) and (35). The training data give little room to the interpretation that stress lapses are allowed in Pintupi, indicating that this constraint has to be high-ranking. WFL ends up rather high because all the forms that the learners will perceive have initial stress, meaning that the first syllable in a word is always footed (no foot = no stress; a stressed syllable is always incorporated in a foot). The ranking of *Clash does not really matter in the constraint set that includes FtNonfin, since FtNonfin is high-ranking and will therefore prevent stress clashes anyway. Nonfinal is crucially ranked below Parse in the GLA learner. The EDCD learner on the other hand has this constraint ranked on the same stratum as PARSE. Still, words with an even number of syllables surface without extrametricality, because FTNONFIN outranks NONFIN. Final syllables in words with an odd number of syllables remain unfooted not due to NONFIN but due to FTNONFIN. Last but not least, WSP is ranked above WBP; this has no effect because *Cμ outranks WBP; therefore weight-sensitivity effects of stress do not surface in syllables with coda consonants. Whether WSP has an effect in forms with long vowels is hard to tell, since long vowels only occur word-initially in Pintupi, being therefore automatically stressed. This is probably the reason why it could shift upwards on the hierarchy of the GLA learner, while it stays undemoted in the EDCD learner. We will see in section 7 what the learners do when asked to produce forms with non-initial long vowels.

(34) A GLA/FTNONFIN learner

FTNONFIN	116.116
*LAPSE	111.136
MAIN-L	110.243
WFL	108.154
FTBIN	106.977
PARSE	105.028
*Cμ	104.139
*CLASH	104.089
NONFINAL	101.947
AFL	101.028
WSP	100.915
WFR	98.053
WBP	95.861
AFR	95.017
MAIN-R	93.846
IAMBIC	81.759

(35) An EDCD/FTNONFIN learner

*CLASH	100
*LAPSE	
AFR	
FTBIN	
FTNONFIN	
MAIN-L	
WFL	
WSP	
*Cμ	99
AFL	
IAMBIC	
MAIN-R	
NONFINAL	
PARSE	
WFR	
WBP	98

EDCD learners

A constraint ranking typical for an EDCD/FTNONFIN learner is shown in (35). The EDCD learners make use of crucial ties (e.g. Tesar & Smolensky 2000:38), which is why they can come up with the correct stress assignment although the constraints are not totally ranked. Crucial ties mean that the violations of all constraints within one stratum are summed up, and the candidate with the least violations in that stratum is the most harmonic one. Nevertheless, the crucial ranking of MAIN-L over MAIN-R and AFL is accomplished, as well as the ranking of FTNONFIN above IAMBIC and PARSE. However, PARSE is ranked on the same stratum as AFL and IAMBIC, and AFR even outranks PARSE and AFL. Other learners that came up with this foot structure had the three constraints ranked on the same stratum. How come that the desired foot structure still shows? The cause for this is that the higher-ranked constraints already eliminate all candidates where the ranking between PARSE and the two alignment constraints could make a decision. Consider tableau (7) again. The line of reasoning was that PARSE has to outrank AFL, otherwise you would get */(ŋál.ku) nin.pa/ instead of /(ŋál.ku) (nìnpa)/; AFL has to outrank AFR to make sure that all feet are left-aligned. The reason why /(ŋál.ku) (nìnpa)/ still surfaces instead of */(ŋál.ku) nin.pa/, although AFR outranks PARSE *and* AFL is that /(ŋál.ku) (nìnpa)/ violates only AFR of all the constraints in that stratum (two violations in total), while */(ŋál.ku) nin.pa/ violates AFR and *LAPSE (three violations in total). Tableau (36) shows the evaluation of *ŋálkunìnpa*, with the constraint ranking of the learner in (35). The candidates of interest are displayed in bold. The first stratum of the hierarchy contains *CLASH, *LAPSE, AFR, FTBIN, FTNONFIN, MAIN-L, WFL, and WSP. The learner at hand saw no reason to demote those constraints. Candidates (34abce) all have three or more violations in that stratum, so candidate (34d), /(ŋál.ku) (nìn.pa)/, wins although it violates AFR twice.

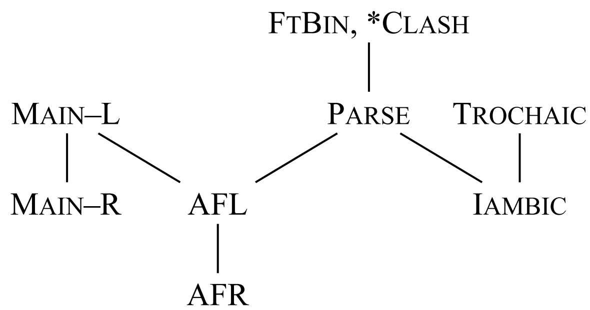
(36) *ŋalkuninpa* 'eating'

| ŋal.ku.nin.pa |	*CLASH	*LAPSE	AFR	FTBIN	FTNONFIN	MAIN-L	WFL	WSP	*Cμ	AFL	IAMBIC	MAIN-R	NONFINAL	PARSE	WFR	WBP
a. / (ŋál.ku) nin.pa /		*	**!								*	**		**	*	**
b. / ŋal.ku (nín.pa) /						**!	*			**	*		*	**		**
c. / ŋal (kú.nin) pa /			*			*	*!			*	*	*		**	*	**
☞ **d. / (ŋál.ku) (nìn.pa) /**			**							**	**	**	*			**
e. / (ŋál.ku) (nìn) pa /			***!	*	*					**	*	**		*	*	**

6.1.2 TROCHAIC learners

Three GLA/TROCHAIC learners and 47 EDCD/TROCHAIC learners came up with the linguist's favourite analysis. The ranking of (15) is for convenience repeated as (37):

(37) A crucial ranking with TROCHAIC



GLA learners

In the grammar of the showcase GLA learner in (38), MAIN-L outranks MAIN-R and AFL. TROCHAIC outranks IAMBIC, while FTBIN and *CLASH outrank PARSE. PARSE in turn dominates AFL that is ranked above AFR. TROCHAIC is top-ranking, while IAMBIC is ranked at the bottom. Maybe in a lot of perception/production evaluations trochees were perceived and iambs were produced; in this way the constraints could have been shifted to the ends of the hierarchy. Each encounter of a learning datum that is perceived as trochaic but produced as iambic would cause FTNONFINAL/TROCHAIC to go up and IAMBIC to go down the hierarchy. If we have a close look at the numbers, we can see that IAMBIC has fallen much farther down the hierarchy than TROCHAIC has gone up; this is due to the interaction with other constraints like e.g. PARSE, FTBIN and NONFINAL, which made IAMBIC go down the hierarchy as well.

