Published in final edited form as:
Cognition. 2006 July ; 100(3): 389–433. doi:10.1016/j.cognition.2005.04.009.

Knowing what a novel word is not: Two-year-olds 'listen through' ambiguous adjectives in fluent speech

Kirsten Thorpe* and **Anne Fernald**
Department of Psychology, Stanford University, Jordan Hall, Building 420, Stanford, CA 94305, USA

Abstract

Three studies investigated how 24-month-olds and adults resolve temporary ambiguity in fluent speech when encountering prenominal adjectives potentially interpretable as nouns. Children were tested in a looking-while-listening procedure to monitor the time course of speech processing. In Experiment 1, the familiar and unfamiliar adjectives preceding familiar target nouns were accented or deaccented. Target word recognition was disrupted only when lexically ambiguous adjectives were accented like nouns. Experiment 2 measured the extent of interference experienced by children when interpreting prenominal words as nouns. In Experiment 3, adults used prosodic cues to identify the form class of adjective/noun homophones in string-identical sentences before the ambiguous words were fully spoken. Results show that children and adults use prosody in conjunction with lexical and distributional cues to 'listen through' prenominal adjectives, avoiding costly misinterpretation.

Keywords

Language development; Speech processing; Ambiguity resolution; Eye movements; Eye-tracking; On-line measures; Speed of processing; Adjectives; Prenominal adjectives; Form class; Prosody; Lexical development; Grammatical development

---

Adults interpret spoken language incrementally, assigning structure to linguistic input continuously as the speech signal unfolds. Although the process of comprehension seems effortless to the fluent adult, incremental processing in fact leads inevitably and continuously to potential problems of ambiguity as the listener attempts to interpret the meaning of a spoken utterance. Because speech unfolds from moment to moment, it is frequently the case that a number of alternative interpretations are possible at a particular point in time, only one of which will be correct. In a sentence such as *Susie fixed her back porch*, there is ambiguity at multiple levels of linguistic structure at different time points. For example, halfway through the verb, *fix*–, the sentence could go on to describe how Susie either fixes or fixates something, and on hearing *Susie fixed her back*–, the listener might prematurely infer that Susie had consulted a chiropractor rather than a carpenter. For adults, the process of recognizing words and resolving temporary ambiguities in spoken language is remarkably fast and efficient, enabled by the use of predictive cues from multiple sources of information in the speech stream. Using extensive knowledge of language and the world, adults can typically rely on linguistic and contextual information to constrain the possibilities of interpretation, so that understanding happens accurately and effortlessly as the sentence unfolds (e.g. Marslen-Wilson & Tyler, 1980; Spivey, Tanenhaus, Eberhard, &

---

*Corresponding author. thorpe@psych.stanford.edu (K. Thorpe).

Sedivy, 2002; Trueswell & Tanenhaus, 1994). However, numerous studies with adults show that when structural and referential ambiguities *cannot* be immediately resolved and lead to misunderstanding, there are measurable costs to the speed and efficiency of comprehension (see Tanenhaus, Magnuson, Dahan, & Chambers, 2000 for a review).

For young language learners, the problem of potential ambiguity is fundamentally more complex given children's limited knowledge of language structure and the fact that many words they encounter in fluent speech are unfamiliar. In this research, we investigate how 24-month-old children begin to develop efficiency in interpreting potentially ambiguous spoken sentences in real time. To monitor the time course of comprehension, we track children's eye movements as they look at pictures of familiar objects while listening to speech naming one of the objects (e.g. Fernald, Pinto, Swingley, Weinberg, & McRoberts, 1998). Two-year-olds are presented with sentences in which an unfamiliar adjective precedes a familiar noun, as in *Where's the zavvy dog?* Because the novel word *zavvy* is preceded by the determiner *the*, it could in principle be either a noun or an adjective. As in the example of Susie's back porch, at the point when the child hears the novel word *zavvy* it could plausibly be interpreted as a new object name rather than as an adjective preceding an object name, at least temporarily. Our central question is whether when encountering a novel adjective in prenominal position, 24-month-olds first mistake it for an object word and then have to revise their interpretation. To investigate whether 2-year-olds experience adverse effects when encountering an ambiguous word, we examine how misinterpretation of the adjective could potentially interfere with rapid and accurate recognition of the familiar noun that follows.

1. Why apprehending adjectives can be difficult for children

In early lexical development, children typically learn words that name concrete objects sooner and more easily than they learn words describing properties of objects. Explanations for this asymmetry have focused on two broad reasons why adjectives may be more difficult to learn than nouns. First, adjectives as a class are semantically more diverse and quixotic than nouns. Second, given the often ambiguous nature of distributional cues to grammatical form class in continuous speech, adjectives can be hard to single out. Thus for both top-down reasons of conceptual complexity as well as bottom-up reasons of surface-level ambiguity, adjectives may be particularly challenging for the young language learner to identify and interpret.

1.1. The semantic subtlety of adjectives

Although we learn in grammar school that adjectives refer to properties of objects such as color and shape, the class of adjectives in English is semantically quite heterogeneous. Words like *blue* and *smooth* would seem to be classic examples of words for physical attributes of objects, while the meanings of *big* and *short* depend more on relative judgments involving comparisons among objects. But just as a mouse and a house can both be described as *big* while differing in absolute size, *blue* and *smooth* can also be associated with different perceptual properties depending on whether these words refer to plums or to oceans. In the case of evaluative adjectives, it is obvious that both a car and a bar could be aptly described as *good* or *slow* or *noisy* yet have little in common in terms of physical or functional attributes. Thus the adjectives in all of these examples are dependent to varying degrees on the nouns they modify for their own scope of interpretation.

