Lexical Selection for Left-Edge Stress in Children[1]

Nancy Hall

University of Essex

1. Lexical selection

It has long been noted that children not only alter the pronunciation of adult words to fit their own grammars, but avoid attempting words whose adult target cannot be produced by their grammars (Schwartz & Leonard 1981). For example, Fikkert (1994:215) reports that in Dutch "at the earliest stage the child has only monosyllabic words, which correspond to monosyllabic adult targets." Thus, the type of word a child chooses to produce indicates the state of her grammar. In this paper I argue that English speaking children go through an un-English-like stage in which the grammar assigns words initial stress, based on corpus evidence that the children attempt long words with initial stress earlier than other types of long words.

This study is based on three children's production of words with 4 or more syllables. Data is taken from the corpus described in Compton & Streeter (1977), which was computerized by Joe Pater, as described in Pater (1997). This corpus includes phonetically transcribed data from Sean, Julia, and Trevor from the ages of about 1 to 3:2. All the children lived in California and were acquiring American English. While the corpus has been previously analyzed for shorter words, 4-syllable words have not to my knowledge been systematically studied before.

2. Child data

In the children's use of 4-syllable words before age 3, there is a bias towards words with initial stress. This bias is revealed in three ways: initially stressed words are the type of long words attempted earliest, the type attempted most often, and the type that forms the greatest part of the child's production lexicon.

[1] Thanks to Joe Pater, Adam Werle, and John Kingston for discussion and help with the corpus work reported here.

Example (1) lists all of the 4-syllable words recorded for all of the children, with the first initials of the children who uttered them. In coding stress, 1 = main stress, 2 = secondary stress, and 0 = no stress. It should be noted that children did not necessarily produce all four syllables of these words; in many cases the words were truncated. The list reflects their choice of 4-syllable *targets*.

(1) all 4-syllable words by stress type (30 4-syllable words)
T = Trevor, J = Julia, S = Sean

1020 (16 words)		2100 (3 words)	
salamander	T	rhinocerous	T, J
everybody	T	harmonica	T, S
helicopter	all	Los Angeles	T
motorcycle	T, S		
elevator	T, S	0100 (4 words)	
cauliflower	J	appreciate	T
alligator	J, S	thermometer	T
caterpillar	J	cooperate	S
pacifier	S	uncomfortable	J
watermelon	S		
babysitter	S	2010 (7 words)	
amplifier	T	application	T
exercises	T	armadillo	J
calculator	S	San Francisco	J, S
operating	S	macaroni	J
decorator	S	decorations	S
		avocados	S
		Cinderella	J

Words with 1020 stress constitute 53% of early 4-syllable lexical items (16 out of 30). Another 23% (7 of 30) have 2010 stress—i.e., still two trochees but with main stress on the second. Altogether, 76% of the 4-syllable words attested at this age range consist of two trochees.

To tell whether children are lexically selecting for a particular stress pattern, it is necessary to compare these percentages to the frequency of each stress pattern in the English language as a whole. I searched the database Pronun, which is coded for stress patterns, for all 4-syllable words. The results of this initial search are not necessarily relevant, however, because many long words of English are ones that children are not likely to hear or be interested in. To narrow the list down to words a child might potentially know, I searched the database Webster's, which codes words for both familiarity (on a 1 to 7 scale) and frequency. The long words the children produce tend to score very high on familiarity but often low on frequency: for example, *hippopotamus* has a familiarity rating of 7 but a frequency of 1 / 1,000,000, reflecting its relative rarity as a topic of adult discourse. Assuming therefore that high-familiarity words are a better representation of a child's potential vocabulary than high-frequency words, I selected all

the words with a familiarity rating of 7. Comparing these two lists yielded 368 4-syllable words with a familiarity rating of 7. Many of these words are still not likely parts of a child's vocabulary (*mortality, generalize)*, but this pared list is at least a better representative.

2.1 Primary stress

Example (2) shows the percentages of high-familiarity 4-syllable English words that have primary stress in each of the possible positions. The second column indicates which of the possible stress patterns in each group are actually attested in the list.

(2) Location of primary stress in high-familiarity 4-syllable English words

initial	(1000, 1010, 1020, 1022, 1002)	20%
antepenultimate	(0100, 0102)	43%
penultimate	(0010, 2010, 2012)	36%
ultimate	(2001)	4%

Interestingly, this distribution is unlike the percentages of each type in the children's vocabularies.