(38) A GLA/TROCHAIC learner

TROCHAIC	117.881
*LAPSE	117.048
MAIN-L	117.027
*Cμ	116.819
FTBIN	116.740
WFL	114.249
*CLASH	112.413
PARSE	107.834
NONFINAL	106.936
WSP	100.000
MAIN-R	94.073
WFR	93.064
AFL	92.415
AFR	90.349
WBP	83.181
IAMBIC	57.590

(39) An EDCD/TROCHAIC learner

*Cμ	100
*CLASH	
*LAPSE	
AFL	
AFR	
FTBIN	
MFL	
NONFIN	
PARSE	
TROCHAIC	
WFL	
WFR	
WSP	
IAMBIC	99
MFR	
WBP	

EDCD learners

In the EDCD learners (e.g. the showcase in (39); strata are marked by solid lines), the crucial rankings are not all borne out. MAIN-L is ranked above MAIN-R, *Cμ is ranked above WBP, and TROCHAIC is ranked above IAMBIC; but FTBIN, *CLASH, PARSE, AFL and AFR are ranked on the same stratum. Again, the exclusion of the non-optimal candidates is taken care of by the total amount of violations on the first stratum.

6.1.3 Summing up

The ranking of FTBIN over PARSE is not always maintained: some learners had these constraints reversed (some GLA/FTNONFIN LEARNERS), or ranked on the same stratum (some EDCD/TROCHAIC LEARNERS). It was already mentioned in section 3 that learners with the FTNONFIN constraint actually do not need FTBIN. High-ranking FTNONFIN will ensure that trochaic feet are disyllabic, and degenerate feet like (cv́) are prohibited.[21] In that way FTNONFIN takes over the function of FTBIN and even *CLASH, so the ranking between PARSE and FTBIN becomes irrelevant. This also explains why FTBIN and PARSE are often very close to each other in terms of ranking values across learners of all conditions. In the EDCD learners, it is due to the crucial ties that the required forms still surface.

[21] Trisyllabic feet are not included in GEN, since the maximum size of feet is not under consideration here. Therefore, the only function of FTBIN here is to prohibit monosyllabic feet.

6.2 Moraic coda consonants

Five of the 50 GLA/FtNonfin learners analysed coda consonants as being moraic (regardless whether they occur in stressed syllables or not). None of the other types of learners came up with this analysis. The foot structure that these learners assign is the same as for the linguist's favourite analysis, but coda consonants are consistently analysed as being heavy, marked with a subscript μ (40). Stress assignment in these forms is clearly weight-insensitive.

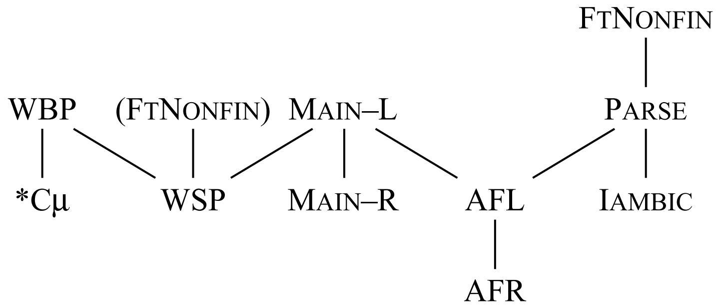
(40) Coda consonants analysed as moraic

overt forms	surface forms
[cv́.cvc]	/ (cv́.cvc$_{\mu}$) /
[cv́c.cvc.cv]	/ (cv́c$_{\mu}$.cvc$_{\mu}$) cv /
[cv́.cvc.cv̀c.cv]	/ (cv́.cvc$_{\mu}$) (cv̀c$_{\mu}$.cv) /

A prototypical ranking of the GLA/FtNonfin learners is displayed in (41), with the crucial ranking of WBP over *Cμ and WSP, which is responsible for the moraic analysis of these learners.

(41) A GLA/FtNonfin learner with moraic codas

FtNonfin	115.277
*Lapse	110.722
Main-L	109.880
WFL	107.706
FtBin	105.024
Parse	104.287
*Clash	103.378
WBP	**102.396**
NonFinal	100.906
WFR	99.094
AFL	98.665
***Cμ**	**97.604**
WSP	**96.065**
AFR	91.244
Main-R	89.070
Iambic	82.690



The difference to the linguist's favourite analysis is that WBP is ranked above *Cμ. Remember that WBP requires coda consonants to be moraic, while *Cμ militates against moraic codas. WSP, the constraint favouring stressed heavy syllables, is crucially ranked below the constraints FtNonfin and Main-L. A reverse ranking would bring about stress sensitivity to heavy syllables (i.e. syllables with coda consonants). In tableau (42) the ranking of the learner in (41) is taken to illustrate the moraic evaluation of a quadrisyllabic word containing a coda consonant. The discussion is limited to two candidates that share the same foot structure but differ in the moraicity of their coda consonants. The first candidate */(yá.lin) (t^{j}à.ra)/ is ruled out by WBP. The second candidate /(yá.lin$_{\mu}$) (t^{j}à.ra)/ wins because of its moraic coda.

(42) *yalint^j ara* ‘north’

| ya.lin.t^j a.ra |	FtNonf	*Lapse	Main-L	WFL	FtBin	Parse	*Clash	WBP	NonFin	WFR	AFL	*Cμ	WSP	AFR	Main-R	Iambic
/ (yá.lin) (t^j à.ra) /								*!	*		**			**	**	**
☞ / (yá.lin$_\mu$) (t^j à.ra) /									*		**	*	*	**	**	**

The learners of this analysis assign stress clearly weight-insensitively, since the weight of the syllable with the moraic coda does not attract stress.

6.3 Moraic coda consonants in stressed syllables only

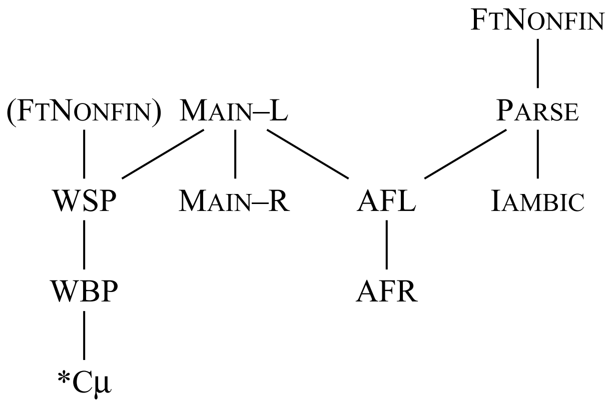
One GLA/FtNonfin learner, 20 EDCD/FtNonfin learners and 3 EDCD/Trochaic learners analysed codas only in stressed syllables as moraic (43). Foot structure was perfectly disyllabic and trochaic, assigned from left to right, just like in the linguist’s favourite analysis.