Because the denotations of words like *big* and *smooth* can vary considerably in relation to different nouns, understanding the meanings of adjectives depends on linguistic knowledge to a greater extent than is the case for concrete nouns. Gentner (1982) refers to this characteristic as 'relational relativity' and argues that verbs, adjectives, prepositions, and

other relational terms are cross-linguistically variable in how they map from concept to word. In contrast, concrete and proper nouns refer to things that are perceptually coherent and easily individuated even by infants. While nouns are used to label categories of objects that share many correlated features, adjectives have far fewer correlated features because they shift semantically in different linguistic contexts (Markman, 1989). As Gasser and Smith (1998) point out, all the things we call *apple* tend to share many more perceptual attributes in common than do all the things we call *red*. The conceptual complexity of adjectives relative to concrete nouns accounts in part for the fact that children are typically slower to learn property terms than object names (Gentner & Boroditsky, 2001).

1.2. Identifying adjectives in continuous speech

In addition to the semantic subtleties that make adjectives difficult for children to apprehend, adjectives are problematic at another level as well. When encountering an unfamiliar word in a sentence, fluent listeners make use of a variety of linguistic cues to decipher what kind of word it might be. The grammatical form class of an unfamiliar noun can usually be identified on the basis of reliable morphosyntactic and positional markers, such as the presence of a determiner preceding the novel word. Determiners, such as *a*, *the*, *this*, *some*, etc. can also be followed by adverbs and adjectives as well as by nouns. However, because nouns occur more frequently overall than do adverbs or adjectives in adult as well as child-directed speech, it is very likely that the word immediately following *a* or *the* will turn out to be a noun. As part of an analysis of child-directed speech described later we found that when the word *the* was used it was followed immediately by a noun 93% of the time[1].

In contrast, the linguistic cues associated with adjectives are less trustworthy. While certain bound morphemes occur frequently with adjectives, including suffixes like *–ish* and *–y* (e.g. *childish*, *gummy*), similar word endings can also occur in nouns (e.g. *goldfish*, *tummy*). Other cues to the form class of adjectives rely on the position of the word within particular syntactic frames. However, these too are not always reliable cues given that nouns and adjectives sometimes appear in identical contexts. Imagine you were asked to guess what kind of English word is replaced by the nonce word *blane* in the following sentences:

1. **a.** That's *blane*.

b. That woman is *blane*.

c. That's a *blane* woman.

d. That's very *blane*.

Based on the syntactic frame in (1a), *blane* could be either the name of a substance (e.g. *That's plastic*), a proper name (e.g. *That's Mary*), or a property term (e.g. *That's big*). In (1b) as well, *blane* could plausibly stand for *plastic* or *Mary* or *big*. Only when the novel word is heard preceding a familiar noun in (1c) is it evident that *blane* must be an adjective. Indeed the most reliable syntactic indicator that a word is an adjective in English is its occurrence in prenominal position. Another unique marker of adjectives is that they can follow qualifiers such as *very*, *too*, or *quite*, as in (1d), although such combinations do not occur with high frequency.

[1]This proportion was based on counts from a subset of the transcripts used for the corpora analysis described in detail in the general discussion section of this paper. Every tenth transcript from the larger analysis of adult speech to five children from the CHILDES database (MacWhinney, 2000) was searched for occurrences of the word *the*, yielding a total of 882 usages, of which 93% were followed directly by a noun.

In the examples above, (1c) provides the most common context that clearly identifies the unknown word as an adjective. On hearing *That's a snod stone*, adults can determine that *snod* is an adjective without knowing that this archaic English word means *smooth*, because *snod* is positioned before the familiar noun *stone*. While prenominal position provides a highly reliable cue to form class for adjectives, it also presents potential problems for the listener interpreting speech from moment to moment, because the diagnostic information does not appear until *after* the unfamiliar word has been spoken. Given that determiners in English are more likely to be followed by nouns than by any other type of word, the best bet for the listener relying on distributional cues alone is that any novel word following *the* or *a* is a noun rather than an adjective. Thus as speech is interpreted incrementally in real time, the listener encountering a determiner followed by *snod* could momentarily mistake this new word for a noun. The alternative would be to delay processing, waiting to hear what comes after *snod* in order to figure out what kind of word it is.

In addition to morphosyntactic and positional cues, prosodic cues also provide information potentially useful in identifying novel words as adjectives in continuous speech. No prosodic patterns in English are uniquely associated with adjectives as opposed to nouns, but when an adjective precedes a noun and is not given contrastive emphasis, it is typically deaccented relative to the noun that follows. Prosodic accent is conveyed by a combination of acoustic features including lengthening of the vowel, a rise in pitch followed by a fall, and an increase in amplitude (Cooper, Eady, and Mueler, 1985; Crystal and House, 1988; O'Shaughnessy, 1979). When the head noun of a phrase receives an accent, the preceding adjective is often deaccented (i.e. shortened in vowel duration, lower in intensity, with gradual pitch decline). If sentence (1c) above were spoken with a typical prosodic contour across the noun phrase, the stronger prominence of the familiar final word relative to the unfamiliar penultimate word could be used by the listener as evidence that *blane* is a prenominal adjective. In fact even before the listener hears the familiar final word, the prosody might be sufficient to indicate that the deaccented unfamiliar word is not the last word in the phrase. Several studies have shown that adults can use this type of prosodic information to determine whether they have heard the final word in a sentence or whether the sentence will continue (Grosjean, 1983; Grosjean & Hirt, 1996; Warren, Grabe, & Nolan, 1995). Thus the experienced listener could know that the novel word *blane* is not phrase-final even before getting to the final noun, a real-time processing cue potentially indicating that the unknown word is in the prenominal position unique to adjectives.

