

The acquisition of compound vs. phrasal stress: the role of prosodic constituents*

IRENE VOGEL

Department of Linguistics, University of Delaware

ERIC RAIMY

Swarthmore College and University of Delaware

(Received 13 December 1999. Revised 23 March 2001)

ABSTRACT

This paper investigates the acquisition of compound vs. phrasal stress (*hót dog* vs. *hot dóg*) in English. This has previously been shown to be acquired quite late, in contrast to recent research showing that infants both perceive and prefer rhythmic patterns in their own language. Subjects (40 children in four groups the averages ages of which are 5;4, 7;2, 9;3 and 11;6 and 10 adults) were shown pairs of pictures representing a compound word and the corresponding phrase. They heard a prerecorded tape with the names of the items, and were asked to indicate which one they heard. In addition to 9 real compounds and corresponding phrases, 9 novel compounds were presented (*rédcup* = invented type of flower vs. *red cúp*). A gradual increase in overall correct scores until age twelve was found along with a significant effect of real vs. novel compounds ($p < 0.001$), and an overwhelming tendency for the younger children to prefer compounds regardless of stress. We conclude that the results are due to the slow development of the ability to use prosodic information to override a strong lexical bias.

INTRODUCTION

With regard to the acquisition of prosody, in particular linguistic rhythm, we find an intriguing discrepancy in the literature. On the one hand, it has been

[*] An earlier version of this paper was presented at the VIIIth International Congress for the Study of Child Language in San Sebastian, Spain. We are grateful to the participants of the Congress for their helpful discussion. In addition, we would like to thank two anonymous referees for their insightful comments. Address for correspondence: Irene Vogel, Department of Linguistics, University of Delaware, 46 East Delaware Ave., Newark DE 19716-2551, USA. e-mail: ivogel@udel.edu

reported that infants not only perceive differences in rhythmic patterns, but that they show preferences for those in use in their own language (Spring & Dale, 1977; Jusczyk & Thompson, 1978; Mehler *et al.*, 1988; Jusczyk *et al.*, 1993; Mandel *et al.*, 1995; Sansavini *et al.*, 1997, etc.). On the other hand, it has been reported by Atkinson-King (1973) that children under the age of twelve years are not able to use rhythmic information to distinguish between compounds and phrases in English, where the only difference is in the stress pattern (e.g. *hót dog* vs. *hot dóg*, respectively). Similar late patterns of development have also been documented for different prosodic structures in Dutch (Vogel & van der Flier, 1979, among others) and British English (Cruttenden, 1974).

Some of the discrepancy is immediately accounted for by the fact that the infant studies only test the physiological ability to discriminate between different rhythmic patterns, while the other studies investigate the contrastive use of prosodic patterns. The contrastive use of prosodic patterns cannot develop until the requisite cognitive maturation and knowledge of grammar have been obtained.

In contrast with the perception facts, children generally seem to produce words with primary stress in the correct location as soon as they begin speaking, even if they omit some of the syllables of a word (e.g. *nána* for *banána*). Furthermore, it seems that children are able to produce novel compounds with the appropriate compound stress pattern by age two (Clark, Gelman & Lane, 1985). It also appears that children do not consistently make errors in phrasal stress, though systematic research would be needed to verify this. In fact, the rare inaccuracies in stress placement observed in young children speaking English, Spanish and other languages appear to be due to factors such as focusing on difficult sounds in a normally unstressed syllable, rather than the inability to perceive or produce the appropriate stress patterns (cf. Hochberg, 1988 and references therein).

This paper addresses the question of why it takes children so long to fully acquire the prosodic patterns of their language, in particular, the rhythmic differences between compound stress and phrasal stress in English. First we present some background on the relevant stress patterns of English and previous perception studies of the contrastive use of prosody. We then discuss the experiment we conducted modeled on that of Atkinson-King, though with some substantive modifications. While our findings replicate Atkinson-King's overall late developmental pattern, the modifications that we have made in the experimental design provide further insight into uninvestigated trends present in Atkinson-King's data. On the basis of our findings, we develop a model for the acquisition of the stress patterns in question. We propose that the reason the ability to distinguish between the stress patterns in question takes so long to acquire, in contrast with lexical stress, is that it requires the acquisition of higher level phonological

constituents. While the acquisition of lexical stress only requires the notion of Phonological Word, as does the acquisition of most segmental phonology, we propose that the constituents that must be acquired for Compound and Phrasal stress are the Clitic Group and the Phonological Phrase, respectively.

THEORETICAL AND EXPERIMENTAL BACKGROUND

Stress and prosodic constituents

In order to fully understand the acquisition of stress it is necessary to clearly identify the object of study. We suggest that one of the reasons for conflicting views on whether stress is acquired early or late is that stress is viewed as a single phenomenon, while in fact, there are several types of stress that must be distinguished. First, as mentioned above, we must distinguish between the purely acoustic aspects of stress and the contrastive use of stress. From the acoustic perspective, stress consists of particular patterns of pitch, duration and loudness, but sensitivity to such patterns does not necessarily mean that one associates a specific meaning with a given acoustic pattern. On the other hand, stress may be used contrastively in language to distinguish meanings, as phonemes do at the segmental level. Acquisition of this contrastive use of stress requires precisely that one associate a specific meaning with a given acoustic (stress) pattern.

Limiting our attention here to contrastive uses of stress, we must next distinguish among different types of meaning conveyed by stress. Restricting ourselves to neutral stress patterns (i.e. those without special emphasis or expressive value), we can identify two types of meaning that can be conveyed by stress in English: a) differences between individual words and b) differences between compound words and phrases. The former accounts for distinctions such as *cónvict* (N) vs. *convíct* (V). The latter accounts for distinctions such as *hót dog* (type of food) vs. *hot dóg* (hot canine). These stress patterns are represented graphically in terms of the metrical grids in (1) and (2) below. The grid marks (x) at the lowest level in the grids demarcate each syllable. At the next higher grid level, a grid mark appears only on those syllables that bear main word stress. In (2), we show additional levels corresponding to the compound and the phrase. In each case, the grid mark at the indicated level represents additional prominence on the corresponding syllable.

(1) Word Stress

word	x			x
syllable	x	x	x	x
	con	vict (N)	con	vict (V)

(2) Compound Stress vs. Phrasal Stress

compound	×		phrase		×
word	×	×	word	×	×
syllable	×	×	syllable	×	×
	hot	dog (compound)		hot	dog (phrase)

It should be noted that only Word Stress correlates with vowel quality differences; Compound and Phrasal Stress leave the vowel quality determined at the word level essentially unchanged.¹

Each of the three types of stress under consideration here is assigned by different principles, and we, therefore, assume that each is associated with a different constituent in the prosodic hierarchy as shown in (3) (cf. Nespor & Vogel, 1986).

- (3) a. Word Stress: Phonological Word (PW)
 b. Compound Stress: Clitic Group (CG)
 c. Phrasal Stress: Phonological Phrase (PPh)

Word Stress in English is either listed in the lexicon or assigned by a fairly complex set of rules along the lines of those in the *Sound Pattern of English* (Chomsky & Halle, 1968). We assume that the domain of this type of stress is the Phonological Word, the constituent that includes the root of a word and any affixes that participate in stress assignment, that is, those traditionally identified as +boundary or level 1 affixes. Given the complexity of English word stress, it is most likely acquired as part of each lexical item, though eventually generalizations may be extracted and represented as rules.