(43) Moraic codas in stressed syllables

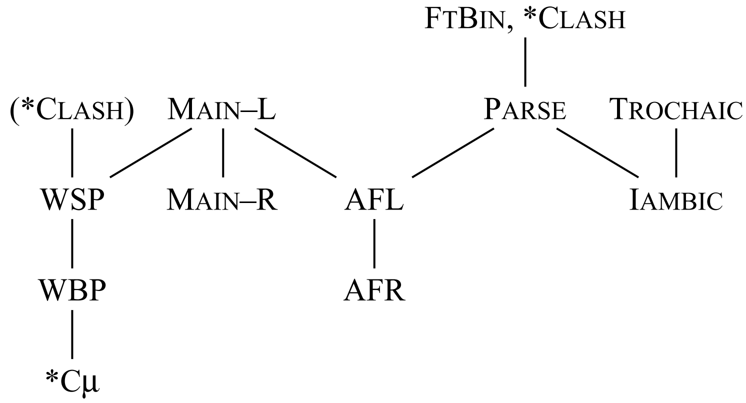
overt forms	surface forms
[cv́c.cvc]	/ (cv́c$_\mu$.cvc) /
[cv́c.cv.cvc]	/ (cv́c$_\mu$.cv) cvc /
[cv́v.cvc.cv̀c.cv]	/ (cv́v.cvc) (cv̀c$_\mu$.cv) /

These learners drew the conclusion that only codas in stressed positions are heavy, which is not surprising. This effect is known as ‘stress-to-weight’ (Myers 1987; Prince 1990), and has its cause in the ranking of WSP above WBP, which has to be ranked above *Cμ in turn. The other constraints are ranked as in the other analyses:

(44) Crucial ranking for moraic codas only in stressed syllables with FtNonfin



(45) Crucial ranking with TROCHAIC



We already observed that only if WBP is ranked above *Cμ, coda consonants can be moraic. If WSP now outranks WBP and *Cμ, only coda consonants that occur in a stressed syllable are moraic, because WSP would cause codas in unstressed syllables not to be moraic. The rankings for prototypical EDCD learners are shown in (47) and (48). We can see here that PARSE, AFL and AFR can be ranked on the same stratum without making a difference in the foot structure of the outputs. The ranking of the single GLA/FTNONFIN learner is shown first in (46).

(46) GLA/FTNONFIN

FTNONFIN	115.849
*LAPSE	111.171
MAIN-L	110.192
WFL	109.149
PARSE	106.133
FTBIN	105.182
*CLASH	103.020
WSP	**102.254**
WBP	**100.892**
NONFINAL	100.866
WFR	99.134
***Cμ**	**99.108**
AFL	94.722
MAIN-R	90.932
AFR	89.728
IAMBIC	80.273

(47) EDCD/FTNONFIN

*CLASH	100
*LAPSE	
FTBIN	
FTNONFIN	
WFL	
WSP	
AFL	99
AFR	
IAMBIC	
MAIN-R	
NONFINAL	
PARSE	
WFR	
WBP	
AFL	
***Cμ**	98

(48) EDCD/TROCHAIC

*CLASH	100
*LAPSE	
AFL	
AFR	
FTBIN	
MFL	
NONFINAL	
PARSE	
TROCHAIC	
WFL	
WFR	
WSP	
IAMBIC	99
MFR	
WBP	
***Cμ**	98

If we take the grammar of the GLA learner in (46) as a basis for evaluation (49), we can see that a candidate where all codas are moraic is ruled out by WSP, while the candidate without moraic codas is ruled out by an additional violation of WBP. The candidate with a moraic coda in stressed position is optimal in this case:

(49) *puḻiŋkalpi* ‘(he fell) finally at the hill’

| pu.ḻiŋ.kal.pi |	FtNonf	*Lapse	Main-L	WFL	Parse	FtBin	*Clash	WSP	WBP	Nonfin	WFR	*Cμ	AFL	Main-R	AFR	Iambic
/ (pú.ḻiŋ$_{\mu}$) (kàl$_{\mu}$.pi) /								*!		*		**	**	**	**	**
☞ / (pú.ḻiŋ) (kàl$_{\mu}$.pi) /									*	*		*	**	**	**	**
/ (pú.ḻiŋ) (kàl.pi) /									**!	*			**	**	**	**

This ‘stress-to-weight’ effect is commonly ascribed to the constraint STRESS-TO-WEIGHT (“if stressed, then heavy”; e.g. Kager 1999). The factorial typology of the three constraints WBP, WSP and *Cμ is sufficient to abandon an additional constraint like STRESS-TO-WEIGHT, though. If we have a look at (48), we see that a ranking of *Cμ over WBP prevents coda consonants from being analysed as moraic. The competing candidates with moraic codas are ruled out by high ranking *Cμ.

(50) *Cμ (WSP) >> WBP (WSP): Codas are not moraic

| cvc.cvc.cv |	*Cμ	WBP	WSP
☞ / (cv́c.cvc) cv /		**	
/ (cv́c$_{\mu}$.cvc$_{\mu}$) cv /	*!*		*
/ (cv́c$_{\mu}$.cvc) cv /	*!	*	

The ranking of WSP does not matter in this case; if coda consonants are not moraic, they do not make a syllable heavy, and only heavy syllables can be stressed by WSP. It can in this case, though, stress syllables with long vowels. This captures a typology with gradience in weight (see Gordon 2002 for a three-way weight hierarchy CVV > CVC > CV found in Klamath, Barker 1964; Kashmiri, Kenstowicz 1994; Chickasaw, Munro & Willmond 1994; and Yapese, Jensen 1977).

If WBP outranks both *Cμ and WSP, all codas within a word are moraic, whether the syllables that contain them are stressed or not:

(51) WBP >> *Cμ, WSP: All codas are moraic

| cvc.cvc.cv |	WBP	*Cμ	WSP
/ (cv́c.cvc) cv /	*!*		
☞ / (cv́c$_{\mu}$.cvc$_{\mu}$) cv /		**	*
/ (cv́c$_{\mu}$.cvc) cv /	*!	*	

In (49) it is crucial that WSP is ranked below WBP, since WSP would require that all unstressed syllables with coda consonants are light. Ranking WSP above WBP gives us exactly this effect, that only stressed syllables can be heavy:

(52) WSP >> WBP >> *Cμ: "Stress-to-weight"

| cvc.cvc.cv |	WSP	WBP	*Cμ
/ (cv́c.cvc) cv /		**!	
/ (cv́c$_\mu$.cvc$_\mu$) cv /	*!		**
☞ / (cv́c$_\mu$.cvc) cv /		*	*

This weight typology corresponds roughly to the one by Morén (2000), which is much more fine-grained. Our *Cμ constraint here translates in his approach to a whole family of structural constraints that distinguish between different sonority classes. Morén's analysis includes furthermore faithfulness constraints to underlying moraic elements, which is captured here by incorporating long vowels into GEN.