Pointing out that nouns and adjectives occur to some extent in different syntactic and prosodic contexts is suggestive but gives no information as to whether these cues are consistently available to young language learners and thus might actually be instrumental in the identification of grammatical categories. To address this question, recent distributional analyses of children's early language input have investigated how the co-occurrence patterns of words in caregivers' speech could yield information about linguistic structure. By categorizing 200 common words in a large corpus of child-directed speech based on their co-occurrence with surrounding words Mintz, Newport, and Bever (2002) succeeded in classifying correctly the majority of nouns and verbs. However, adjectives were frequently misclassified as nouns, an unsurprising outcome given that nouns are much more frequent than adjectives and both appear in similar contexts. While lexical co-occurrence patterns alone may not be sufficient for distinguishing adjectives from nouns, prosodic cues used in conjunction with co-occurrence patterns could offer more reliable information. Research with connectionist models has shown that combining multiple probabilistic cues is more informative than using single cues in isolation. Christiansen, Allen, and Seidenberg (1998) provided their model with information about phonemes, lexical stress, and boundaries between utterances in a corpus of child-directed speech. Although each individual source of

information on its own provided relatively unreliable cues to word boundaries, they were highly effective when used in combination.

2. Learning the meanings of adjectives

Because of their complexity at the conceptual level and their potential ambiguity at the level of surface structure, adjectives should by all accounts be harder for children to learn than nouns. Indeed observational studies show that young language learners acquire several nouns before producing any adjectives (Fenson et al., 1994; Nelson, 1976). Moreover, numerous experimental studies of word learning have found a disparity between young children's ease in learning novel object names and their difficulty in learning new property terms. The typical procedure in this type of study is for the experimenter to present the child with a distinctively shaped novel object with a special property such as a nubbly texture, saying either *This is a TIV* or *This is a TIV one*. The child is then shown at least two novel test objects, one matching the original stimulus in kind (i.e. same shape, different texture) and another matching in property (i.e. same texture, different shape), and asked to find either *another TIV* or *another TIV one*. If the novel word in the phrase *a TIV one* is interpreted as an adjective, the child should choose the property match over the kind match. While children under 3-years-old reliably interpret *This is a TIV* as referring to a category name, they fail to interpret *This is a TIV one* as referring to a property, extending the novel word to the test object that matches in kind rather than in texture (Hall, 1994; Hall, Waxman & Hurwitz, 1993; Klibanoff & Waxman, 2000; Taylor & Gelman, 1988; Waxman & Kosowski, 1990, Waxman & Markow, 1998). Although consistent with the observation that adjectives are generally harder to learn than nouns, the magnitude and robustness of these negative findings are somewhat surprising given that by 30 months of age English-learning children are typically speaking upwards of 50 descriptive words (Dale & Fenson, 1996).

By presenting children with a novel word that is framed syntactically as either a count noun (*a TIV*) or an adjective (*a TIV one*), the goal of these experiments is to determine whether the child can use syntactic cues to figure out the meaning of the unfamiliar word. The logic of this question rests on assumptions about two capacities: First, that young children can make use of surface structure cues, such as word position and morphology to identify the form class of the novel word as either noun or adjective; and second, that these children have abstract grammatical categories for both nouns and adjectives, such that categorizing the novel word syntactically leads to inferences about its meaning as either an object name or a property term. Note that the second of these capacities presupposes the first, since correct identification of the form class of the word is essential for determining its meaning. However, the child's appreciation of how surface structure cues relate to grammatical category is generally taken for granted in this literature, even though this relation is far from transparent for the many reasons described earlier. The puzzling failure of young children to interpret the novel word in the phrase *a TIV one* as referring to a property of the novel object, not the object itself, is typically explained in terms of the conceptual complexity of adjectives as compared to nouns, without consideration of the prior problem of how to identify the novel word as an adjective in the first place. Is it possible that children might parse this ambiguous phrase incorrectly as *a TIVwun*, and thus interpret the novel word as a noun? Given that the novel word in the phrase *a TIV one* is naturally accented, the compound *TIVwun* would have the strong/weak metrical pattern characteristic of 80% of all bisyllabic nouns in English (Cutler & Carter, 1987). Since stressed syllables are perceived even by 9-month-old infants as marking the onset of bisyllabic words (Jusczyk, Cutler, & Redanz, 1993), it is possible that the prosodic pattern in *a TIV one* might bias young listeners toward a noun interpretation, independent of any conceptual difficulties associated with adjectives as opposed to nouns. The present research explores this possibility by

investigating how accent patterns influence children's tendency to respond to a novel adjective as if it were a noun.

3. The cost of misinterpretation in on-line speech processing

Spoken language processing by fluent listeners requires not only rapid lexical access but also efficiency in resolving ambiguity at multiple levels simultaneously. Many studies have used speed of processing tasks to monitor how adults respond when they encounter a temporary syntactic ambiguity, as in the famous example from Bever (1970), *The horse raced past the barn fell.* When hearing this 'garden path' sentence, the listener first interprets the word *raced* as the main verb, when in fact *raced* functions as the beginning of a reduced relative clause. When listeners parse the sentence incorrectly at the point of ambiguity, they then have to revise their interpretation on the basis of subsequent information (e.g. Frazier & Rayner, 1982; Crain & Steedman, 1985). Although previous research has focused primarily on syntactic indeterminacy, temporary ambiguities are widespread at other levels in speech as well (Tanenhaus et al., 2000). Given that sentences are processed incrementally as they unfold (Marslen-Wilson & Tyler, 1980), a prenominal adjective in English can be temporarily ambiguous until the following head noun becomes clear later in the sentence, as in the earlier example of Susie fixing her back porch. Such ambiguity can lead to misinterpretation that reduces the speed and efficiency of understanding.