By contrast, both Compound Stress and Phrasal Stress follow very simple patterns in English. In fact, Phrasal Stress appears to be determined on the basis of universal principles. As noted in Nespor & Vogel (1986), prominence is assigned at the right edge of the PPh in right branching, or SVO, languages such as English and Italian, while it is assigned at the left edge in left branching, or SOV, languages such as Japanese and Turkish. Since phrasal stress correlates with another, independently defined, parameter of language, it essentially ‘comes for free’ in language acquisition; once a child acquires the basic word order of his/her language, s/he also knows how to assign phrasal stress.²

[1] In fact, there are often additional allophonic differences associated with differences in the placement of primary word stress. For example, in comparing *átom* and *atómic*, in addition to the stress difference, we find: a full vs. a reduced vowel in the first and second syllables, and a flap vs. an aspirated alveolar stop in the second syllable.

[2] An alternative view is that the difference in stress patterns is precisely what helps babies identify the basic word order of their language and thus is a means of ‘bootstrapping’ (cf. Christophe *et al.*, 1997). It is not clear, however, what type of empirical evidence could distinguish between the two views, so at this point, the choice depends on more general theoretical preferences. What is relevant here is the existence of a relationship between word order and phrasal stress, regardless of its direction.

Compound stress is assigned differently from both word level stress and phrase level stress. It is rule governed, but it is not predictable from other characteristics of a language. In English, stress is assigned to the first member of a compound (e.g. *dish washer*), but this is not true in all SVO languages.³ For example, in Italian, another SVO language, it is assigned to the last member of a compound (e.g. *lava piatti* 'dish washer'). We are assuming that the domain that includes the two (or more) members of a compound is the Clitic Group since stress cannot be assigned to compounds at either the PW or the PPh level given that there are different stress assignment rules operating at each of these levels. Thus, from a theoretical perspective, what must be learned in order for a speaker to distinguish between compound and phrasal stress is the identity of the Clitic Group and the principles that assign stress at this level in a particular language.

Acquisition of contrastive stress

The acquisition of word stress in English has received little attention since it seems that children do not regularly make mistakes either in its production or perception. That is, there is no analogue to common segmental substitutions such as the use of [d] for [ð] or [t] for [k].

Nor is there evidence that pairs such as *récord/recórd* are misunderstood because of their stress pattern. Furthermore, while children between one and five years of age produce significant numbers of slips of the tongue involving individual segments, indicating the existence and manipulability of individual segments, slips of the tongue involving stress or intonation are extremely rare in the same children's speech. At least in a corpus collected by Jaeger (1999, to appear) they do not appear at all until age four, and then account for no more than 1% of all the slips of the tongue. Indeed, in an analogous adult corpus, such slips of the tongue are rare as well, only accounting for 2% of all slips (Jaeger, 1999, to appear). Thus, it seems that at least at the word level, English-speaking children acquire the necessary stress patterns as they master their vocabulary. By the same token, it appears that adult English speakers store the stress pattern of a word as part of its lexical representation.

At first glance, it would appear to be the same for compound and phrasal stress. As mentioned above, it has been reported that even before two years of age children productively create compounds and appropriately stress the first word (Clark *et al.*, 1985). The paucity of research on the acquisition of phrasal stress might be taken as an indication that children do not seem to make substantial errors in this realm either. A study by Wieman (1976), however, indicates that the acquisition of both compound and phrasal stress

[3] Actually, stress also occasionally occurs on the second word of compounds in English (e.g. *Italian American*), but stress on the first syllable is by far the more widespread pattern.

may be a bit less straightforward than appears at first glance. In two word utterances, it was found that ‘the semantic relations being expressed by the child are more important to his stress assignment than are any adult-type syntactic category labels ...’ (p. 284). In fact, there appears to be a stress hierarchy, whereby locatives and possessives are the most likely items to be stressed (e.g. *play muséum*, *rábbit house*), followed in decreasing order by objective, attributive, verbal and agentive elements. New information also was found to be stressed, and thus could cause the general pattern to be overridden (e.g. *sée marble*, when a missing marble was surprisingly found) (p. 285). In most compounds in English, including those discussed in Clark *et al.*, the new information is provided by the first member. Thus, the fact that compounds have been found to be produced correctly by two-year-olds may actually be a reflection of the tendency to stress new information, rather than an indication that the children have mastered the stress pattern of compounds *per se*.

In any case, Atkinson-King’s (1973) study seems to be in conflict with claims that stress is mastered very early, finding that children are not able to use a difference in stress patterns to perceptually distinguish between compounds and phrases until age twelve. That is, while children may be able to produce compounds with initial stress (*hót dog*) and phrases with final stress (*hot dóg*) by age two, when presented with these different stress patterns, it takes many more years before they can reliably identify the corresponding compound/phrasal meaning. Given pictures of a ‘food item’ and a ‘hot canine’, as in Figure 1, Atkinson-King reported that four-year-

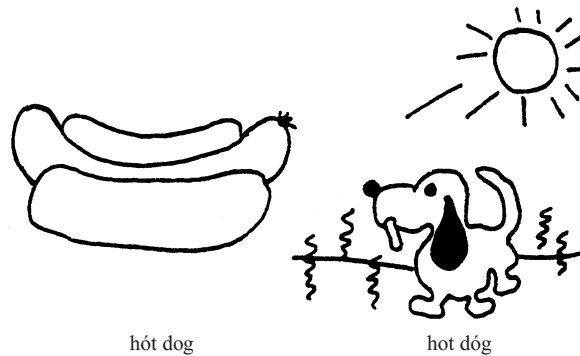


Fig. 1. *hót dog* (compound) vs. *hot dóg* (phrase).

olds randomly chose either picture, whether they heard a stimulus with stress on the first word (*hót dog*) or a stimulus with stress on the final word (*hot dóg*). Beyond that, Atkinson-King found a gradual increase in the children’s scores with age, with subject groups consisting of children aged four, six, eight, ten

and twelve. Only the twelve-year-olds scored close to 100%, as did the adult subjects.

While it may be surprising that the ability to distinguish between compound and phrasal stress is acquired so late, this pattern is consistent with two other studies investigating children's ability to use prosodic information to disambiguate different types of structures.⁴ It is reported that in Dutch another type of contrast that is signalled only prosodically also takes until age twelve to be acquired (Vogel *et al.*, 1979; Vogel & van der Flier, 1979; Bol & van der Flier, 1980). In the Dutch study, sixty-seven children between four and twelve years and eight adults were asked to distinguish between the two meanings of sentences such as those in (4).

- (4) Ambiguous Dutch sentence
- a. Wie 'helpt de man? Whom is the man helping?
Who helps the man (Wie = object; de man = subject)
 - b. Wie helpt de 'man? Who is helping the man?
Who helps the man (Wie = subject; de man = object)

The two meanings of such sentences are distinguished by the position of the main sentence accent, indicated by the accent marks in (4). When the accent falls on the verb (*helpt* 'helps' in this case), we get the meaning in (4a), in which *wie* 'whom' is the object; when the accent falls on the noun (*man* 'man'), we get the meaning in (4b), in which *wie* 'who' is the subject.

