In J. Beckman (ed.), *Proceedings of the North Eastern Linguistic Society 25*, 13-25.
Amherst, MA: GLSA, University of Massachussets. 1995.

Markedness and the Development of Prosodic Structure*

Katherine Demuth

Brown University

1. Introduction

It has long been noted that children's early words are truncated in form, and that those forms show a certain degree of variability in shape. In this paper I argue that children's early word productions can best be understood in terms of output constraints on surface form. First, I show that children's early grammars allow for the emergence of the unmarked form of syllables (Core Syllables) and prosodic words (Minimal Words), and that these can be thought of as different stages of prosodic development. I then demonstrate how the prosodic development of children's early words can be naturally accounted for in terms of prosodic constraints on output form, where parsing of segmental input into output forms becomes more highly ranked over time. I conclude by showing how a given constraint ranking can provide an account of variability in form at a given point in time, whereas a shift in constraint ranking is needed to account for development over time. In so doing, I demonstrate how the process of language acquisition can be thought of as an optimization problem, and how Optimality Theory (Prince & Smolensky 1993) might be adapted to provide for a theory of both competence *and* performance.

* I thank Tom Bever, Jane Fee, Paula Fikkert, Mark Johnson, Clara Levelt, Mark Liberman, John McCarthy, and audiences at Brown University, CUNY Graduate School, the University of Pennsylvania, and the University of Arizona for comments and discussion.

2. Stages in the Development of Prosodic Structure

Much of the work on early phonological development has focused on issues at the segmental level (e.g. Yeni-Komishian, Kavanagh & Ferguson 1980, Ferguson, Menn, & Stoel-Gammon 1992). Only recently have researchers begun to examine the form of children's early words from the perspective of developing prosodic structure, focusing specifically at the level of syllables and words (Demuth 1994, 1995a; Fee 1994, in press).[1] Demuth (1994) identifies four stages in the development of prosodic structure for English and Dutch, and suggests that similar stages of development may be found in the acquisition of all languages. These Stages are presented in Table 1.

Stage I.	Core Syllables - CV (No vowel length distinctions)
Stage II.	Minimal Words/Binary Feet
a.	Core Syllables - CVCV
b.	Closed Syllables - CVC
c.	Vowel length distinctions - CVV
Stage III.	Prosodic Words - larger than a binary foot
Stage IV.	Prosodic Words - target form

Table I. Stages in the Development of Prosodic Structure

Interestingly, Stage I contains the unmarked form of *syllables* - that is, Core (CV) syllables. Similarly, Stage II is characterized by the unmarked form of *prosodic words* - that is, Minimal Words, composed of binary feet (cf. Broselow 1982, McCarthy & Prince 1990, 1991, 1993b). Stage III is composed of prosodic words that are larger than just a binary foot. Target words are eventually produced at Stage IV. The first three stages are illustrated in following examples from Dutch (Fikkert 1994), and similar stages are found in English (Fee & Ingram 1982, Fee 1994, in press).

(1) Stage I - CV

	Child	Adult Target		
a.	[ka:], [kɑ]	/kla:r/	klaar	J (1;4-1;5)
b.	[da:], [dɑ]	/da:r/	daar	
c.	[ti:], [tɪ]	/dɪt/	dit	

(2) Stage IIa - (C)VCV

a.	['a:pə]	/a:p/	aap	T (1;4-1;6)
b.	['bo:tɔ]	/bo:t/	boot	
c.	['tɛinə]	/trɛin/	trein	N (2;5-2;7)
d.	['tœynɑ]	/toyn/	tuin	

---

1 Waterson (1971, 1987) was one of the first to consider prosodic aspects of early acquisition. Others have taken a 'rhythmic' approach, where it has been proposed that children's early utterances are composed of sequences of Strong-(weak) trochaic feet (Allen & Hawkins 1980, Gerken 1991, 1993, Gerken & McIntosh 1993, Wijnen, Kirkhaar & den Os 1994). Although descriptively adequate in characterizing certain aspects of later, sentence-level rhythmic phenomena, it lacks explanatory adequacy at the earliest, word-level stages of acquisition. A more abstract phonological account is therefore needed for both descriptive adequacy and explanatory coverage at the onset of early words.