3.1. How adults interpret adjectives in on-line comprehension

By recording eye movements as adult listeners hear sentences referencing objects in the scene in front of them, experimenters can monitor the time course of speech comprehension from moment to moment as listeners resolve temporary ambiguities (e.g. Tanenhaus et al., 2000). In one such eye-tracking experiment, participants were presented with sentences such as *Put the apple on the towel into the box*, which contains a temporary ambiguity at the first prepositional phrase (Tanenhaus et al., 1995). Because the phrase *on the towel* can initially be interpreted either as the destination of *put* or as a modifier of *apple*, this ambiguity could lead a listener down a garden path of misinterpretation. Referential support (e.g. having two apples present in the visual array, one of which is on a towel) facilitates disambiguation, enabling listeners to correctly interpret the phrase in real-time. However, when the situation is left temporarily ambiguous (e.g. just one apple is present) adults' visual fixation patterns reflect their initial misinterpretation. This and other examples demonstrate how researchers use gaze patterns in response to speech to measure the efficiency of processing and the inefficiency of misinterpretation.

Similar types of efficiency have been found in studies of lexical ambiguity resolution. Allopenna, Magnuson, and Tanenhaus (1998) demonstrated that when words referring to two objects in a scene shared common initial phonemes (e.g. *beaker* and *beetle*), both objects were considered as potential referents up to the point in the speech stream when they were clearly disambiguated (e.g. at the sound of /k/ or /t/). In contrast, in unambiguous conditions when identification was possible at the start of the target label, looks to the appropriate referent occurred more rapidly. Efficiency in resolving referential indeterminacy has also been explored using adjectives as target stimuli. Because prenominal adjectives often serve to restrict the domain of possible reference, Eberhard, Spivey-Knowlton, Sedivy, and Tanenhaus (1995) asked whether listeners could process adjectives incrementally to identify a particular referent within a larger set of possibilities. When adult participants heard instructions such as *Touch the plain blue square*, their responses differed depending on what other objects were present in the visual array. When there was only a single object that was both plain *and* blue, listeners could identify the referent upon hearing *blue*; however, with several plain blue objects present, participants waited to make their choice

NIH-PA Author Manuscript

NIH-PA Author Manuscript

until they heard the word *square* (also see Sedivy, Tanenhaus, Chambers, & Carlson, 1999). At ambiguous moments in speech such as these, efficient processing requires the listener to *postpone* responding until more information becomes available before committing to a particular interpretation. This tendency to hold off in responding to an ambiguous word will be referred to here as 'listening through'.

3.2. On-line speech processing by young language learners

With the refinement of eye-tracking techniques for use with infants, it is now possible to monitor the time course of spoken language understanding by very young language learners as well as by experienced adults. Using a looking-while-listening procedure with infants from 15 to 24 months of age Fernald et al. (1998) found dramatic gains in the speed and accuracy of word recognition over the second year. Infants looked at pictures of familiar objects while listening to speech naming one of the objects. Fifteen-month-olds responded inconsistently and shifted their gaze to the appropriate picture only after the offset of the target word, while 24-month-olds were faster and more reliable, initiating a shift in gaze before the target word had been completely spoken. A recent longitudinal study following infants from 12 to 25 months found that on-line measures of efficiency in speech processing were correlated with numerous more traditional measures of lexical and grammatical development (Fernald, Perfors, & Marchman, in press). Success at word recognition in degraded speech is also correlated with vocabulary size in the second year (Zangl, Klarman, Thal, Fernald, & Bates, 2005), further evidence that efficiency in speech processing is related to other dimensions of early language development.

The increase in response speed over the second year enables infants to identify unambiguous words more quickly based on partial phonetic information, rather than waiting until the word is complete. However, one consequence is that the young language learner is increasingly confronted with problems of temporary ambiguity. When Allopenna et al. (1998) presented adults with objects that included candy and a candle and asked them to *Pick up the can-*, they waited to hear the next speech sound before orienting to the appropriate object. That is, they postponed their response until the final syllable of the target word made it clear which object was the intended referent. The child who hears *Where's the doll?* in the presence of a doll and a dog is also faced with a temporary ambiguity, given that *doll* and *dog* overlap phonetically and thus are indistinguishable for the first 300 ms or so. Swingley, Pinto, and Fernald (1999) found that 24-month-olds in this situation also delayed their response by about 300 ms until disambiguating information became available. This result leads to one of the questions motivating the present study: If 2-year-olds can postpone their response to familiar words that initially overlap, waiting until the ambiguity is resolved before orienting to an appropriate target object, are they also able to avoid responding prematurely to other kinds of ambiguity in the speech stream?

A closely related issue is how to determine when a 2-year-old is in fact 'misinterpreting' an ambiguous word. When adults misinterpret an ambiguous word or a garden-path sentence, the resulting cost in processing time can be quantified in terms of disruptions in reading fluency, lexical decision and other measures not easily used with young children. What form will disruption take when infants encounter ambiguities, especially those they cannot resolve? In particular, when the child encounters a novel word preceded by the determiner *the* and followed by a familiar noun, will the novel word be misinterpreted as a potential object name? And if so, what kinds of behavior would justify the inference that the unfamiliar word is being considered as an object name rather than as a potential modifier? These questions are addressed in relation to our findings, but a relevant background study should be mentioned here. Swingley and Fernald (2002) investigated how 24-month-olds would respond upon hearing a target word that matched neither of the pictures available in the array. For example, while looking at pictures of a car and a dog, children heard three

types of target word: *car* or *dog* (familiar/present), *shoe* (familiar/absent), or *gizmo* (unfamiliar/absent). Children shifted quickly and reliably to the correct picture in response to familiar/present words, and more slowly and erratically in response to unfamiliar/absent words. The surprising finding was that children also shifted quickly upon hearing a familiar/absent word, then switched back and forth when they failed to find a match. This result suggests that if they recognized the target word as an object name, 24-month-olds initiated a search for a potential referent even when the word matched neither of the pictures available.