As in the study of the two stress patterns in English, the Dutch children were asked to indicate which of two pictures corresponded to the stimulus they heard. For example, in the case of the sentences in (4), they saw a picture of a man helping a woman up from the ground for (4a), and a picture of a woman helping a man up from the ground for (4b). In this study, it seemed that the youngest subjects responded without regard for the stress pattern. Various response patterns were observed in subsequent age groups until, finally around age twelve, the subjects performed at close to 100%, as did the adults.

Similarly, Cruttenden (1974) reported that it was only after age ten that British children could listen to the first portion of a football match report on the news and determine which team had won on the basis of the intonation pattern – before the scores were actually announced. Depending on the prosody of the material preceding the second score as in *Everton 2 Liverpool ...*, adults can predict with complete accuracy whether Liverpool beat

[4] We are grateful to an anonymous referee who pointed out that there appears to be a related discrepancy in the development of another prosodic phenomenon, tone. While lexical tone appears to be acquired early, along with individual lexical items, more complex tonal rules, including Tone Sandhi, appear to be mastered a number of years later (cf. Tse, 1978; Clumeck, 1980, and Erbaugh, 1992 for Chinese, Tuaycharoen, 1977 for Thai, Demuth, 1984, 1989, 1991, 1992 for Sesotho).

Everton, lost to Everton or tied with Everton. When twenty-eight boys between seven and ten years of age were asked to predict the football scores, however, they were unable to do so. '[A]t age 9 at least 50 percent of children show nothing approaching complete competence' (Cruttenden, 1974:229).

Such results, combined with the early ability of children to produce the compound vs. phrasal stress patterns under investigation here raise intriguing questions as to why there should be such a large discrepancy between the production and perception of prosody, and why the perceptual ability appears to develop so late.⁵ To investigate these questions, we conducted the experiment described in the next section. This experiment, based on that of Atkinson-King, was designed both to determine whether we would get similar results, and to gain further insight into the development of prosody, in particular, linguistic stress.

METHOD

Experimental design

Following Atkinson-King, our study consisted of minimal pairs differing only in their stress pattern: compound vs. phrasal. Subjects heard one or the

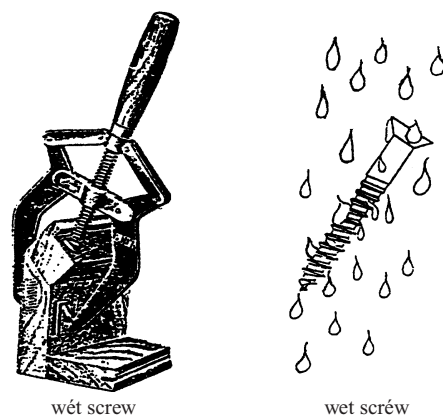


Fig. 2. *wét screw* (compound) vs. *wet scrév* (phrase).

other item pronounced and had to indicate which of two pictures corresponded to what they heard. While Atkinson-King used only real compounds and their corresponding phrases, we introduced some novel compounds and their

[5] Further contributing to the intriguing nature of the issue is Ashby's (1992) anecdotal report that his eight-year-old child had metalinguistic knowledge of the difference in stress between compounds and phrases. In addition, the article reports a discussion between this child and his six-year-old brother in which both children produced both compound and phrasal stress patterns on the term *blue bottle*, a type of fly.

corresponding phrases (e.g. *wét screw* vs. *wet scréw*), as illustrated in Figure 2.

This innovation was made since it seemed that a number of the ‘real’ compounds were unknown to the younger children. Given that we might expect unknown items to behave differently from known items, by creating novel compounds we were able to provide stimuli for the older subjects that were parallel to the unfamiliar compounds for the youngest children.

In developing the novel items, we strove to construct images of the would-be compounds that reflected the general properties of typical noun compounds in English. Given that the semantics of English compounds is often unpredictable (e.g. *greenhouse*, *hotdog*), we aimed for items that bore some relation to their invented names, without actually being fully transparent. In so doing, we were furthermore trying to isolate the subjects’ use of the stress pattern. Within the context of our forced choice paradigm, we expected that if subjects recognized the compound stress pattern they would choose a relatively plausible, though unknown, item to go with it, rather than the picture of the phrasal meaning which would involve a mismatch in the stress pattern.

Subjects

The subjects for our study were forty children attending an elementary school in Delaware, grouped by age as shown in Table 1 below. Subsequent

TABLE 1. *Subject groups information*

Group	No. of subjects	Age range	Age mean
Pre-kindergarten	10	4;9–5;9	5;4
1st grade	10	6;11–7;7	7;2
3rd grade	10	8;11–9;6	9;3
5th grade	10	11;2–12;3	11;6
Adult	10	N/A	N/A

Testing occurred in the spring of the school year, thus subjects appear to be older than would be expected for their grade level.

reference to each group will be made using only the years portion of the average age of the children in that group; for example, ‘five-year-olds’ refers to the youngest group, with the average age 5; 4. We also ran the experiment on a control group of ten adults who work at the same school or who are parents of children attending the school. The adults are all over twenty, but more exact age information is not included here due to the reticence on the part of some of the adults to report their age. All the subjects are monolingual speakers of American English, with no history of hearing or language problems.

Stimulus materials

The stimuli consisted of eighteen pairs of items in which one member of the pair had the compound stress pattern and the other had the phrasal stress pattern. In nine of the pairs, the compound was a real item (e.g. *red head*), while in the other nine, the compound was a novel item (e.g. *red cup* = a type of flower). (See Appendix 1 for a complete list of the stimuli.)

The stimuli were recorded by a female speaker, the first author, in two neutral sentential contexts, as illustrated below. Note that the word 'sticky' in (5b) refers to 'Post-It' notes made by 3M, that are sticky along one side.

- (5) The *hót dog* is over here.
 The *hót dog* gets the sticky.

In preparing the stimulus tape, the speaker recorded all the stimuli three times each, in a soundproof room. The second author then independently listened to the items and selected the best exemplar of each target for inclusion in the experiment. Two additional adult native speakers of English listened to the final version of the stimulus tape to ensure that all the items sounded natural.

Acoustic analysis shows that there was no significant difference between the real and novel stimulus items on any measure. By contrast, there were highly significant differences between the stress patterns of the compound and phrasal items. Comparison of the pitch (F_0) at the beginning and the end of the rhyme of the stressed syllable in the first and second words (W1 and

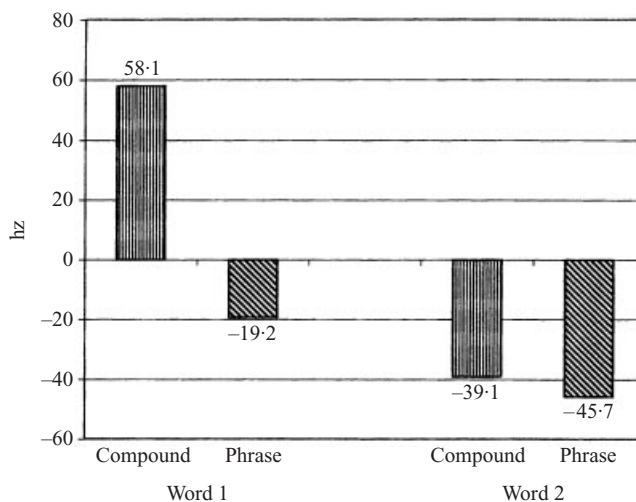


Fig. 3. Change in pitch (hz) from beginning to end of word for word 1 and word 2.