(3) English high-familiarity lexicon vs. children's lexicon by stress type

Stress type	English words (n = 368)	Children's words (n = 30)
1σσσ	20%	56%
σ1σσ	43%	20%
σσ1σ	36%	23%
σσσ1	4%	0%

Antepenultimate stress is the most common pattern in the English lexicon, followed by penultimate and initial stress. But for children, initial stress is more than twice as common as any other kind. The children have more initially stressed lexical items than would be expected based on the percentage of items having these stress patterns in the adult English lexicon.

Token frequency in the children's speech and the order of acquisition of items also indicate that the children prefer initially stressed words. Example (4) shows the total number of utterances of words of each type, including repeated utterances of the same word. More than half of the tokens of 4-syllable words are of words with initial main stress. Julia is an exception to this pattern; she attempts penultimate-stressed words slightly often.

(4) Total instances of 4-syllable words of each stress type

	1σσσ	σ1σσ	σσ1σ	σσσ1
Julia	8	3	12	0
Sean	17	5	4	0
Trevor	35	13	4	0
total	50 (55%)	21 (23%)	20 (22%)	0

All the children attempt initially stressed 4-syllable words before any other kind. Example (5) shows the earliest utterance of each 4-syllable word for each child. Stress is marked if it is not initial. Trevor and Julia both produce only initially stressed words before age 2; Sean before age 2;5. Again, the pattern is weaker with Julia, who produces only one initially stressed word before beginning on other types.

(5) Age of first utterance of each 4-syllable word.

child	word	age	stress pattern, if *not* initial
Trevor	motorcycle	1;5.5	
	helicopter	1;5.30	
	rhinocerous	2;0.8	σ1σσ
	thermometer	2;0.14	σ1σσ
	harmonica	2;0.14	σ1σσ
	salamander	2;0.27	
	application	2;5.26	σσ1σ
	everybody	2;6.15	
	appreciate	2;8.5	σ1σσ
Julia	cauliflower	1;10.14	
	rhinocerous	2;0.14	σ1σσ
	alligator	2;0.27	
	San Francisco	2;2.13	σσ1σ
	macaroni	2;2.16	σσ1σ
	helicopter	2;5.5	
	caterpillar	2;8.25	
	uncomfortable	2;11.22	σ1σσ
Sean	alligator	1;6.9	
	pacifier	1;7.17	
	helicopter	1;8.8	
	motorcycle	1;11.28	
	watermelon	2;1.19	
	harmonica	2;5.12	σ1σσ
	San Francisco	2;6.15	σσ1σ
	decorations	2;9.14	σσ1σ
	babysitter	2;9.18	
	elevator	3;1.8	
	cooperate	3;2.11	σ1σσ

There are three pieces of evidence, then, that children have a lexical selection strategy that favors word with initial main stress: they attempt these words first, they attempt them most often, and they attempt a greater number of lexical items with this stress pattern than would be predicted based on the frequency of these items in the adult language.

2.2 Initial secondary stress

Of the 14 long words in (1) without initial main stress, 10 have initial secondary stress. As with initial primary stress, the number of words with initial secondary stress is disproportionate to the number of such words in the English lexicon.

Using the list of familiar 4-syllable words described above and the Pronun coding for secondary stress, I obtained the following estimate of the percentages of familiar long English words that have initial stress.

(6)

	English	children
Primary stress on first syllable	70	16
Secondary stress on first syllable	132	10
Stressless first syllable	161	4

Stressless first syllables are the most common pattern in 4-syllable words. The children's pattern is the opposite: they prefer primary stress initially, but if not primary then secondary stress. The children's lexical selection strategy appears to favor words with *any* initial stress, whether primary or secondary.

3. Analysis

A possible reason that children go through a stage of preferring initial stress in long words is that their grammar needs to pass through such a setting on the path from its initial state to its adult state. In Optimality Theory (Prince & Smolensky 1993), the grammar is modeled as a set of ranked and violable constraints. Constraints are universal, meaning that all constraints are present in all grammars, although in any one grammar, most constraints are so low ranked that they have no effect. The process of phonological acquisition consists of finding the correct constraint ranking. There are believed to be constraints favoring strictly initial stress, since this pattern is found in many languages. If, during the process of reranking, these constraints are temporarily high-ranked, they can cause the child to briefly have a grammar in which words must have initial stress. I will show that such a stage is very likely to occur.

McCarthy & Prince (1993) propose a set of markedness constraints of the basic form in (7) align phonological and morphological boundaries.