If 24-month-olds initially misinterpret prenominal adjectives as potential object names, the Swingley and Fernald (2002) results give us an indication of what behavior we might expect to observe. Upon hearing *Where's the zavvy*… while looking at a picture of a car, the child interpreting the novel word as an object name should shift away from the car to search for a referent. However, when the sentence continues and *zavvy* is followed by *car*, the child would then have to recover from such a 'false alarm' to respond correctly to the target noun once they hear it. This recovery would most likely delay children's speed of orienting to the target picture. Thus, if 2-year-olds are unable to listen through a prenominal adjective and instead mistake it for a potential object name, we would expect slower and less accurate responding to the target noun that follows the ambiguous word.

4. Summary and overview of experiments

In three experiments, we explore how both inexperienced and experienced listeners respond to potentially ambiguous prenominal adjectives in fluent speech, and what prosodic cues might help them to avoid misinterpretation. Experiment 1 investigates whether 24-month-olds are able to listen through familiar and unfamiliar adjectives preceding a familiar object name, or whether the efficiency of their response to the target noun is disrupted by the prenominal word. Since placing emphasis on a prenominal adjective may make it sound prosodically more like a final noun, we also ask whether accented adjectives interfere more with identification of the following target noun than deaccented adjectives. Experiment 2 explores the nature and extent of the disruption in target word recognition that might occur if children *did*, in fact, interpret the prenominal word as a potential object name. In Experiment 3, we test adult participants in a gating procedure to determine what prosodic cues experienced listeners are able to use in identifying ambiguous homophones as either prenominal adjectives or sentence-final nouns.

5. Experiment 1

In Experiment 1, infants listened to naturally spoken sentences in which an adjective preceded the familiar target noun at the end of a sentence, along with control sentences that contained no adjective. Two factors were manipulated: whether the prenominal adjective was familiar or unfamiliar to the child, and whether or not accent was placed on the adjective. The adjectives in the familiar condition (*good*, *nice*, *new*, *pretty*) were chosen according to two criteria: First, these adjectives were all uninformative, in that they did not refer to the perceptually accessible properties of the target objects. Second, norming data from the MacArthur-Bates Communicative Development Inventory (Fenson et al., 1994) indicated that at least 50% of children at this age produce these four adjectives. In our sample, 67% of the children used at least one of these words productively and 48% used two or more, according to maternal report. Because these words are among the most common adjectives heard in speech to toddlers, we could assume that they were familiar to all the participants, even if some children were not yet using them in production. In contrast, the adjectives in the unfamiliar condition (*glib*, *lace*, *skimpy*, *faux*) were assumed to be novel words because none was present in any child's productive vocabulary, and all are typically rare in child-directed speech.

6. Method

6.1. Participants

Participants in Experiment 1 were 64 24-month-olds (*M*=105 weeks; range=103–108 weeks). Twenty-seven were girls, and all participants were from homes in which English was the primary language spoken. An additional eight participants were excluded from the sample due to inattentiveness (i.e. failure to look at either of the pictures when the speech stimuli were presented on more than 50% of the trials, *n*=3), fussiness (i.e. crying, or struggling to get down from their caretaker's lap, *n*=4), or technical difficulties (*n*=1).

6.2. Speech stimuli

Each stimulus sentence contained one of eight target words (*dog*, *bunny*, *duck*, *cat*, *flower*, *car*, *ball*, *shoe*) preceded by a familiar adjective, an unfamiliar adjective, or no adjective. Each target word appeared in all three sentence types, yielding 24 critical trials. Each familiar adjective was paired with one of the unfamiliar adjectives, and each pair was then yoked to one pair of target words and corresponding pair of pictures. For example, *good* and *glib* were used with both members of the pair *doggy* and *bunny*. To add variety and maintain children's interest in the pictures, a second uninformative filler sentence followed each target sentence (e.g. *Look at that!*). Six filler trials with six additional target nouns and corresponding pictures were also interspersed among the test trials.

A female native speaker of American English prerecorded multiple tokens of each sentence type with each target word. All stimulus sentences were spoken slowly and clearly. Two versions of the stimulus sentences were recorded. One version of the sentences containing prenominal adjectives had accent placed on the noun and the adjective was deaccented. This pattern is typical in conversational English when the adjective is not intended to express contrast (Bolinger, 1986)[2]. The other versions of the same sentences had accent placed on the adjective preceding the target noun. To ensure prosodic comparability across speech stimuli, the final tokens used in testing were selected based on acoustic measurements of the duration of the carrier phrase (*Where's the…*) and the prenominal adjective. In the final stimulus set, all carrier phrases were comparable in duration within 32 ms, and adjectives were matched within 68 ms within each type (deaccented or accented). In a few instances, minor editing of the stimuli was required to achieve a satisfactory match using Peak™ LE 2.62 waveform analysis and editing software. The mean duration of adjectives in the familiar and unfamiliar adjective conditions was 392 ms for the deaccented adjectives and 484 ms for the accented versions. The mean duration of the target nouns in both accent level conditions was 496 ms.

6.3. Visual stimuli

The visual stimuli were digitized full-color photographs of objects named by the target words, presented on gray backgrounds. Pictures were presented in pairs (dog/bunny, duck/cat, flower/car, ball/shoe) and matched approximately for size and relative salience within pairs.

[2]It is difficult to determine from corpora of parental speech just how often children hear prenominal adjectives used contrastively, given that detailed information about prosody is not typically provided. Other contextual information can also be critical in determining whether a word is used contrastively, such as what objects are present, what the speaker is attending to, etc, details that are also not available in most corpora. What is clear from the linguistic structure alone is that children do frequently hear prenominal adjectives used in a non-contrastive way, as in *What a big doggy* and *You're a good helper*.

6.4. Apparatus and procedure

The experiment was conducted in a room containing a three-sided cloth-walled booth measuring 1×1.2×2 m tall. Flush with the back of the booth were two 15-inch color computer monitors aligned side-by-side 60 cm apart on which the pictures were presented during testing. The caretaker sat facing the monitors with the child positioned on her lap, approximately 45 cm from the back of the booth. Speech stimuli were played over a concealed central speaker on the floor directly in front of the child. The child's gaze patterns were recorded with a concealed video camera positioned above the speaker between the two monitors.