W₂) of our stimuli (both real and novel) showed the following patterns. In compounds, the pitch rises substantially on W₁ and drops moderately on W₂, while in phrases, the pitch drops slightly on W₁ and drops moderately on W₂, as shown in Figure 3. The drop in pitch on W₂ in the compounds vs. the phrases is not significant, indicating that as far as the pitch pattern is concerned, the burden of the difference between the two patterns is borne by W₁ (cf. Appendix 2 for details).

Examination of the other reliable acoustic measure of stress in English, duration, also revealed significant differences between the compound and phrasal patterns. While the duration of W₁ in compounds and phrases is essentially the same, the duration of W₂ shows a highly significant difference ($p < 0.001$) in the two cases: in phrases, the duration is an average of 61 msec longer than in compounds, as shown in Figure 4. Thus, as far as duration is

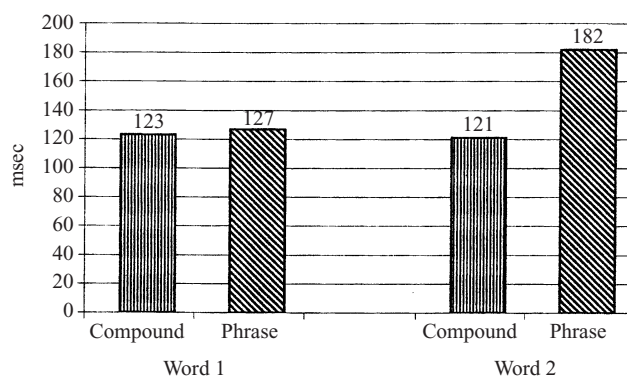


Fig. 4. Duration in msec of word 1 and word 2.

concerned, the burden of the difference between the two patterns is borne by W₂ (cf. Appendix 3 for details).

Corresponding to the audio presentation of each stimulus was a pair of pictures – one representing the compound meaning and the other representing the corresponding phrasal meaning (e.g. ‘item of food’ vs. ‘hot canine’ for *hot dog*), as seen above in Figure 1. The stimuli were presented to each subject in one of four random orders. Each subject heard only one instance of a stimulus pair (e.g. either *hót dog* or *hot dóg*), for a total of eighteen items per subject.

Procedure

The subjects were interviewed individually. They were told only that they would hear some words on a tape recorder, and they had to put a ‘sticky’ (Post-It) on the picture that showed what they heard. The actual experiment was preceded by several practice items to ensure that the subjects understood

the task; the scores on the practice items were not, however, included in our analysis. The practice items consisted of two of the 'real' stimuli and two of the 'novel' stimuli. The subjects heard both versions of the practice stimuli, for a total of eight practice items. We did not tell the subjects what to listen for, nor did we explicitly train them on the test items. In this last respect, our experiment differs from that of Atkinson-King (1973), in which all the test items were identified and named contrastively for each subject as part of the testing procedure. In fact, during the introduction phase of Atkinson-King's experiment, several repeated and maximally contrastive pronunciations of each member of a pair of stimuli were presented (Atkinson-King, 1973:24-25).

RESULTS

Overall, our results agree with those of Atkinson-King (1973) in that our subjects, too, achieve the adult level of behaviour on the distinction between compounds and phrases around age twelve (average age of our subjects 11;6). Furthermore, we observed a gradual development of this behaviour over the preceding years. In a finer grained analysis, however, we found that the development of stress is not simply a matter of a gradual improvement of overall correct scores. Most significantly, we found a crucial effect for whether a compound is real or novel.

First let us consider Figure 5, which shows the overall percentage of correct scores across the different age groups.

Despite the gradual increase in performance with age, a Scheffé *post hoc* test (Table 2) reveals that, on the basis of overall correct answers, the subjects fell into only two distinct groups. The lines in bold face are those showing significant differences.

Specifically, Table 2 shows is that there is not a significant difference in overall correct responses among the three youngest groups of children, ages five, seven and nine. Neither is there a significant difference in overall correct responses between the oldest children (age eleven) and the adults. By contrast, the groups of children aged five, seven and nine are each significantly different from both the groups of eleven-year-olds and adults.

Such overall correct scores, however, can reveal only the most general developmental trend. Let us now consider the results of a three-way ANOVA: group (5) × compound type (2) × stress type (2) with correct answers as the dependent variable. As Table 3 shows, we found significant effects for each of these factors.

These results reveal, most notably, a highly significant effect for each of the individual factors: group, compound type (real vs. novel) and stress type (compound vs. phrasal). Similarly, each of the interactions involving both compound type and stress type yielded significant results.

ACQUISITION OF COMPOUND VS. PHRASAL STRESS

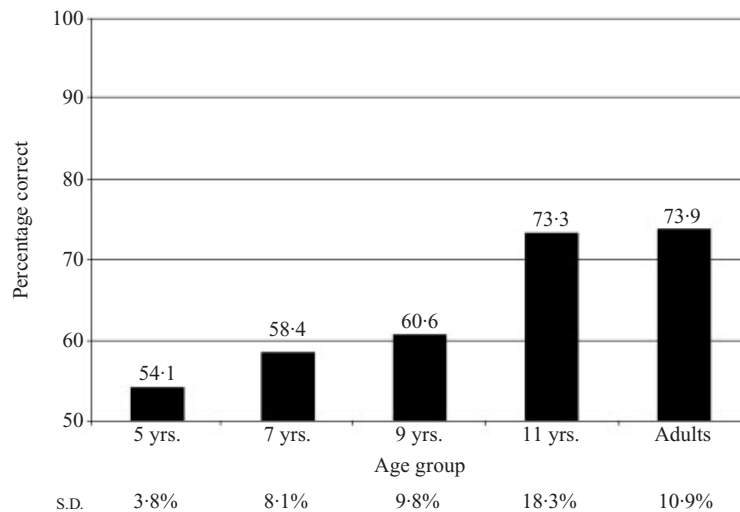


Fig. 5. Mean performance by group.

TABLE 2. *Scheffé post hoc test: subject groupings for overall percentage correct*

Groups	Probability level
5 yrs-7 yrs	0.965731
5 yrs-9 yrs	0.966846
5 yrs-11 yrs	0.001089
5 yrs-Adults	0.001004
7 yrs-9 yrs	1.000000
7 yrs-11 yrs	0.014567
7 yrs-Adults	0.013648
9 yrs-11 yrs	0.014905
9 yrs-Adults	0.013969
11 yrs-Adults	1.000000

TABLE 3. *Three-way ANOVA: group (5) × compound type (2) × stress type (2)*

Independent variable	p value
group	≤ 0.0001
compound type	0.0012
stress type	≤ 0.0001
group × compound type	0.3378
group × stress type	0.5753
compound type × stress type	≤ 0.0001
group × compound type × stress type	0.0465

The highly significant effect for real vs. novel compounds revealed in Table 3 requires that we examine these groups of stimuli separately. This same result also leads to the conclusion that the appropriate comparison of data between the younger and the older subjects involves only those compounds that were known to all subjects. To determine (albeit indirectly) which items were known to the subjects, we gave the parents along with the standard permission slip, a list of words including, among many others, the target compounds, some novel compounds and other compounds. The parents were asked to indicate which words they believed their child knew, but not to actually check this with the child, as we did not want the children to be exposed to any of the words in advance. As a result of these parental questionnaires and our own observations, we determined that several real compounds needed to be excluded since they were unknown to many of the five-year-old subjects. Essentially, all the parents of the youngest subjects reported that their children did not know *redhead* and *greenhouse*. In addition, we found that many were unfamiliar with the word *blackboard*, using the term *chalkboard* instead. That is, even though such items are real compounds for the older subjects, they are essentially novel compounds for the youngest subjects. The parents of the seven-year-olds all reported that their children knew all of the real compounds.