	e.	['bo:mɪ], ['bo:mi:]	/bɑ'lɔn/	ballon	C (1;10)
	f.	['ɦɑfə]	/ʃi:'rɑf/	giraf	
	g.	['nɛnɛ]	/ko:'nɛin/	konijn	T (1;5)
	h.	['ʋɑfɑ]	/ʃi:'rɑf/	giraf	
(3)	Stage IIb - CVC				
	a.	[teif], [de:s]	/'de:ze/	deze	J (1;6-1;7)
	b.	[pu:s]	/pu:s/	poes	
	c.	[pa:s]	/pa:rt/	paard	
	d.	[a:p], [ɑp]	/a:p/	aap	
(4)	Stage IIc - VV~VC_{son}				
	a.	[tɛi]	/trɛin/	trein	J (1;10-2)
	b.	[ty:]	/stu:l/	stoel	
	c.	[da:]	/da:r/	daar	
	d.	[bo:]	/bo:m/	boom	
	e.	[mɑm], [mɔm]	/ma:n/	maan	
	f.	[œʊ]	/œyl/	uil	
	g.	[hɑm]	/ha:n/	haan	
	h.	[bɔm], pɔm]	/bo:m/	boom	
(5)	Stages II and III (σσ'σ)				
	a.	['fɔm]	/ˌte:lə'fo:n]	telefoon	R (1;10-2;1)
	b.	['tɪfo:m]			
	c.	[ˌmi:kə'fo:n]	/ˌmi:kro:'fo:n/	microfoon	
	d.	['ho:tɑ]	/'o:li:ˌfɑnt/	olifant	J (2;1-2;4)
	e.	[fɑut]			
	f.	['o:fɑ'fɑn]			

Stage I of prosodic development is a sub-Minimal Word - that is, a CV form only, where vowel length is not distinctive (Fikkert 1994). This is shown in (1) above. Then there is a stage where (C)VCV structures are used, as shown in (2). This is the beginning of the Minimal Word Stage (Stage IIa), where words now have the structure of binary feet. For children leaning languages that have only open syllables, this will be the Minimal Word Stage. However, for children learning languages that have closed syllables, the (C)VCV stage may be realized only briefly as means for producing Minimal Words when coda consonants cannot yet be represented. However, once children can represent coda consonants, Minimal Words take the form of CVC structures, as shown in (3). This is Stage IIb. Shortly afterwards, children begin to control vowel length, producing CVV Minimal Words, as shown in (4). This is Stage IIc. Finally, children's words begin to take the form of prosodic words which are larger than a binary foot, as shown in (5). This is Stage III.

Development in the shape of children's early words is therefore a principled, not random, process. And, although there may be variation in the shape of early words at any one point in time (i.e. both Stage IIa and IIb may be coextensive), that variation is only of a particular sort. In other words, children's early words seem to exhibit an increasing ability to handle more complex prosodic structures over time. Any theory of acquisition must be able to account for these facts. Demuth (1995a) captures this development in terms of the Prosodic Hierarchy (Selkirk 1984, Nespor & Vogel 1986), as illustrated in Table 2.

Stage I.	>	Stage IIa.	>	Stage IIb,c.	>	Stage III.
						Pw
						|
		Ft/Pw		Ft/Pw		Ft
		|		|		|
σ		σ		σ		σ
				|		|
				μ		μ

Table 2. The Development of Prosodic Representations

There are certain problems with this structural account, however, in terms of both descriptive and conceptual adequacy. First, if the prosodic hierarchy is available as part of Universal Grammar, it is not clear why children would initially use only part of that representation. Second, although children's development of prosodic structure has been characterized here in terms of 'stages', these stages are partially overlapping rather than discrete. In other words, at any one point in time, there may be a certain amount of variation in the form that a certain word will take, as shown in (5) above. Thus, any theory of acquisition must be flexible enough to account for this type of variation in output form. This is difficult to achieve with the representational approach, where once a given structure comes 'on line', all subsequent productions should have access to it, and can be expected to take that form. Finally, it is not clear, given the representational approach, what 'triggers' children's eventual access to different levels of prosodic structure. If children hear all of the segments of an input form, and have access to the prosodic hierarchy as part of Universal Grammar, why do they not produce fully formed prosodic words from the onset of speech?[2] Do children not perceive certain (unstressed, unfooted) syllables, and therefore not produce them? Demuth (1995a) argues that this is not the case. Rather, there is ample evidence that children perceive weak syllables even though they don't produce them. I therefore assume that, by the time children begin to speak, lexical representations for the words they hear are segmentally (if not morphologically) wellformed. Do children then have a different lexical representation as their output form, as suggested by Kiparsky & Menn (1977)? Again, I suggest that this is not the case. Rather, I argue that children have one lexical representation that serves both as an input and output form at the segmental level of analysis. The shape of children's early words must then be affected by certain surface constraints on output form, where unmarked structures are the first to emerge.