(7) Generalized Alignment

Align (Cat1, Edge1, Cat2, Edge2) $=_{def}$
$\forall$ Cat1 $\exists$ Cat2 such that Edge1 of Cat1 and Edge2 of Cat2 coincide
Where
Cat1, Cat2 $\in$ prosodic categories $\cup$ grammatical categories
Edge1, Edge2 $\in$ {Right, Left}

Among other things, constraints in this family govern the placement of stress. Some constraints that will be relevant include:

(8) Align (foot, R/L, PrWd, R/L)

The right / left edge of every foot is aligned with the right / left edge of a prosodic word.
Henceforth: ALIGN-FT-L, ALIGN-FT-R

(9) Align (PrWd, R/L, σ, R/L)

The right / left edge of every prosodic word is aligned with the right / left edge of a stressed syllable.
Henceforth: INITIAL STRESS, FINAL STRESS

Walker (1996) argues for the existence of the constraints in (9) based on the typology of prominence-based stress systems. Many languages always stress the rightmost or leftmost syllable of a word. Tinrin, in (10), is an example of a language with initial stress.

(10) Tinrin: initial stress (L = light syllable, H = heavy syllable)

a.	ĹL	ŋ́ɖi	'(in the) swamp'
b.	ĹHL	ám^{w}a:ti	'chief'
c.	H́LL	ɔ́:juo	'chair'

Tinrin's edge stress does not seem to derive from foot alignment, because in cases like (10b), neither (Ĺ) nor (ĹH) would be a canonical foot. Rather, Tinrin has INIITAL STRESS as an undominated prosodic constraint. The same constraint exists in English, but is low ranked and hence not very active, so that edge stress is routinely violated. It is, however, assumed to be present in the grammar and hence in the child's initial grammar. Although it will end up low-ranked, it can still influence the course of development.

Other standard constraints that I will use include:

(11) PARSE-σ (Prince & Smolensky 1993)
All syllables are parsed by feet.

MAX
Every segment of the input has a correspondent in the output.

Tesar (1997) argues that a child begins with all markedness constraints ranked above all faithfulness constraints. In the course of acquiring the adult grammar, the child must demote the constraints that are violated in English. In the initial grammar, all the prosodic constraints are equally high ranked, as shown in (12).

(12) Initial state of the grammar

ALIGN (PRWD, L, σ́, L), ALIGN (PRWD, R, σ́, R), >> MAX
ALIGN (FT, L, PRWD, L), ALIGN (FT, R, PRWD, R),
PARSE-σ, etc.

At the early stage shown in (12), undominated markedness constraints demand that all feet be aligned with the right edge of the word and also with the left, and for the first and last syllable of every word to be stressed. All these constraints are satisfied in the earliest child outputs, which are monosyllabic.

(13) Trevor's production of 'cookie' at 0;9.11

/cookie/	ALIGN-FT-L	ALIGN-FT-R	INITIAL STRESS	FINAL STRESS	MAX
→ (ká)					*
(cóokie)				*!	

As children advance beyond monosyllables, they go through a stage when outputs consist of exactly one foot. Since English is trochaic, the shape of these outputs can be σ́σ or σ́. For example, *hippopotamus* might be truncated to *pómus*, and *giraffe* to *wáf*. As Pater (1997) argues, this pattern suggests that PARSE-σ and ALIGN-FT-L or ALIGN-FT-R are ranked above MAX at this stage. Because all feet must be aligned to a word edge, there can only be one foot per word. Since PARSE-σ is undominated, there cannot be stray syllables. It is not clear whether the child is aligning feet with the right or left edge of the word (or both), since any of these strategies would produce the same result.

(14)

/hippopotamus/	ALIGN-FT-R	ALIGN–FT-L	PARSE-σ	INITIAL STRESS	MAX	FINAL STRESS
(pómus)						*
(hìppo)(póta)mus	*!*	*	*		σσσ	*
(mús)					σσσσ!	

To produce 2-syllable words, the child must have already demoted the constraint FINAL STRESS, which would block *pómus* in favor of *mús*. But crucially, there is no reason to demote INITIAL STRESS. Since English feet are trochaic, they always have stress on the first syllable. At the one-foot stage, the child is able to maintain the hypothesis that English might have a strictly initial stress pattern like that of Tinrin in (12). This is not contradicted by the outputs she is able to produce at this point.

The next major step the child must make is to allow more than one foot per word. The grammar in (15a) must change to that in (15b): PARSE-σ must demote below MAX, in order to allow unparsed syllables, and ALIGN-FT-L/R must demote below PARSE-σ, to allow multiple feet. The order of these changes does not concern us here.