While we cannot provide a direct correlation between a child's knowledge of a particular compound and his/her behaviour since we only have indirect information about the subjects' vocabulary, a general pattern can nonetheless

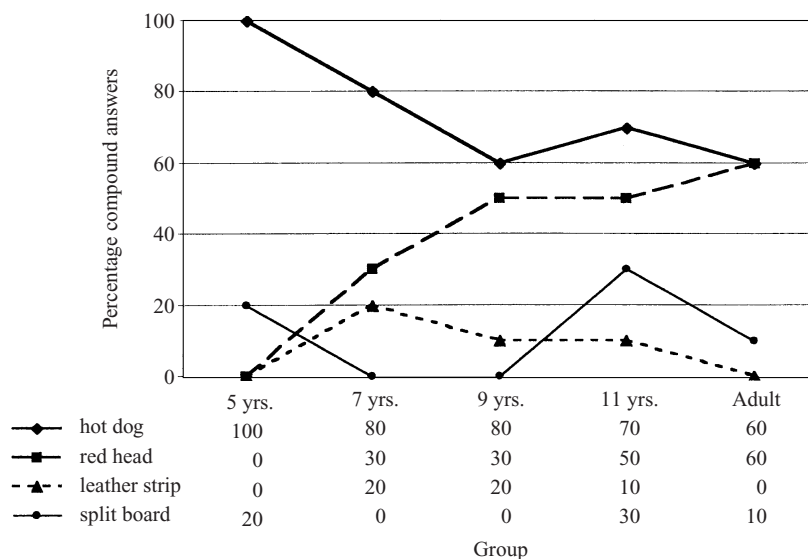


Fig. 6. Percentage compound answers: known, unknown and novel.

be observed. Indeed, if we examine the youngest subjects' performance on *redhead*, an unknown though real compound, we see that the treatment of this item is the same as that of novel items, as distinct from the treatment of known items such as *hotdog*. This is shown in Figure 6, which tracks the percentage of compound answers as opposed to phrasal answers for four prototypical items: *redhead* and *hotdog*, as well as two novel compounds, *leather strip* and *split board*, for comparison. A perfect score here is 50 % since the percentage of compound choice shown refers to the full set of stimuli, only half of which were presented with the compound stress pattern. By around age seven, *redhead* approaches the 50 % mark as does *hotdog*, while the novel compounds never approach this level. Thus, while the real but unknown compound *redhead* started out like a novel item, with a score of 0 %, it eventually diverged from the novel items, which continued to exhibit very low percentages of compound choice.

We thus now restrict our attention to the real compounds known by all subjects (henceforth 'known items') that appeared among our set of target

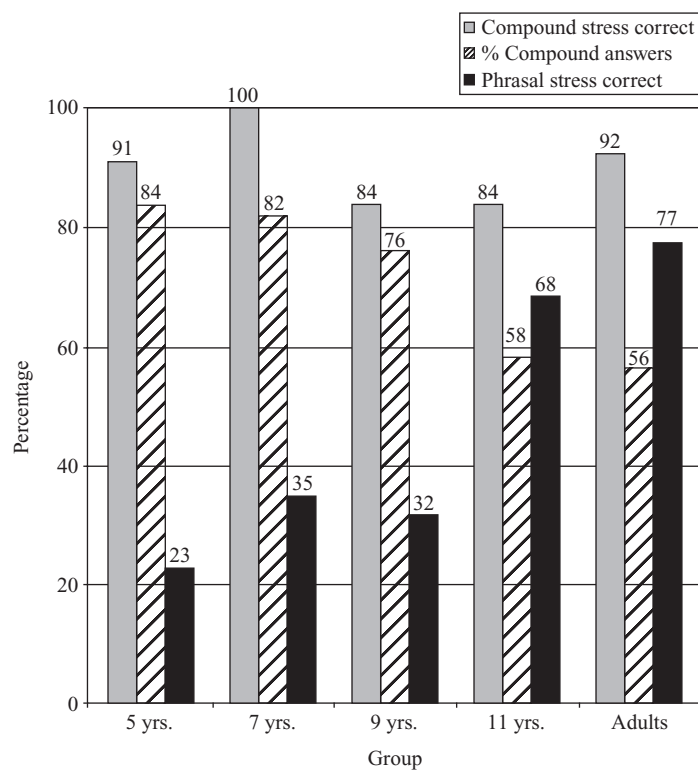


Fig. 7. Compound answers for known items: compound, combined stress and phrasal.

stimuli. Figure 7 presents the percentage of compound and phrasal answers given for the known items only. For each age group, we first show the percentage of compound responses on the items presented with compound stress only. A perfect score here would be 100%. The second column shows the compound responses on the full set of items, presented with both compound and phrasal stress. A perfect score here would be 50% since it is expected that those items presented with phrasal stress should NOT get a compound response. Finally, the third column shows the phrasal responses for those items presented with phrasal stress. The perfect score here would be 100%, as in the first column.

Our findings show that the adult subjects scored very close to the ideal results, and the eleven-year-olds were not far behind. The younger subjects, however, performed substantially less well.

DISCUSSION

Despite the similarity in the overall developmental trend between Atkinson-King's and our study, there is a noteworthy difference. Indeed, we believe that the analysis of this difference is what provides the crucial insight into the acquisition of compound vs. phrasal stress. While Atkinson-King reported that her oldest children (around age twelve) and the adults performed at nearly 100%, the corresponding groups of children and adults in our study only achieved about 74% correct responses. The fact that Atkinson-King explicitly trained her subjects on the experimental stimuli before the experiment, while we did not, most likely accounts to some extent for the better scores on her test. More important, we suggest, is the fact that our experiment was inherently harder due to our inclusion of novel compounds.

Although we cannot investigate the effect of the difference in training in the two experiments, we are able to evaluate the effect of novel compounds introduced in our experiment. As was seen in Table 3, we found a highly significant difference in performance between the real and novel compounds. Since we combined both types in our tabulation of the percentage of correct responses in Table 1, we can conclude that our lower overall scores are indeed due, at least partially, to the inclusion of the novel compounds, on which performance was substantially poorer.

Furthermore, given the strong effect of real vs. novel compound type we observed across the board, we must now cast some doubt on the conclusions of Atkinson-King's study, most notably where the youngest children are concerned. Since it is likely that not all of the compounds in Atkinson-King's study were known to all of the subjects at the outset of the experiment, the results confound the equivalent of real and novel compounds at least for the youngest group. We suspect that the 'random' performance reported by Atkinson-King for this age group would instead show significant patterns if

the results were analysed separately for the compounds that were known to the children as opposed to those that were not known. While Atkinson-King attempted to address this issue in advance by training the subjects at the beginning of the experiment, it is not clear that such training results in the same type of processing that established knowledge of an item does. In fact, Atkinson-King did mention in passing that there was a compound bias for three of her six stimuli, and a phrasal bias for the other three, though she did not investigate the implications of this pattern. Interestingly, the three items that showed the bias towards the compound interpretation are those that we would expect to be known by all the subjects (*high chair*, *blackboard*, *hot dog*); the others were the items that we might expect not to be known by the younger subjects (*Red Sox*, *greenhouse*, *redhead*). In fact, *greenhouse* and *redhead* were reported to be unknown to our youngest subjects in our parental questionnaire; we did not test *Red Sox*.