6.4 Final syllable extrametricality

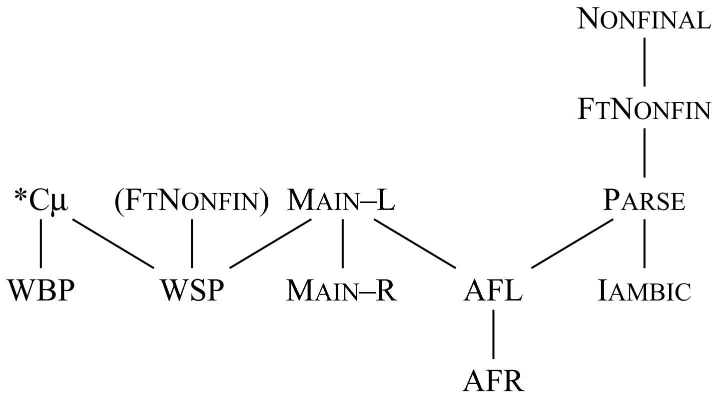
As established in section 3, final syllables in Pintupi are unfooted if the word has an odd number of syllables, because a monosyllable is too small to be parsed into a foot. This results in some forms with syllable extrametricality by accident, / (cv́v.cv) <cv> /, so to say. A linguist would not analyse forms like that as being the result of the involvement of a constraint like NONFINAL. However, learners of the language could misinterpret those forms as an occurrence of extrametricality caused by NONFINAL. And indeed, eleven of the 50 GLA/FTNONFIN and 29 of the GLA/TROCHAIC learners analysed *all* final syllables as being extrametrical, even at the cost of having degenerate feet (53).

(53) All final syllables extrametrical

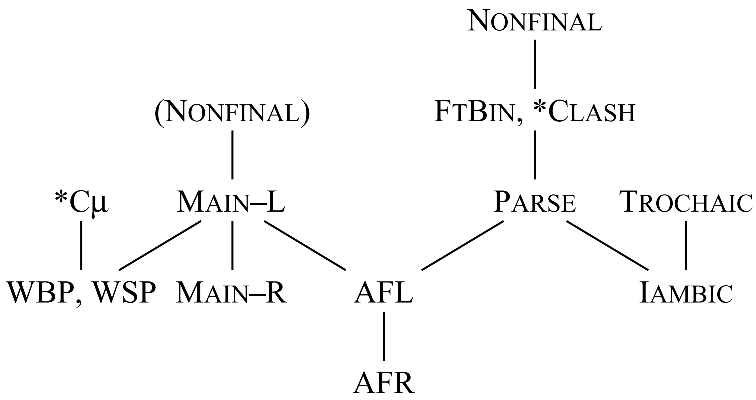
	overt forms	surface structures
a.	[cv́.cv]	/ (cv́) cv /
b.	[cv́v.cv.cvc]	/ (cv́v.cv) cvc /
c.	[cv́v.cv.cv̀.cvc]	/ (cv́v.cv) (cv̀) cvc /

This is not surprising if one considers the Pintupi data and bears in mind that learners have the constraint NONFINAL to their disposal: the learners might perceive forms as having syllable extrametricality, because all words have unstressed final syllables. So some of the learners ended up with a ranking that put NONFINAL above FTNONFIN/TROCHAIC:

(54) A crucial ranking for extrametricality with FtNonfin



(55) A crucial ranking for extrametricality with Trochaic



Across the board, learners with this kind of analysis had Parse outranking FtBin, as in the learners in (56) and (57). How come? FtBin is pushed downwards the hierarchy once Nonfinality is high, because learners will interpret di- and quadrisyllabic forms probably as having monosyllabic feet, and will produce them as such. These forms then constantly violate FtBin. The same applies to Parse: each form that contains an extrametrical syllable violates Parse once. FtBin is below Parse, since the learners probably came up with forms containing a lot of degenerate feet. These forms would violate FtBin, but not Parse.

(56) A GLA/FtNonfin learner

NonFinal	117.690
*Lapse	111.671
Main-L	110.130
WFL	109.076
Parse	107.054
*Cμ	104.002
*Clash	103.017
FtNonfin	101.627
WSP	98.269
WBP	95.998
AFL	90.765
FtBin	90.762
Main-R	89.776
Iambic	88.099
WFR	82.310
AFR	80.269

(57) Another GLA/Trochaic learner

NonFinal	114.646
*Clash	106.920
*Lapse	105.893
Trochaic	104.565
Main-L	104.282
WFL	103.250
Parse	103.173
*Cμ	102.931
WSP	98.797
WBP	97.069
AFL	95.949
FtBin	94.859
Iambic	93.425
Main-R	88.742
AFR	87.015
WFR	85.354

The ranking of Main-L >> Main-R, FtNonfin >> Iambic, and Parse >> AFL >> AFR is nevertheless borne out. *Cμ outranks WBP, so that coda consonants are non-moraic.

The effect of this ranking is exemplified in (58). The candidate with disyllabic feet is ruled out by top-ranking NonFinal, rendering the candidate with an extrametrical syllable as the winner:

(58) *ŋalkuninpa* 'eating'

| ŋál.ku.nìn.pa |	NonFin	*Lapse	Main-L	WFL	Parse	*Cμ	*Clash	FtNonf	WSP	WBP	AFL	FtBin	Main-R	Iambic	WFR	AFR
/ (ŋál.ku) (nìn.pa) /	*!									**	**		**	**		**
☞ / (ŋál.ku) (nìn) pa /					*			*		**	**	*	**	*	*	***

It remains an open question why none of the EDCD learners came up with an analysis like this.

6.5 Moraic codas and final syllable extrametricality

10 of the 50 GLA/FtNonfinal learners analysed coda consonants as being moraic and final syllables as extrametrical. This results in degenerate feet even in words with an even number of syllables:

(59) Moraic codas and syllable extrametricality

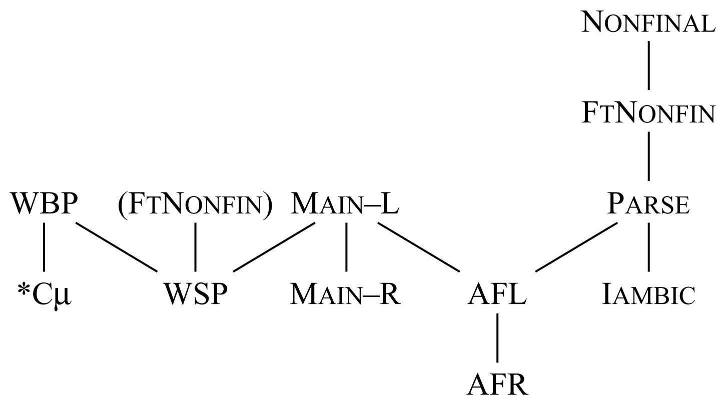
overt forms	surface forms
[cv́v.cvc]	/ (cv́v) cvc$_\mu$ /
[cv́c.cv]	/ (cv́c$_\mu$) cv /
[cv́c.cvc.cvc]	/ (cv́c$_\mu$.cvc$_\mu$) cvc$_\mu$ /
[cv́.cvc.cv̀.cvc]	/ (cv́.cvc$_\mu$) (cv̀) cvc$_\mu$ /