At the beginning of the experiment, the caretaker and child were escorted by an experimenter into the testing room. After the caretaker signed a consent form and the child became comfortable with the surroundings, they were seated in the testing booth. A curtain was drawn between the two so that the child's responses would not be influenced by the caretaker. When the child was attentive the trials began. On each trial the pictures appeared on the monitors for 2 s before the sound stimuli began, to give the child a chance to see both pictures before hearing the sentence. Trials were separated by a 1-s pause when the screens went blank. Half the children heard versions of the stimulus sentences with the prenominal adjective deaccented, and half heard the accented prenominal adjective versions. The 30 trials were presented in two blocks of 15 each. Within each block, trial types were presented in a quasi-random order such that trials of the same type did not occur consecutively and the target picture did not appear on the same side more than three trials in a row. Each picture served equally often as target and distracter, and appeared equally often on the left and right sides. Four different orders were counterbalanced to control for possible effects of the sequence of trials.

6.5. Coding

A digital time-code accurate to 33 ms was recorded onto the videotape, as well as visual markers indicating the onset and offset of the pictures and the speech stimulus on each trial. Highly trained observers, blind to trial type, target word, and side of target picture, coded the videotape off-line in a frame-by-frame analysis of the session. Using custom software, coders indicated on each frame whether the infant's eyes were oriented to the left picture, to the right picture, between the pictures, or away from both. The time at which the infant initiated each shift in gaze, and the duration of each fixation was calculated by the computer. Eye movement data were then temporally aligned with the onset of the target word on each trial. This yielded a high-resolution record of the time course of gaze patterns to the pictures as the stimulus sentences unfolded.

Inter- and intra-observer reliability checks were conducted routinely for all coders. To assess the reliability of our coders, we selected sets of 8 trials on which there were at least two or more gaze shifts for 25% of the subjects, and these trials were coded independently by two coders. Two measures of inter-observer reliability were made. The first measure was the proportion of frames (33 ms units) over each trial on which the two coders agreed within one frame. In this case, agreement was 97%. However, since this analysis included many frames on which the child was simply maintaining fixation on one of the pictures, we also calculated a more stringent test of reliability. This more conservative measure focused only on the shifts in gaze, ignoring the steady-state fixations in each trial on which agreement was inevitably high. By this measure coders agreed within one frame 92% of the time.

6.6. Measures of speech processing efficiency

Because children could not know which of the two pictures presented would be named by the target word on a given trial, they were by chance equally likely to be looking at the

NIH-PA Author Manuscript

NIH-PA Author Manuscript

target or the distracter picture at the onset of a target word. At that point, two correct responses were possible: if a child happened to be looking at the target picture already (i.e. target-initial trials), then she should continue to do so upon hearing the target word; however, if she happened to be looking to the distracter picture (i.e. distracter-initial trials), she should shift gaze to the target picture upon hearing the target word. Three measures were used to assess different aspects of children's efficiency in identifying the familiar target noun at the end of each stimulus sentence: reaction time and accuracy reflected how rapidly and correctly children oriented to the appropriate picture, while shift tendency reflected children's tendency to respond inefficiently by shifting from one picture to the other prematurely before the onset of the target word.

6.6.1. Accuracy—The accuracy of the child's response on each trial was calculated based on the time spent looking at the target picture as a proportion of total time the child looked at either the target or distracter picture. Looking time was averaged over the 333–1800 ms window following the onset of the target word. This time window was chosen based on previous research using this same procedure (e.g. Fernald, Swingley, & Pinto, 2001). First, looking at either picture during the first 333 ms following target word onset was not measured, on the assumption that these looks had been initiated before the word was heard, and were irrelevant to linguistic processing of the target word. Second, only looking to the target that occurred within 1800 ms of the first target word onset was analyzed, giving children this age ample time to make a response without including too much extraneous looking information.

6.6.2. Reaction time—Reaction times were calculated only for distracter-initial trials. These were trials when the child happened to be fixating the wrong picture at target word onset and then shifted correctly to the target picture. Reaction time could not be calculated on target-initial trials, as a correct response on those occasions was for the child to stay fixated where they were already looking. Again, in measuring reaction time only shifts that occurred in the window from 333 to 1800 ms from target word onset were included in the analyses, because shifts made before 333 were considered too early to be in response to hearing the target word, and those made after 1800 were considered too slow to be responses and were more likely to be random noise.

6.6.3. Shift tendency—This measure was designed to capture children's tendency to 'false alarm', i.e. to shift from one picture to the other during a prenominal adjective before hearing the target noun. If a child initially mistakes a prenominal adjective for a noun, we would expect the child to shift between pictures in search of a match (e.g. Swingley & Fernald, 2002). Thus, a high shift tendency during prenominal adjectives would reflect greater uncertainty on the part of the child. On the other hand, if the child does not interpret the prenominal word as a possible noun and waits to hear the target word before responding, we would expect the shift tendency during prenominal adjectives to be low. Shifts in gaze were counted during the window starting at 333 ms from the onset of the prenominal adjective through the first 333 ms of the target noun, before the response could have been elicited by the target noun itself. The shift-tendency measure reflected the proportion of total trials on which shifts occurred during this window.

7. Results

Fig. 1A–C show the time course of children's orienting to the named target picture in the deaccented and accented adjective conditions, with time from target word onset on the *x*-axis. These graphs depict changes in the proportion of total trials on which children were looking at the correct picture at every 33 ms time interval as they listened to the stimulus sentence. Each curve represents children's aggregate responses in one adjective-type

condition (familiar, unfamiliar, or none) for either the deaccented or accented adjective group. In all conditions, children were initially at chance at the beginning of the target word. As the target noun unfolded, orienting to the target picture began to increase rapidly as children correctly identified the named object.