Once we separate the known items from the others, we can begin to gain insight into what is being learned as children acquire the distinction between compound and phrasal stress in English. As far as the novel or unknown items are concerned, the overwhelming interpretation is the phrasal one. There is no significant change according to age with regard to this behaviour. The known items, to the contrary, do show a developmental change with age.

While Figure 7 revealed an exceptionally high percentage of correct scores among the youngest children (average ages 5;4 and 7;2) on the compounds produced with compound stress, closer examination of the results shows that this does not mean that these children have grasped the relevant stress pattern. To the contrary, what we see is a strong bias in favor of a compound interpretation when subjects have a lexical entry consisting of the sequence of the two roots in question, with little regard for the stress pattern. Indeed, the phrasal interpretation is rarely chosen at all, even when the phrasal stress pattern is provided.

The five-year-old subjects thus seem to be applying a strategy that in effect says ‘choose a compound answer if the item can possibly correspond to a compound word in your vocabulary – regardless of the stress pattern.’ So, when these subjects hear *hot + dog*, they ignore the stress pattern and simply choose the lexical item that consists of the segmental string they hear. Recall that, unlike lexical stress, compound and phrasal stress do not have substantial segmental effects on an item. By contrast, when the same subjects hear the unfamiliar item *red + head*, they do not choose the compound meaning since they have no lexical entry that matches this string. Instead, they choose the phrasal meaning consisting of the literal interpretation of each of the two members of the compound as a type of default interpretation. This is essentially the same behaviour we observed with the novel compounds among subjects of all ages.

By comparing the percentage of correct compound answers with the

percentage of overall compound answers in Figure 7, we see that by about seven years, the subjects are beginning to show some sensitivity to the stress pattern of the items they know. That is, these subjects no longer always pick a compound meaning for a string that corresponds to a lexical item if the stress pattern is not consistent with the compound meaning. Thus, the overall selection of the compound meaning for hot dog diminishes to 80%; and by age nine, it is only 60%. It should be recalled that the perfect score on this measure is 50%, since only half of the stimuli were presented with compound stress. This is evidence that the seven- and nine-year-olds are developing the ability to overcome the early overwhelming bias to select a compound interpretation in disregard of the stress pattern. The eleven-year-olds and adults no longer show a bias toward selecting a compound interpretation regardless of stress. That is, they appropriately choose the compound meaning only when compound stress is heard.

The development of the reliance on stress in the interpretation of the known compounds is not paralleled in the interpretation of the novel compounds. Instead, the novel items generally remained at a very low level of compound interpretation since there are no lexical entries for them to match with. Nonetheless, a slightly higher percentage of compound answers was found among the eleven-year-olds for the novel compounds when compound stress was provided. What we suggest is happening here is that these children are at a stage of general development where they are expected to learn massive amounts of new material and many technical terms in school, and thus are more inclined to assume that the novel items were indeed called by the compound name we presented.

Interestingly, some of the adults seemed to be inclined in this direction, too, as we can conclude anecdotally on the basis of some of the comments they made during the experiment. Thus, when faced with a novel item such as *split board*, several adult subjects commented 'I don't know what this thing is called, but I guess it could be a *split board*' and they pointed to the picture of the piano sounding board, as opposed to the picture of the piece of wood split down the middle (i.e. the *split board*). More of the adults, however, commented 'I don't know what a *split board* is, so I'll pick this one', placing the sticky on the picture corresponding to the literal (phrasal) meaning.

In sum, what is revealed in Figure 7 is that the mechanism underlying the gradual acquisition of compound vs. phrasal stress is the ability of the subjects to overcome the bias to call every string a compound if it matches the segmental characteristics of a stored lexical item. The leap in development does not appear, however, until eleven years. At this point we observe, first of all, appropriate selection of the compound interpretation when the compound stress pattern is heard. In addition, we observe a substantial jump (to 68%) in the eleven-year-old group in the selection of the phrasal interpretation when the phrasal stress pattern is presented.

We must thus understand the acquisition process under investigation as the development of the ability to choose a compound interpretation only when compound stress is heard, and by the same token to choose the phrasal interpretation when phrasal stress is heard, even if a string segmentally matches a lexical item. The reason it takes so long for this knowledge to be acquired, we propose, is not that the stress patterns are inherently difficult, but rather that it requires a shift in perceptual strategies.

That is, it seems that, at first, when children hear a word they attend primarily to segmental information. If they can find a match at this level, they are satisfied; they do not consider prosodic differences to be relevant at this point. Eventually, however, a shift takes place and children become able to attend to, and make contrastive use of, both the segmental and prosodic qualities of the speech they hear. At this point, they are able to associate the occurrence of stress on the first word with a compound structure and the occurrence of stress on the last word with a phrasal structure. From a more theoretical perspective, this indicates that the subjects have developed the generalizations that correspond to different stress assignment rules for different types of strings. As mentioned above, we interpret this as involving the acquisition of two distinct post-lexical stress domains, the Clitic Group and the Phonological Phrase, each with its own principles of stress assignment.

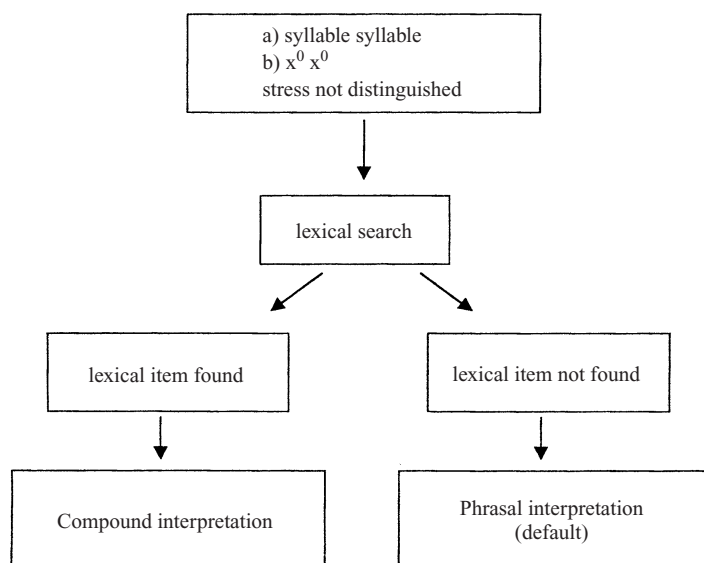


Fig. 8. Compounds vs. Phrases: stress not distinguished.

ACQUISITION MODELS

On the basis of our experimental results, we can identify several stages in the acquisition of compound vs. phrasal stress in terms of the changing roles played by segmental and rhythmic information in the different groups of subjects. We present our findings in the form of two flow charts, Figures 8 and 9 below, roughly corresponding to the original split of the subjects into two groups identified on the basis of overall correct responses shown in Table 1 and Figure 1 above. We believe, however, that the development is gradual, with overlapping behaviours, not that there is a jump between the first and second models.