FootNonfinal learners

Final syllable extrametricality is due to the ranking of NonFinal >> FtNonfin. Moraic codas come along with the ranking of WBP >> *Cμ:

(60) A prototypical GLA learner with extrametricality

NonFinal	119.727
*Lapse	113.150
Main-L	112.772
WFL	111.817
Parse	110.369
*Clash	106.862
FtNonfin	104.390
WBP	102.036
*Cμ	97.964
FtBin	97.132
WSP	95.167
Main-R	90.240
AFL	89.522
Iambic	85.247
WFR	80.273
AFR	75.895

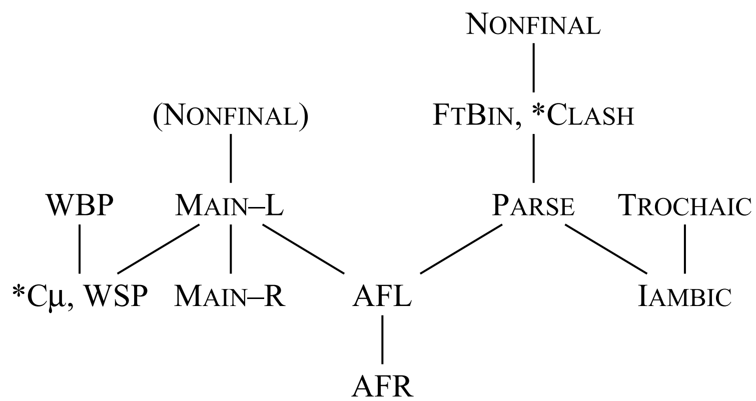


Trochaic learners

While none of the GLA/Trochaic learners came up with an analysis of moraic codas, quite a number of them (17 learners) came up with an analysis that had both moraic codas and extrametrical syllables. High-ranking NonFinal is anticipated, along with a reversal of the ranking between *Cμ and WBP. It turns out that NonFinal is indeed top-ranking and WBP is ranked above *Cμ. Like noted in section 6, Parse can outrank FtBin without having an effect:

(61) The crucial ranking for syllable extrametricality and moraic codas

NonFinal	114.347
Main-L	107.284
*Lapse	107.113
*Clash	106.062
Trochaic	105.539
WFL	104.161
Parse	103.617
WBP	101.062
*Cμ	98.938
FtBin	96.217
AFL	95.611
Main-R	95.611
WSP	94.256
Iambic	92.418
AFR	88.098
WFR	85.653



As shown in (62), NONFINAL is ruling out all candidates that have disyllabic feet in final position. The decision between the candidates with extrametrical syllables is made by WBP, leaving the candidate with moraic codas as the optimal output.

(62) *puɭiŋkalpi* ‘(he fell) finally at the hill’

| pu.ɭiŋ.kal.pi |	NONFIN	MAIN-L	*LAPSE	*CLASH	TROCHAIC	WFL	PARSE	WBP	*Cμ	FTBIN	AFL	MAIN-R	WSP	IAMBIC	AFR	WFR
/ (pú.ɭiŋ$_\mu$) (kàl$_\mu$.pi) /	*!								**		**	**	*	**	**	
/ (pú.ɭiŋ) (kàl$_\mu$.pi) /	*!							*	*		**	**		**	**	
/ (pú.ɭiŋ) (kàl.pi) /	*!							**			**	**		**	**	
☞ / (pú.ɭiŋ$_\mu$) (kàl$_\mu$) pi /							*		**		**	**	*	*	***	*
/ (pú.ɭiŋ) (kàl$_\mu$) pi /							*	*!	*		**	**		*	***	*
/ (pú.ɭiŋ) (kàl) pi /							*	*!*		*	**	**		*	***	*

None of the EDCD learners analysed final syllables as being extrametrical. However, few of them demoted NONFINAL. The few learners that actually did demote NONFINAL to a lower stratum must have had a stage where they would have assigned final stress to some forms in production, resulting in a mismatch with what they must have heard.

Final syllables extrametrical and optionally moraic codas

One GLA/TROCHAIC learner could not decide whether coda consonants should be moraic or not. Both moraic and non-moraic codas could surface in a single form, independently of whether they occurred in a stressed syllable or not.

(63) Undecided weight

overt forms	surface forms
[cv́c.cvc]	/ (cv́c$_\mu$) cvc /
[cv́c.cvc.cvc]	/ (cv́c$_\mu$.cvc) cvc /
[cv́c.cv.cv̀c.cv]	/ (cv́c.cv) (cv̀c$_\mu$) cv /

This draw is due to an exact equal ranking of *Cμ and WBP, something uncommon with the GLA (64). Since the GLA learners all learned with some plasticity noise we can be sure that in this particular case *Cμ and WBP did not move during the learning process.

(64) A GLA/TROCHAIC learner

NONFINAL	112.746
*LAPSE	107.166
MAIN-L	106.835
*CLASH	106.467
TROCHAIC	105.049
PARSE	102.751
WFL	101.931
*Cμ WBP	100.000
AFL	96.361
MAIN-R	95.211
WSP	95.050
FTBIN	94.404
IAMBIC	93.053
AFR	90.939
WFR	87.254

6.6 Interim summary

To sum up shortly, all 200 virtual learners created grammars that describe the data they heard during the training phase. Nevertheless they created different grammars which is reflected in different surface structures. This does no harm, since all these surface structures translate to uniform overt outputs. Let us now have a look at what the learners do when they have to abstract away from the familiar forms, and have to determine the stress pattern of words they have not been trained on.