In the following section I show how a constraint-based approach to early word production provides a natural means of understanding the early stages of prosodic development.

---

2 A parallel problem is found with positing impoverished initial structural representations in syntax, based on the absence of functional categories. This has led to a flurry of recent debate, where theorist fall roughly into two groups along 'limited early structure' (VP only) vs. 'early full structure' (IP (& CP)) lines (cf. Meisel 1992, Lust, Suñer, & Whitman 1994 for various views). Demuth (1995b) provides an optimality theoretic perspective on these issues (see also Grimshaw 1994).

3. A Constraint-based Theory of Prosodic Development

I propose that the shape of children's early words might best be accounted for in terms of wellformedness conditions on linguistic structures, where a given output form is 'optimal' for a certain ranking of linguistic constraints. Specifically, I assume that the initial output states for children learning all languages will be realized in terms of unmarked prosodic structures, and that these will change during the course of development in accord with language-particular realizations of phonotactic and prosodic constraints. I therefore make the following assumptions concerning unmarked prosodic structures and their 'emergence' at the early stages of acquisition:

(6) Acquisition Assumptions

a. Core Syllables (CV) are the unmarked form of syllable structure provided by Universal Grammar (Clements & Keyser 1983, Maddieson 1984, Steriade 1982).

b. Minimal Words (binary feet) are the unmarked form of Prosodic Words provided by Universal Grammar (McCarthy & Prince 1993).

c. Following Jakobson (1941/68), we predict that children will move from unmarked to more marked prosodic structures in the course of acquisition.

d. Stages of acquisition can be characterized by a set of constraints (Prince & Smolensky 1993), where unmarked values emerge first, and where development involves the reranking of constraints over time.

Furthermore, in accord with optimality theoretic assumptions (Prince & Smolensky 1993), I assume that the following also hold:

(7) Optimality Theoretic Assumptions

a. Constraints are all available as part of Universal Grammar

b. A full set of phonological representations (segmental, prosodic) is 'available' as part of Universal Grammar

c. Constraints can be violated (i.e. they are 'soft' constraints)

(7a) states that the set of linguistic constraints that characterizes natural language is available, in principle, to any language learner, as part of Universal Grammar. Thus, though certain constraints may play no active role in a particular language, they might be found at certain stages of language development. In fact, it has often been assumed, as part of the Continuity Hypothesis (e.g. Pinker 1984), that all stages in the learning of a target language will involve only structures available as part of natural language. Furthermore, (7b) states that all phonological representations (e.g. all levels of the Prosodic Hierarchy) are also available as part of Universal Grammar, and should in theory be accessible to early language learners. In other words, children don't have to *learn* the Prosodic Hierarchy, but they will have to determine *if and how different levels of that structure are instantiated* in the language being acquired. Finally, (7c) states that constraints are not 'hard' or absolute, but rather soft - that is, they can be violated, at least to a certain degree. It will be shown that this allows for a certain amount of variability in output form. Accounting for variability in children's early productions, especially at the phonological and morphophological level, has been one of the most serious challenges for theories of language acquisition. In the rest of this section I will show that a constraint-based approach to the problem of early words allows for just the right amount of constrained variability in the surface forms that words can take.

Given the markedness considerations in (6) and the optimality theoretic assumptions in (7), what does an initial grammar actually look like in terms of constraints and their relative rankings, and how do constraints come to be reranked over time? The constraints that are relevant to the realization of prosodic words are given in (8).

(8)

FTBIN	Feet are binary at some level of analysis (σ, μ)
FILL	Syllable positions are filled with segmental material
*COMPLEX	Consonant clusters are not allowed
ALIGN	Align (category 1 with edge 1, category 2 with edge 2)
NO-CODA	Syllables may not have codas
PARSE-SEG	Underlying segments are parsed (into syllable structure)
$\text{ALIGN}_{\text{PrWd}}$	(Lex^0, L/R, PrWd, L/R): Binary foot = unmarked form of PrWd

The most important constraints for the problem at hand are the last three: NO-CODA, where syllables cannot have a coda, thereby yielding the unmarked (CV) form of syllables, PARSE-SEG, where the segments of the target word should be parsed into segments in the output form, and $\text{ALIGN}_{\text{PrWd}}$, where the unmarked form of a prosodic word is realized as a binary foot, or a Minimal Word (cf. McCarthy & Prince 1993). We will see that the other constraints play a role at Stages I and IV, to be discussed at the end of this section.