In Figure 8, two options are shown for the initial stage of development of the Compound vs. Phrasal Stress distinction. Option *a* represents the behaviour of the youngest children who seem to treat the stimuli as unanalysed wholes. These subjects do not realize that a string such as *hot dog* contains internal morphological structure, but rather take it as a simple sequence of two syllables: $[\sigma \sigma]$. This was borne out by informal discussion with a five-year-old child⁶ who found it totally implausible that there could be two words in items such as *hot dog*, *red head*, *high chair*, *blue jay*. A related result was also reported by Berko (1958). When she asked children about the meanings of certain compound words, the majority had no explanation, or provided a personal explanation that did not necessarily reflect the actual constituents of the compound. Indeed, as mentioned earlier, in English compounds there is frequently not a transparent relationship between the meanings of the individual words and the meaning of the whole.

In addition to the absence a semantic analysis, the stress difference between the compounds and phrases is disregarded at this stage, possibly because it is not associated with any segmental allophonic variation, as opposed to lexical stress which typically does induce such variation. Thus, on hearing a stimulus string, a lexical search is performed. If a match is found, the item is identified as a word since words, but not phrases, are listed in the lexicon. If no match is found (i.e. in the case of unknown real compounds and novel compounds), it seems that the individual components are searched separately, and a default phrasal interpretation results. In the youngest subject group, there is greater individual variation with regard to what is listed in the lexicon, so some children find a match for items such as *red head* while others who do not know this term do not. The fundamental process, however, is the same.

At the next stage, represented as Option *b*, children begin to realize that the strings in question consist of two words (X⁰s), even the compounds.

[6] This is the daughter of the first author, who served as an eager and helpful consultant, but not as one of the actual subjects.

ACQUISITION OF COMPOUND VS. PHRASAL STRESS

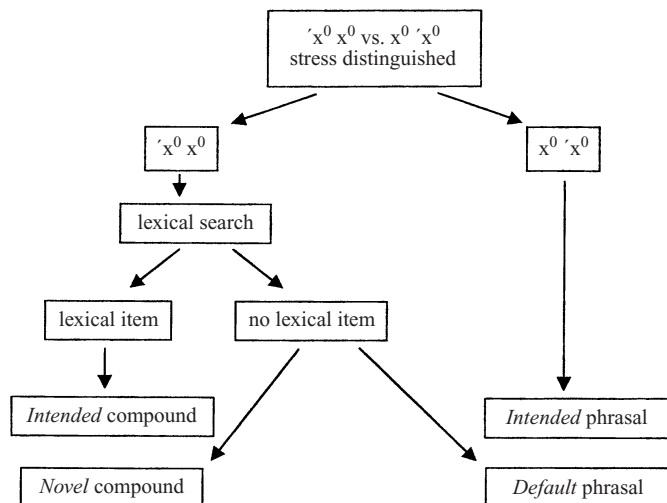


Fig. 9. Compounds vs. phrases: stress distinguished.

Although we do not have reaction time data, we observed during the testing that somewhat older children started to hesitate on the task, realizing that there was a ‘problem’, while the youngest group did not even notice a challenge. We thus find that subjects begin to recognize the two members of a compound as such, though they are not yet able to focus on the stress difference. This is typical of the both the seven-year-olds and the nine-year-olds who still show high percentages of compound interpretation both on the items with compound stress (100 % and 84 %) and on the combined set of items with both stress patterns (82 % and 76 %). These subjects, like the first group, do a lexical search and if they find a (compound) word, they select this interpretation regardless of the stress pattern of the stimulus. If no match is found, the phrasal interpretation is chosen, as a default. By contrast, the phrasal interpretation is correctly chosen only in about one third of the cases, or fewer.

A major change takes place with the eleven-year-olds, who show a large difference between the percentage of compound answers given for items with compound stress, as opposed to the full set of items including those with phrasal stress: 84 % and 58 %, respectively, as seen above in Figure 7. This behaviour is similar to that of the adults, whose corresponding scores were 92 % and 56 %. Thus, as shown in Figure 9, the starting point in the response of the eleven-year-olds and the adults involves two components a) the recognition of different stress patterns and b) the recognition of distinct types of lexical items among the stimuli (i.e. compounds and the individual words of the phrases).

Once the stress pattern of a stimulus item is determined, one of two paths is chosen depending on whether initial (compound) stress or final (phrasal) stress is heard. If phrasal stress is heard, each word is accessed separately, and there is a direct path to the literal phrasal interpretation, which is the intended interpretation in this case. If compound stress is heard, a lexical search is made for a match with the stimulus item. If such a match is found, the compound is identified and the intended compound interpretation is given. If no match is found, as occurs with the novel items with compound stress, we observed two strategies, as mentioned above. Either the novel item is accepted as a compound, with the novel compound interpretation depicted in the stimulus drawings, or it is rejected as a compound and it is assigned a default phrasal interpretation, despite its compound stress pattern. While it seems that the eleven-year-olds may be most open to the possibility of unknown items having novel compound words as names, in general, there seem to be more personal considerations that lead subjects to be more or less inclined to treat the novel items with compound stress as compounds. Most frequently, however, subjects were inclined to reject a compound interpretation, despite compound stress, if the item was unknown to them.⁷

CONCLUSIONS

What our experimental study of compound vs. phrasal stress shows is that the subtleties of this distinction are indeed mastered relatively late – around age twelve, as claimed by Atkinson-King (1973). Our investigation, however, provides a finer grained analysis of the acquisition process than that of Atkinson-King. First, we revealed a fundamental difference between the behaviour of subjects of all ages on known and unknown items. Then, examination of the known items revealed an overwhelming preference for a compound interpretation, regardless of stress. Once different stress patterns were recognized they were used increasingly to permit subjects to accurately select either the compound or the phrasal meaning of a stimulus item. We interpret this as the acquisition of knowledge of the relevant prosodic constituents, the Clitic Group and the Phonological Phrase, and in particular, of the stress patterns associated with each. What is perhaps more surprising is the fact that once the difference between compound and phrasal stress is acquired for known items it is not consistently generalized to novel items. Instead, there appears to be a very strong tendency among most adults to disregard the stress pattern if they happen not to have a lexical compound corresponding to a string they hear for the first time.

[7] One referee suggests that it is possible for children, and even adults, to disregard stress under particular circumstances due to the relatively lighter functional load borne by stress, as opposed to segmental distinctions, in English. It is not possible to evaluate this position at present, in the absence of a well articulated theory of functional load with regard to child language.

Finally, given the above results, we are now in a position to address the apparent paradox mentioned above whereby children can produce novel compounds with appropriate stress by age two, as reported by Clark *et al.*, (1985), but take years longer to distinguish compound from phrasal stress in perception. What we propose is that the knowledge required to distinguish between compound and phrasal stress is quite distinct from, and more abstract than, the ability exhibited by young children when they produce the novel compounds with the correct stress pattern. Specifically, it requires knowledge of higher level prosodic constituents and the rules that govern them. Thus, it seems that in the early production of compounds, stress assignment is part of the compound formation rule itself, though it remains for future research to determine acoustically precisely how accurate the early productions are. Only subsequently is the stress pattern separated from the morphological operation that combines individual lexical items (X⁰s) into larger lexical items (i.e. compound words). In the case of English, this also requires that children develop the (abstract) awareness that the same strong-weak compound stress pattern is used in compounds of varying structures: in A + N (e.g. *hót dog*), N + N (e.g. *tóy store*) and N + V (e.g. *báby sit*). Only in this way, can the general stress assignment rules associated with the different prosodic constituents become formulated as phonological rules distinct from the morphological rules of compound formation.