7 Generalizations to unheard forms

In line with computational linguistic tradition the virtual learners were asked to make generalizations (e.g. Manning & Schütze 1999:577), i.e. predict the stress pattern of words they were not trained on. This provides evidence for to what extent the learners are able to abstract away from the type of forms they heard in the training phase to a grammar accounting for the language. As outlined in 5.2, the learners has been trained on a set of 17 types of di- to quadrisyllabic words. After the training phase, the virtual learners were tested on what stress and foot structure they assign to forms they had not heard before. Among them were all di- to quadrisyllable combinations of cv-, cvc- and cvv-syllables that are allowed in Pintupi, as listed in (26). They were then tested what stress and foot structure they assign to forms that contain long vowels in any syllable within the word, like [cv́.cvv.cvc] or [cv́.cv.cvv] (discussed in 7.1). Furthermore they were tested on what stress they assign to forms with more than four syllables (discussed in 7.2). The five-syllable forms consist of all possible combinations of syllable forms (forms like *puḻiŋkalatʲu*, but also non-attested forms of Pintupi with ling vowels in non-initial position), and the six- and seven-syllable forms consist of cv- strings. 86 learners generalized the stress pattern correctly, meaning that they could transfer the stress pattern they have been trained on to forms with more than four syllables and to forms with long vowels in other syllables than the initial one (these forms are actually not attested in Pintupi, but imagine a real-life scenario where speakers of Pintupi are confronted with loanwords with that kind of syllable structure). All of them were GLA learners. None of the EDCD learners were able to generalize correctly. An overview of the results is given in (65):

(65) Summary of generalizations

Correct generalizations	GLA learners		EDCD learners		Total
	FTNONF	TROCHAIC	FTNONF	TROCHAIC	
a. Linguist's favourite:	14	0	0	0	= 14
b. Moraic codas:	5	0	0	0	= 5
c. Moraic codas in stressed syllables:	0	1	0	0	= 1
d. Extrametricality:	10	29	0	0	= 39
e. Extrametricality & moraic codas:	10	17	0	0	= 27
	= 39	= 47	= 0	= 0	= 86

39 of the GLA/FTNONFIN learners and 47 of the GLA/TROCHAIC learners were successful in transferring the correct pattern to forms the learners were not trained on, that is to unheard forms.

7.1 Generalizations to unattested forms in Pintupi

Some examples are given in (66) for correct generalizations to forms with non-initial long vowels that the GLA learners produced:

(66) Correct generalizations to forms with non-initial long vowels

overt forms	surface structures
[cv́.cvv]	/ (cv́.cvv) / or / (cv́) cvv /
[cv́c.cvv]	/ (cv́c.cvv) / or / (cv́c) cvv /
[cv́.cv.cvv]	/ (cv́.cv) cvv /
[cv́.cvv.cvc]	/ (cv́.cvv) cvc /
[cv́.cvc.cvv]	/ (cv́.cvc) cvv /
[cv́v.cv.cvv]	/ (cv́v.cv) cvv /
[cv́.cvv.cv̀.cvc]	/ (cv́.cvv) (cv̀.cvc) / or / (cv́.cvv) (cv̀) cvc /
[cv́.cvc. cv̀.cvv]	/ (cv́.cvc) (cv̀.cvv) / or / (cv́.cvc) (cv̀) cvv /

The constraint ranking responsible for the weight-insensitive treatment of forms with non-initial long vowels is MAIN-L and FTNONFIN above WSP in the group of FTNONFIN learners and MAIN-L and *CLASH above WSP in the group of the TROCHAIC learners. MAIN-L >> WSP guarantees that the foot with main stress will be aligned to the left edge of the word, so that a heavy syllable cannot attract stress away from the edge. FTNONFIN >> WSP ensures that there won't be a stress clash as in (65i). In the TROCHAIC group, *CLASH takes care of that.

Consider a learner of the GLA/FTNONFIN-group that was able to generalize correctly. Its stress pattern is perfectly weight-insensitive. In (65) we can see an effect of WFL, ruling out a candidate */cv (cv́v.cv) cvc/ with stress on the heavy syllable (remember that long vowels have inherently two moras). Even without this constraint, the same candidate would be ruled out by PARSE. The optimal candidate is (65a), /(cv́.cvv) (cv̀.cvc)/, which has not a moraic coda. Its direct competitor, */(cv́.cvv) (cv̀) cvc/, is ruled out by *Cμ, which is ranked above WBP.

(67) Correct generalization to unattested forms

| cv.cvv.cv.cvc |	FTNONFIN	*LAPSE	MAIN-L	FTBIN	WFL	PARSE	*Cμ	*CLASH	NONFIN	WFR	AFL	WSP	WBP	AFR	MAIN-R	IAMBIC
☞ a. / (cv́.cvv) (cv̀.cvc) /									*		**	*	*	**	**	**
b. / (cv́.cvv) (cv̀) cvc /	*!			*		*				*	**	*	*	***	**	*
c. / (cv́.cvv) (cv̀.cvc$_{\mu}$) /							*!		*		**	**		**	**	**
d. / (cv.cv́v) (cv̀.cvc) /	*!							*	*		**		*	**	**	*
e. / cv (cv́v) (cv̀.cvc) /	*!				*	*		*	*		***		*	**	**	*
f. / cv (cv́v.cv) (cv́c$_{\mu}$) /	*!				*	*	*		*		***			*	*	*
g. / cv (cv́v.cv) cvc /					*!	**				*	*		*	*	*	*
h. / cv (cv́v) (cv̀.cvc$_{\mu}$) /	*!				*	*	*	*	*		***	*		**	**	*
i. / (cv́) (cv̀v) (cv̀.cvc$_{\mu}$) /	*!			*			*	*	*		***			*****	***	*

7.2 Generalizations to longer forms

The learners were also tested whether they could produce the correct stress pattern when asked to produce words with more than four syllables. Some examples for correct generalizations to forms with five to seven syllables are given in (68):

(68) Correct generalizations to forms with more than four syllables

overt forms	surface forms
[cv́.cvv.cv̀v.cv.cvc]	/ (cv́.cvv) (cv̀v.cv) cvc /
[cv́.cvv.cv̀.cvv.cvc]	/ (cv́.cvv) (cv̀.cvv) cvc /
[cv́c.cvv.cv̀v.cvc.cvv]	/ (cv́c.cvv) (cv̀v.cvc) cvv /
[cv́.cv.cv̀.cv.cv̀.cv]	/ (cv́.cv) (cv̀.cv) (cv̀.cv) / or / (cv́.cv) (cv̀.cv) (cv̀) cv /
[cv́.cv.cv̀.cv.cv̀.cv.cv]	/ (cv́.cv) (cv̀.cv) (cv̀.cv) cv /

Consider the grammar of a successful GLA/TROCHAIC learner. The form with six syllables has an extrametrical syllable due to high-ranking NONFINAL, but stress is nonetheless correct: the first syllable has primary stress, the third and fifth syllable have secondary stress. MAIN-L makes sure that stress is aligned with the left word edge, while PARSE makes sure that there are three feet in the word:

(69) Generalization to forms with six syllables

| cv.cv.cv.cv.cv.cv |	NonFin	*Clash	Main-L	*Lapse	Trochaic	WFL	Parse	*Cμ	WSP	WBP	FtBin	AFL	Main-R	Iambic	AFR	WFR
☞ a. / (cv́.cv) (cv̀.cv) (cv̀) cv /							*				*	*** ***	****	**	**** ***	*
b. / (cv́.cv) (cv̀.cv) (cv̀.cv) /	*!											*** ***	****	***	*** ***	
c. / cv (cv́.cv) (cv̀.cv) cv /			*!			*	**					****	***	**	****	*
d. / (cv́.cv) cv (cv̀.cv) cv /							**!					***	****	**	*** **	*
e. / (cv́.cv) cv.cv (cv̀.cv) /	*!			*			**					****	****	**	****	

7.3 Incorrect generalizations

114 learners failed when tested to generalize correctly to unheard forms. Among them were all EDCD learners. Their grammars are not deep enough to transfer the correct stress pattern to unheard forms. Many of the learners that failed to generalize displayed a strong tendency for weight-sensitivity in forms with long vowels in non-initial position and produced forms like */(cv.cv́v).cv.(cv̀v.cvv)/.