(15) a. One-foot stage: PARSE-σ, ALIGN-FT-R/L >> MAX
b. Adult grammar: MAX >> PARSE-σ >> ALIGN-FT-L/R.

For the prosody to become fully adult, INITIAL STRESS must also demote, to a position below the other stress constraints. But at what point does this happen? The child has no need to demote this constraint until the progression out of the one-foot stage has begun. Until words can exceed the σ́σ template, the question of whether stress is always initial is a moot point. I propose that the reranking in (15) is what triggers the demotion of INITIAL STRESS—and hence, that INITIAL STRESS does not demote until after the reranking in (15) is already underway. During this lag time, children take the conservative approach of only accepting words with initial stress as they begin to produce their first long words.

4. Conclusion

Children's preference for initial stress can be explained as an accidental byproduct of the change from the initial grammar to adult prosody. Children are born with the universal constraint INITIAL STRESS high-ranked, like all markedness constraints. During the early stages of grammatical development, other markedness constraints happen to enforce initial stress independently. Children don't need to assume that INITIAL STRESS is violated, so they leave it high ranked. The demotion of INITIAL STRESS will only be triggered by the demotion of other constraints, in the period when outputs begin to exceed a foot. Since reranking does not happen overnight, INITIAL STRESS may continue to exert influence on the grammar for a while even as it gradually demotes, causing children to go through a period where they favor words with initial stress.

In this way the child pattern demonstrates one of the primary assumptions of Optimality Theory: that all constraints exist in all languages. A constraint can be completely inactive in the adult grammar, yet still influence the course of acquisition.

According to this theory, the preference for initial stress is rooted in the grammar, and tied to the fact that English is trochaic. Since English doesn't have σσ́ feet, the child is not forced early on to decide that English allows non-initial stress. In iambic languages, the child's progression is predicted to be different. If a child acquiring an iambic language goes through a stage of producing only σ́ and σσ́ words, then she would have to demote INITIAL STRESS at this point. She would not be expected to go through a stage of preferring initial stress later on, and in fact would likely prefer final stress.

Another possible explanation for why children prefer intial stress, not pursued here, might appeal to the child's perceptual abilities. A number of linguists have

proposed that edge stress makes it easier for the listener to identify the beginning of a new word, and hence to isolate it from the stream of speech (Trubetzkoy 1939:277, Jakobson *et al.* 1952, Martinet 1960:87, Garde 1968:98, Walker 1996). Perhaps children don't usually pick out all the words from speech around them, but successfully isolate a higher percentage of the initially stressed words. They might prefer initially stressed words because these are simply the ones that they hear most often. If this is the case, then the preference for initial stress is expected to be universal. A cross-linguistic study of the acquisition of long words, especially one encompassing iambic languages, would help to decide between the perceptual account and the grammatical account proposed here.

References

Compton, A.J. & M. Streeter. 1977. Child Phonology: Data Collection and Preliminary Analysis. *Papers and Reports on Child Language 7*. Stanford University, Stanford, California.

Fikkert, Paula. 1994. *On the Acquisition of Prosodic Structure*. Dordrecht: ICG Printing.

Garde, Paul. 1968. *L'Accent*. Paris: Presses Universitaries de France.

Jakobson, Roman, Gunnar Fant, & Morris Halle. 1952. *Preliminaries to Speech Analysis*. Cambridge, Mass.: MIT Press.

Martinet, André. 1960. *Elements of general linguistics.* Chicago: University of Chicago Press.

McCarthy, John and Alan Prince. 1993. Generalized alignment. *Yearbook of Morphology*, pp. 79–153.

Pater, Joseph. 1997. Minimal violation and phonological development. *Language Acquisition* 6(3), 201-253.

Prince, Alan and Paul Smolensky 1993/2004. Optimality Theory: Constraint interaction in Generative Grammar. Technical report, Rutgers University and University of Colorado at Boulder, 1993. Rutgers Optimality Archive 537, 2002. Revised version published by Blackwell, 2004.

Schwartz, Richard G. & Laurence B. Leonard. 1981. Do children pick and choose? An examination of phonological selection and avoidance in early lexical acquisition. *Journal of Child Language* 9:319-336.

Tesar 1997. Multi-recursive constraint demotion. ms, Johns Hopkins University

Trubetzkoy, N. 1939. *Principles of Phonology*. Originally published in German. Translated (1969) by C. A. M. Baltaxe. Berkeley: University of California Press.

Walker, Rachel. 1996. Prominence-based stress. ms, University of California, Santa Cruz.

nhall@essex.ac.uk