The data from the *deaccented* adjective condition are shown in Fig. 1A. The curve marked by squares shows the time course of children's responses to target nouns in the 'none' control condition, when no adjective was present. The curves marked by circles and triangles represent responses to familiar and unfamiliar deaccented adjectives, respectively. The high degree of overlap among these curves suggests that children were equally efficient in recognizing the target noun regardless of whether or not an adjective was present, and whether or not the adjective was familiar. In Fig. 1B, the data from the *accented* adjective condition present a somewhat different pattern, with variations in slope suggesting differences in the speed of responding to control and adjective sentences. To enable direct comparison of children's response to all four sentence types with prenominal adjectives, Fig. 1C combines the deaccented and accented adjective conditions in Experiment 1, omitting data from the control sentences. Note the gradual decline in speed of response to the target word across the four conditions.

7.1. Accuracy

For purposes of statistical analysis, accuracy was quantified as the mean proportion of looking time to the correct picture, averaged over the window from 333 to 1800 ms following target word onset. Accuracy scores were calculated for each child in the deaccented and accented conditions and in each of the three adjective-type conditions (see Table 1). These accuracy scores were entered into a 3 (adjective type: none, familiar, unfamiliar) ×2 (accent level: deaccented, accented) mixed ANOVA with adjective-type as a within subjects factor and accent level as a between-subjects factor. There was no main effect of adjective type on children's accuracy. However, accuracy scores tended to differ by adjective type for children in the accented condition, but not in the deaccented condition, as shown by the marginally significant interaction between adjective type and accent condition, $F(2, 124)=2.52$, $P=0.085$. The ANOVA also revealed a main effect of accent level, $F(1, 62)=13.82$, $P<0.001$, indicating that children in the deaccented adjective condition were more successful overall in target word recognition than children in the accented adjective condition.

While there was no effect of adjective type in the deaccented adjective condition, planned comparisons revealed a significant effect of adjective type, $F(2, 62)=3.14$, $P<0.05$ within the accented adjective condition. When an unfamiliar accented adjective preceded the target noun, 24-month-olds were significantly less accurate in target word recognition (M=64.8%, SD=14.2%) than when no prenominal word was present (M= 73.2%, SD=12.3%), $t(31)=2.52$, $P<0.05$. However, when the accented adjective was familiar (M=67.3%, SD=15.4%), children's accuracy did not differ significantly from the control condition, $t(31)=1.63$, $P>0.50$.

These accuracy results for Experiment 1 indicate that when children encountered deaccented prenominal adjectives in continuous speech, their efficiency in recognizing the subsequent target noun was not affected, regardless of whether or not the adjective was familiar to them. However, when children heard sentences with *unfamiliar accented* adjectives, their ability to identify the subsequent target noun was disrupted, as compared to their performance on sentences with no prenominal word. Accurate responding in the familiar accented adjective condition did not differ reliably from the other two conditions.

7.2. Reaction time

Mean reaction times to shift correctly from the distracter to the target picture were calculated for each child for all three sentence types in both the deaccented and accented conditions, as shown in Table 2. When mean reaction times were compared among conditions in a mixed design ANOVA, the main effect of adjective type was significant, $F(2, 112)=4.97$, $P<0.01$. Children responded most rapidly to target words with no prenominal adjective, and were slower to respond to target words preceded by familiar and unfamiliar adjectives. The main effect of accent was not significant $F(1, 56)=1.19$, $P> 0.10$, and there was no significant interaction of adjective type and accent, $F(2, 112)= 0.76$, $P>0.10$.

Although there was no significant effect of accent level, planned comparisons were conducted to test our hypothesis that accented adjectives would be more disruptive to young listeners than deaccented adjectives. These pair-wise comparisons provided a more sensitive comparison between conditions, since children who provided reaction time data in only two of the three conditions could also be included. For children who heard deaccented prenominal adjectives, there were no significant differences in reaction time among sentence types, $F(2, 54)=2.22$, $P>0.10$. However, reaction times for children in the accented adjective condition were significantly affected by sentence type, $F(2, 58)= 3.65$, $P<0.05$. Pair-wise comparisons showed that response times to target words following both familiar accented adjectives ($M=721$, SD=15) and unfamiliar accented adjectives ($M=752$, SD=17) were significantly greater than response times in control sentences ($M=655$, SD=14), $t(29)$ none/unfamiliar=−2.66, $P<0.05$; and $t(29)$ none/familiar=−1.99, $P<0.05$. Thus, children were slower to respond when the target word followed an *accented* prenominal adjective, whether it was familiar or unfamiliar, although when the target noun was preceded by a *deaccented* adjective they responded just as quickly as when no adjective was present at all.

7.3. Shift tendency

To compare children's tendency to shift gaze prematurely before the target noun was spoken for each trial type, we calculated the proportion of trials on which shifts occurred during the prenominal adjective. A 2 (adjective type)×2 (accent level) mixed factors ANOVA with adjective type as a within-subjects factor and accent as a between-subjects factor revealed a main effect of adjective type, $F(1, 62)=6.27$, $P<0.05$, but no main effect of accent, and no significant interaction of adjective type and accent level.

Consistent with the previous analysis, planned comparisons showed that in the deaccented condition there was no difference in shift tendency during familiar (M= 32.4%, SD=22.0%) vs. unfamiliar prenominal adjectives (M=36.1%, SD =20.3%). However, in the accented condition children were significantly more likely to shift prematurely, or 'false alarm' in response to an unfamiliar adjective (M=42.0%, SD= 25.6%) than to a familiar adjective (M=29.8%, SD=22.3%), $t(31)=-2.68$, $P<0.05$. Thus when children heard an unfamiliar adjective that was deaccented, they did not shift prematurely; however, when the unfamiliar adjective was accented, the false-alarm rate increased and children were more likely to shift before hearing the target word.