(70) A short-sighted EDCD learner

*Clash	100
*Lapse	
AFL	
AFR	
FootBin	
FtNonf	
Main-L	
Nonfinal	
Parse	
WFL	
WFR	
WSP	
*Cμ	99
Iambic	
Main-R	
WBP	98

From the data they were trained on they did not infer the crucial ranking of Main-L, *Clash and FtNonfin above WSP, as the successful learners did. Crucial ties were not helping in this case.

Another point of failure was the alignment of feet. Many of the longer forms contained feet that were not properly stringed together, but left out syllables. This comes about with the equal ranking of AFL and AFR. AFL would have to outrank AFR in order to properly align the feet. The very same learner as displayed in (68) therefore produced */(cv́.cv).cv.(cv̀.cv).cv/ and even an iambic foot for the seven syllable form */(cv.cv́).cv.cv.cv.cv.cv/. The occurrence of the iamb happened because the constraints in the first stratum could not evaluate an optimal candidate, and the decision was left to the lower ranked IAMBIC constraint.

A general reason for failure could be that any of the ingredients in this modelling of stress is deficient. The constraints might be poor description of their function, the learning algorithms could be wrong, the training set could have been too impoverished or OT as a a theory of learning could be inadequate. Since there are quite a few successful learners, the reason for the failure of the other learners suggests that the ingredients are quite sufficient, though.

8 Concluding discussion

Simulations on the learnability of languages in an Optimality Theoretic framework gives linguists the possibility of testing claims made in cross-linguistic research and the study of child language acquisition. Learnability limits the amount of e.g. possible stress systems in a different way than factorial typology does. While factorial typology is the set of all possible rankings of constraints that result in different languages, learnability limits the typology of all possible languages by restricting the range to the constraint rankings that are learnable (see also Boersma 2003).

In the simulations of this paper, all learners acquired the stress pattern in the sense that they produced stress on the correct syllable within a word, i.e. their overt production was the same than the overt forms in the target language. Despite the fact that the overt forms were the same, the learners came up with different analyses, though. More than half of the learners (all in all 103) came up with an analysis similar to the one a linguist would come up with for Pintupi stress. Five learners analysed coda consonants as moraic, but apart from that assigned the same foot structure as a linguist would assign it. Twenty-four learners treated only stressed codas as moraic, nonetheless assigning disyllabic feet from left to right. These three groups of learners can be clustered together as one group in terms of foot structure, resulting in a total of 132 learners that came up not only with the desired stress pattern but also with the desired foot structure. The remaining learners assigned a different foot structure in that they always left final syllables unfooted. Some of them assigned this foot structure in combination with moraic codas. Stress assignment was nevertheless correct and weight-insensitive.

There are several reasons why different analyses were possible. The data that the learners encountered do not give enough evidence as to whether codas in Pintupi are moraic or not, so some learners interpreted codas as being moraic (WBP >> *Cμ), while others did not (*Cμ >> WBP).

Evidence for/against syllable extrametricality in Pintupi is not unambiguous, either. The data in Pintupi only give explicit evidence against footing of final syllables in the case of words with an odd number of syllables, but they do not give explicit evidence for final syllable parsing in words with an even number of syllables. Syllable extrametricality comes about with a ranking of NONFINAL above FTNONFIN and FTBIN.

The possibility of interpreting incoming forms in a different way results from the fact that the learners encountered the same data, but in a different order. Depending on which forms you encounter a lot in the beginning, your perception changes to the extent your grammar changes. This applies even more for the GLA learners, since they learned with a plasiticy decrement, i.e. they took bigger learning steps in the beginning and were slowing down over time.

A further reason for the variation in analyses lies in the different characters of the learners. Of the four learning types the GLA/FTNONFIN learners came up with five different analyses. The GLA/TROCHAIC-group came up with three analyses (a subset of the ones that the GLA/FTNONFIN learners came up with), while the EDCD learners only came up with two analyses (again a subset), irrespectively whether they had FTNONFIN or TROCHAIC implemented. This suggests that the kind of learning strategy rather than the difference in constraints is responsible for the variation.

With respect to the different constraint sets, neither FTNONFIN nor TROCHAIC can be excluded as the constraint on trochaic feet. TROCHAIC learners came up with less different analyses than FTNONFIN learners, but were nonetheless successful.

While all of the learners were able to assign stress correctly to forms of two to four syllables, not all of them were able to transfer this stress pattern correctly to forms that they have not been trained on. Eighty-six of the GLA learners were able to do so. They were able to generalize to longer forms (forms with five to seven syllables) and to forms that are not attested in Pintupi (forms with long vowels in non-initial position). None of the EDCD learners could apply the correct stress pattern to unheard forms. Most of the learners that failed to generalize correctly tended to stress long vowels that occurred anywhere in the word. This weight-sensitivity resulted in stress on the wrong syllable within the word. The incorrect generalizations occurred because the data that the learners have been trained on do not give enough evidence for the fact that Pintupi stress is consequently weight-insensitive, since long vowels only occur in initial position, and are therefore always stressed. The results indicate that the learners that could not transfer the correct stress pattern have not yet truly acquired the Pintupi stress system. Further training with so far unheard forms would probably lead to an improvement of their grammars. A reason why all the EDCD learners failed to transfer the desired stress pattern to unheard forms could be that they did not manage to create a total ranking (Tesar & Smolensky 2000) of constraints. Their grammars had usually only two or three strata, four in rare cases. The crucial rankings required at least four up to five strata.

Another effect became apparent in the simulations of acquisition here. An exhaustive list of constraints does seem to facilitate learning: one constraint not directly applying to the phenomenon can substitute the effect of another constraint crucial for the analysis: an analysis using TROCHAIC needs FTBIN and *CLASH to be higher-ranking so that disyllabic feet can surface. An analysis with FTNONFIN can take over the role of FTBIN: if FTNONFIN is ranked above PARSE feet are naturally disyllabic. Final syllables are unfooted not because they are too small, but because they would violate high-ranking FTNONFIN. This applies only to languages with trochaic feet; in iambic languages, FTBIN might still make a difference.