To demonstrate how a constraint-based approach to acquisition would work, I take as a possible target form the three syllable Dutch word *olifant* /ˈo:li:ˌfɑnt/ 'elephant', and consider the possible surface forms it might take at Stages IIa, IIb, and III. To do this, I consider some actually attested forms from the Dutch corpus - the forms in (5d-f) - repeated in (9). Although these forms coexisted at a certain point in development (discussed in (15) below), each of these forms is also typical of the word shapes found at different stages of development. I therefore take these forms, along with the hypothetical Stage I form [fɑ], and the final target form (Stage IV) as the types of forms that children might output at a given stage of development. In other words, these are potential 'candidate' forms which might be generated by the child's grammar (the GEN function) at a certain point in development. I will then show how some of these candidate forms are disallowed by the constraint ranking at a given point in time, and how others are 'optimal'.

(9)

Stage I	[fɑ]	/ˈo:li:ˌfɑnt/	'elephant'
Stage IIa	[ˈho:tɑ]		
Stage IIb	[fɑut]		
Stage III	[ˈo:fɑˈfɑn]		
Stage IV	[ˈo:li:ˌfɑnt]		

At Stage IIa, early words are maximally and minimally composed of binary feet, and coda consonants have not yet appeared. The NO-CODA constraint and $\text{ALIGN}_{\text{PrWd}}$ constraint must be ranked fairly high. Specifically, they must be ranked more highly than the constraint PARSE-SEG, which is violated several times. At this point the prosodic word is equivalent to a binary foot. That is, the unmarked form of prosodic words has emerged. Note also that the unmarked form of prosodic words can emerge at this point precisely because PARSE-SEG is ranked very low - otherwise this would not be an optimal form.

(10) Stage IIa: NO-CODA, $\text{ALIGN}_{\text{PrWd}}$ . . . . >> . . . . PARSE-SEG

(PRWd = unmarked ≈ binary FT)

Given these constraints, and this particular ranking of constraints, the optimal output will then be CVCV in shape - a binary foot with no codas (a bisyllabic foot). Other possible forms that the child might generate at this point will be less optimal, as shown in (11).[3]

(11) Stage IIa - CVCV Minimal Words

NO-CODA, $\text{ALIGN}_{\text{PrWd}}$ . . . . >> . . . . PARSE-SEG

/ˈo:li:ˌfant/		NO-CODA	$\text{ALIGN}_{\text{PrWd}}$	PARSE-SEG
i.	[fa]			*****
ii. ☞	[ˈho:ta]			***
iii.	[faut]	*!		***
iv.	[ˈo:faˈfan]	*!	*	*
v.	[ˈo:li:ˌfant]	*!	*	

The hypothetical form in (i) is unacceptable as it violates the unmarked (binary foot) form of PRWd. The form in (iii) is unacceptable at Stage IIa because it violates NO-CODA. The forms in (iv) and (v) violate both NO-CODA and the unmarked form of PRWd. Thus, at Stage IIa, the form in (ii) is optimal, even though it violates PARSE-SEG several times.[4] At initial stages of development, PARSE-SEG must be ranked very low.

At Stage IIb, Minimal Words are still being used, but coda consonants now appear. I suggest that the NO-CODA constraint has therefore been demoted, now reranked below PARSE-SEG, or at least ranked lowly enough so that it will no longer exert any effect.

(12) Stage IIb: $\text{ALIGN}_{\text{PrWd}}$ . . . >> . . . PARSE-SEG . . . >> . . . NO-CODA

(PRWd = unmarked ≈ binary FT)

Interestingly, such a ranking of constraints allows for variation in the surface realization of Minimal Words. That is, both CVCV and CV(V)C forms are 'optimal' outputs for the constraint ranking provided in (12) - both are minimal words and violate PARSE-SEG to the same degree. In other words, the ranking of constraints provided in (12) allows for the possibility of free variation in output for a given input form, i.e. both (ii) and (iii) are optimal forms. This is precisely the type of free variation that is frequently attested at various stages of acquisition.

---

3 This presumes that onsets are obligatory, but this and the exact specification of particular segments is irrelevant to the present discussion of prosodic structure. See Rice & Avery (1993) for a discussion of the acquisition of segments from a perspective that is similar to the prosodic approach taken here (although yet to be formalized from an OT perspective).

4 I have calculated the number of PARSE-SEG violations as a function of how many C's and V's occur in the output form - that is, how may *segments* actually surface. Those segments may bear little relation to the actual segmental/featural content of the input form. Although this is also a problem to be accounted for in any theory of acquisition, I take it to be partially independent of the prosodic issues discussed here.