REFERENCES

- Ashby, M. (1992). Metalinguistic awareness of nuclear accent and of compound stress in an 8-year-old. *European Journal of Disorders of Communication* 27, 175-9.
- Atkinson-King, K. (1973). *Children's acquisition of phonological stress contrasts*. Unpublished doctoral dissertation: University of California, Los Angeles.
- Berko, J. (1958). The child's learning of English morphology. *Word* 14, 150-77.
- Bol, G. & van der Flier, E. (1980). De ontwikkeling van intonatie in het Nederlands: de interpretatie van ambigue vie-vragen. *Spektator* 10, 310-22.
- Chomsky, N. & Halle, M. (1968). *The Sound Pattern of English*. New York: Harper & Row.
- Clark, E. V., Gelman, S. A. & Lane, N. M. (1985). Compound nouns and category structure in young children. *Child Development* 56, 84-94.
- Clumuck, H. (1980). The acquisition of tone. In G. Yeni-Komshian, J. Kavanagh & C. A. Ferguson (eds), *Child Phonology, Vol. 1*. New York: Academic Press. 257-75.
- Cristophe, A., Guasti, T., Nespou, M., Dupou, E. & van Ooyen, B. (1997). Reflections on phonological bootstrapping: its role for lexical and syntactic acquisition. *Language and Cognitive Processes* 12, 585-612.
- Cruttenden, A. (1974). An experiment involving comprehension of intonation in children from 7-10. *Journal of Child Language* 1, 221-31.
- Demuth, K. (1984). *Aspects of Sesotho language acquisition*. Bloomington: Indiana University Linguistics Club.
- Demuth, K. (1989). Problems in the acquisition of grammatical tone. *Papers and Reports on Child Language Development* 28, 81-8.
- Demuth, K. (1991). Acquisition of the Sesotho tonal system. *Afrikanistische Arbeitspapiere* 26, 5-35.
- Demuth, K. (1992). The acquisition of Sesotho. In D. I. Slobin (ed.), *The Crosslinguistic Study of Language Acquisition. Vol. 3*. Hillsdale, NJ: Erlbaum.

- Erbaugh, M. S. (1992). The acquisition of Mandarin. In D. I. Slobin (ed.), *The Crosslinguistic Study of Language Acquisition. Vol. 3*. Hillsdale, NJ: Erlbaum.
- Hochberg, J. (1988). First steps in the acquisition of Spanish stress. *Journal of Child Language* **15**, 273–92.
- Jaeger, J. (1999). Early slips of the tongue and children's developing phonology. Paper presented at the International Congress for the Study of Child Language. San Sebastian: Spain.
- Jaeger, J. (to appear) *Kids' slips: using children's slips of the tongue to understand language development*. Erlbaum.
- Jusczyk, P. & Thompson, E. (1978). Perception of a phonetic contrast in multisyllabic utterances by 2-month-old-infants. *Perception and Psychophysics* **23**, 105–9.
- Jusczyk, P., Cutler, A. & Redanz, L. (1993). Infants sensitivity to predominant stress patterns in English. *Child Development* **64**, 675–87.
- Mandel, D. R., Jusczy, P. & Pisoni, D. (1995). Infants recognition of the sound patterns of their own names. *Psychological Science* **6**, 315–18.
- Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncin, J. & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition* **29**, 143–78.
- Nespor, M. & Vogel, I. (1986). *Prosodic Phonology*. Dordrecht: Foris Publications.
- Sansavini, A., Bertoncini, J. & Giovanelli, G. (1997). Newborns discriminate the rhythm of multisyllabic stressed words. *Developmental Psychology* **33**, 3–11.
- Spring, D. R. & Dale, P. S. (1977). Discrimination of linguistic stress in early speech. *Journal of Speech and Hearing Research* **20**, 224–32.
- Tse, J. K. P. (1978). Tone acquisition in Cantonese: a longitudinal case study. *Journal of Child Language* **5**, 191–204.
- Tuaycharoen, P. (1977). *The phonetic and phonological development of a Thai baby: from early communicative interaction to speech*. Unpublished doctoral dissertation, University College London.
- Vogel, I., Bensink, C., Bol, G., van der Flier, E. & Nagteggall, D. (1979). The development of intonation in Dutch: the interpretation of ambiguous wie-questions. *Publications of the Institute for General Linguistics* **25**, 1–39.
- Vogel, I. & van der Flier, E. (1979). How Dutch children learn to use intonation to disambiguate who-questions. Paper presented at the Annual Linguistics Society of America Meeting. Los Angeles: CA.
- Wieman, L. A. (1976). Stress patterns in early child language. *Journal of Child Language* **3**, 283–6.

APPENDIX I

STIMULUS LIST

Real items	Novel items	Novel compound description
hot dog*	wet screw	type of drill
green house*	split board	part of piano sounding board
black board*	wood cradle	item that carries wood
red head*	long stem	type of saw
high chair*	bear/bare foot	animal's foot
blue jay	red cup	type of flower
toy store	blue banner	type of fish
orange tree	leather strip	type of scraping tool
big birds†	flat bridge	type of paddle

* used in Atkinson-King 1973.

† the compound form is the name of a well-known character on the television programme Sesame Street.

ACQUISITION OF COMPOUND VS. PHRASAL STRESS

APPENDIX 2

ACOUSTIC INFORMATION – PITCH CHANGE FOR WORD 1 AND
WORD 2*(A) Change in pitch (hz) for word 1 in compounds and phrases*

Word 1	Compound	Phrase
Mean	58.117	-19.2
Variance	365.974	359.46
Observations	18	18
Pearson correlation	0.353	
Hypothesized mean difference	0	
df	17	
t Stat	15.142	
$p(T \leq t)$ one-tail	≤ 0.001	
t Critical one-tail	1.74	

(B) Change in pitch (hz) for word 2 in compounds and phrases

Word 2	Compound	Phrase
Mean	-39.106	-45.667
Variance	446.672	2090.672
Observations	18	18
Pearson correlation	-0.573	
Hypothesized mean difference	0	
df	17	
t Stat	0.461	
$p(T \leq t)$ one-tail	0.325	
t Critical one-tail	1.74	

APPENDIX 3

ACOUSTIC INFORMATION – DURATION OF WORD 1 AND WORD 2

(A) Duration in msec of word 1 in compounds and phrases

Duration	Compound	Phrase
Mean	123	127
Variance	0.002	0.001
Observations	18	18
Pearson correlation	0.712	
Hypothesized mean difference	0	
df	17	
t Stat	0.024	
$p(T \leq t)$ one-tail	0.491	
t Critical one-tail	1.74	

(B) Duration in msec of word 2 in compounds and phrases

Duration	Compound	Phrasal
Mean	121	182
Variance	0.005	0.003
Observations	18	18
Pearson correlation	0.593	
Hypothesized mean difference	0	
df	17	
t Stat	-4.5074221	
$p(T \leq t)$ one-tail	≥ 0.001	
t Critical one-tail	1.74	