In sum it can be said that learners of one and the same language may not end up with exactly the same grammar. The virtual learners here were exposed to the same data, only in a different order of presentation, and ended up with five different analyses. Real children acquiring a language will be exposed to data not only differing in order of encounter, but that also differ qualitatively. This suggests that the variation in grammars of real speakers could be even broader.

9 References

Apoussidou, Diana & Paul Boersma (2003). The learnability of Latin stress. *IFA Proceedings* 25, 101-148. ROA 643.

Apoussidou, Diana & Paul Boersma (2004). Comparing two Optimality-Theoretic learning algorithms for Latin stress. *Proceedings of the 23rd Western Coast Conference on Formal Linguistics*.

Boersma, Paul (1997). How we learn variation, optionality, and probability. *IFA-Proceedings* **21**: 43–58.

Boersma, Paul (1998). Functional phonology: Formalizing the interactions between articulatory and perceptual drives. Ph.D. dissertation, University of Amsterdam.

Boersma, Paul (2003). Review of Tesar and Smolensky (2000): *Learnability in Optimality Theory*. Phonology **20**: 436-446.

Boersma, Paul & David Weenink (1992-2005). *Praat: doing phonetics by computer (Version 4.4)*. [Computer program]. Retrieved from www.praat.org.

Bosson, James E. (1964). Modern Mongolian. *Uralic and Altaic Series* 38. Bloomington: Indiana University.

Broselow, Ellen, Su-I Chen & Marie Huffman (1997). Syllable weight: convergence of phonology and phonetics. *Phonology* 14:47-82.

Chomsky, Noam (1986). *Knowledge of language: Its nature, origin, and use*. New York: Prager.

Chomsky, Noam & Morris Halle (1968). *The Sound pattern of English*. New York, Evanston and London: Harper & Row Publishers.

Daelemans, Walter, Steven Gillis & Gert Durieux (1994). The acquisition of stress: A data-oriented approach. *Computational Linguistics* 20:421-451.

Dresher, Elan & Jonathan Kaye (1990). A computational learning model for metrical phonology. *Cognition* 34:137-195.

Elenbaas, Nine & René Kager (1999). Ternary rhythm and the lapse constraint. *Phonology* 16:273-329.

Everett, Daniel (1988). On metrical constituent structure in Pirahã. *Natural Language and Linguistic Theory* 6:207–46.

Everett, Daniel & Keren Everett (1984). Syllable onsets and stress placement in Pirahã. *West Coast Conference on Formal Linguistics* 3:105–16.

Green, Thomas & Michael Kenstowicz (1995). On lapses (The lapse constraint?). Unpublished manuscript, MIT. ROA 101-0000.

Gordon, Matthew (2002). A phonetically-driven account of syllable weight. *Language* 78, 51-80 [also in *UCLA Working Papers in Phonology 2*].

Hammond, Michael (1984). Constraining metrical theory: a modular theory of rhythm and destressing. Ph.D. dissertation, Univeristy of California, Los Angeles. [Published 1988, New York: Garland.]

Hansen, K.C. & L.E. Hansen (1969). Pintupi Phonology. *Oceanic Linguistics*, Vol. VIII, 2:153-170.

Hayes, Bruce (1980). A metrical theory of stress rules. Doctoral dissertation, Linguistics Department, MIT.

Hayes, Bruce (1989). Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20:253-306.

Hayes, Bruce (1991). *Metrical stress theory: principles and case studies*. Ms., UCLA.

Hayes, Bruce (1995). *Metrical stress theory: Principles and case studies*. Chicago: University of Chicago Press.
Kager, René (1992). Shapes of the generalized trochee. *Proceedings of WCCFL* XI.
Kager, René (1994). Ternary rhythm in alignment theory. Unpublished manuscript, University of Utrecht. ROA 35-1094.
Kager, René (1999). *Optimality Theory*. Cambridge University Press.
Kenstowicz, Michael (1995). Base-identity and uniform exponence: alternatives to cyclicity. Unpublished manuscript, MIT. ROA 103-0000.
Liberman, Marc (1975). The intonational system of English. Doctoral dissertation, MIT.
Liberman, Marc & Alan Prince (1977). On stress and linguistic rhythm. *Linguistic Inquiry* 8:249-336.
Manning, Christopher D. & Hinrich Schütze (1999). *Foundations of Statistical Natural Language Processing*. Cambridge, Mass.: MIT Press.
McCarthy, John & Alan Prince (1993). Prosodic Morphology I: constraint interaction and satisfaction. Ms., Linguistics Department, University of Massachusetts, Amherst, and Linguistics Department, Rutgers University, New Brunswick, NJ.
Morén, Bruce (2000). The puzzle of Kashmiri stress: implications for weight theory. *Phonology* 17:365-396.
Myers, S. (1987). Vowel shortening in English. *Natural Language & Linguistic Theory* 5:485-518.
Phillips, J. (1973). Syntax and vocabulary of mother's speech to young children: age and sex comparisons. *Child Development* 44: 182-185.
Prince, Alan (1980). A metrical theory for Estonian quantity. *Linguistic Inquiry* 11:511-562.
Prince, Alan (1983). Relating to the grid. *Linguistic Inquiry* 14:19-100.
Prince, Alan (1990). Quantitative consequences of Rhythmic Organization. *CLS* 26, 2:355-398.
Prince, Alan & Paul Smolensky (1993). Optimality Theory: constraint interaction in generative grammar. Ms., Rutgers University, New Brunswick, NJ, and University of Colorado, Boulder. Technical Report RuCCS-TR-2. Rutgers Center for Cognitive Science.
Sapir, Edward & Morris Swadesh (1960). Yana dictionary. *University of California Publications in Linguistics* 22. Berkeley: University of California Press.
Selkirk, Elisabeth (1984). Phonology and syntax: the relation between sound and structure. Cambridge, Mass.: MIT Press.
Tesar, Bruce (1995). Computational Optimality Theory. Doctoral dissertation, Department of Computer Science, University of Colorado, Boulder. ROA-90.
Tesar, Bruce (1998). Using the mutual inconsistency of structural descriptions to overcome ambiguity in language learning. *Proceedings of the North East Linguistic Society 28*, eds. Pius N. Tamanji and Kiyomi Kusumoto, 469-483. GLSA, University of Massachusetts, Amherst.
Tesar, Bruce & Paul Smolensky (1993). The learning of Optimality Theory: An algorithm and some basic complexity results. ROA-52.
Tesar, Bruce & Paul Smolensky (1998). Learnability in Optimality Theory. *Linguistic Inquiry* 29:229-268.
Tesar, Bruce & Paul Smolensky (2000). *Learnability in Optimality Theory*. Cambridge: MIT Press.

Walker, Rachel (1995). Mongolian stress: Typological implications for nonfinality in unbounded systems. *Phonology at Santa Cruz* 4:85–102.