8. Discussion: Experiment 1

The results of Experiment 1 show that encountering a deaccented prenominal adjective in the speech stream did not disrupt 24-month-olds' ability to identify the noun that followed. In fact, children responded just as quickly and accurately to a familiar target noun when it was preceded by a deaccented adjective as when it was not, even when the prenominal word was an unfamiliar adjective that could potentially be mistaken for a novel noun. One

interpretation of this finding is that 24-month-olds waited for the target word before responding because they were able to make use of prosodic cues indicating that the deaccented adjective is not likely to be the final word in the sentence and thus is not an object name. In typical English sentences with broad focus such as the deaccented adjective sentences used in Experiment 1, pitch accent tends to fall on the final content word (Bolinger, 1986). When the sentence-final word is accented, the penultimate content word, in this case the prenominal adjective, is deaccented, i.e. relatively shortened in duration, lower in amplitude, and steady in pitch (e.g. Cooper et al., 1985). Thus, the lack of prosodic emphasis on prenominal adjectives in sentences in the deaccented condition of Experiment 1 could have provided young listeners with a prosodic cue that the object name was still to come.

In contrast, when prenominal adjectives in Experiment 1 were accented rather than deaccented, significant disruption occurred in response times, and when the accented adjective was unfamiliar, accuracy was reduced and more false alarm responding occurred as well. One interpretation of these findings is that when infants heard the word *the* followed by an unfamiliar accented word, they were more likely to perceive the novel word as a potential object name and thus to initiate a search for the referent. This false-alarm response was strongest in the case of accented, unfamiliar words for two reasons. First, because sentence accent in English falls on final nouns much more frequently than on prenominal adjectives, a strongly accented word following *the* is, in fact, highly likely to be a noun. Second, although children’s extensive familiarity with the adjective *pretty* might preclude a noun interpretation, the novel word *skimpy* is lexically ambiguous to them and could plausibly be a noun. Thus prosodic and lexical cues may converge to bias the child toward a noun interpretation when a novel prenominal word is accented. Moreover, prosodic and lexical cues may indicate to the child that a deaccented prenominal adjective is *not* a potential object name because the informative target word is still to come.

9. Experiment 2

The goal of Experiment 2 was to provide support for or against the interpretation that the disruption in on-line processing found in Experiment 1 reflected children’s tendency to misinterpret accented unfamiliar adjectives as potential object names. Here we assessed the nature and extent of the interference that would occur if the word preceding the target noun was, in fact, unambiguously interpretable as an object label. The prenominal adjectives used in Experiment 1 were replaced with prenominal *nouns*, giving children full lexical information about the type of word they were hearing. For example, instead of *Where’s the nice bunny?* children heard, *Where’s the ball monkey?* The prenominal nouns used in place of adjectives were familiar object names, and none of the corresponding objects was depicted in the visual stimuli. The rationale for this manipulation was that a child’s response to hearing a known object name in this position in the sentence should be similar to the effect on processing that would occur if a child misinterpreted a prenominal adjective as an object name. Thus, children in Experiment 2 encountered prenominal nouns spoken with the same prosodic characteristics of either accented or deaccented prenominal adjectives. As in Experiment 1, each of these prenominal words was followed by a familiar target noun naming one of the two pictures accessible to the child on that trial. We predicted that encountering a noun spoken with prosody typical of a prenominal adjective before the target word would cause children to respond more slowly and less accurately to the target word. This result would suggest that the reason familiar adjectives and deaccented unfamiliar adjectives did not interfere with target word recognition in Experiment 1 was because these adjectives were not easily mistaken for object names; and conversely, that delayed responses to the target words preceded by an accented unfamiliar adjective reflected children’s confusion as to whether the novel word could possibly be an object label.

10. Method

10.1. Participants

Participants were 27 24-month-olds participated in Experiment 2 (*M*=105 weeks, range 103–106 weeks). Thirteen were girls, and all participants were from predominantly English speaking families. Two additional participants were excluded from the sample because they failed to complete the first half of the trials due to fussiness.

10.2. Auditory and visual stimuli

The procedures for constructing, and presenting auditory and visual stimuli were identical to those used in Experiment 1, with four important changes. First, the prenominal adjectives used in Experiment 1 were replaced with familiar nouns (*ball*, *car*, *duck*, *flower*, *doggy*, *bunny*) preceding the sentence-final target words. It is important to note that the objects named by the prenominal nouns were never pictured, and that these words never began with the same consonant as the name for the target or the distracter picture. Second, because some of the target words from Experiment 1 were used as prenominal nouns in Experiment 2, six new target words were chosen with corresponding pictures presented in pairs (pig/monkey, cookie/apple, and bird/truck)[3]. A third change was that while accent-level was a between-subjects factor in the first experiment; it was a within-subjects factor in Experiment 2. Parallel to Experiment 1, each sentence containing a prenominal noun was recorded in two versions. In one type of trial, accent fell on the prenominal noun (e.g. *Where's the BALL monkey?*) while in the other; the prenominal noun was deaccented (e.g. *Where's the ball MONKEY?*). The prenominal nouns in the accented condition of Experiment 2 were comparable in duration (*M*=480 ms) and in pitch contour to the accented adjectives in Experiment 1, and in the deaccented condition, the prenominal nouns were comparable in duration (*M*=380 ms) and pitch contour to the deaccented prenominal adjectives of Experiment 1. As in the previous experiment, there was no prenominal word preceding the target noun in the control condition (e.g. *Where's the MONKEY?*). The final change was to reduce the number of critical test trials in each condition from eight to six and to insert additional filler trials. This was done to reduce the overall proportion of anomalous trials, since they might be particularly difficult or frustrating for participants.

10.3. Procedure and apparatus

The procedure and apparatus were identical to those in Experiment 1.

10.4. Coding

Coding procedures were identical to those in Experiment 1.

10.5. Measures of speech processing efficiency

Accuracy, reaction time, and shift tendency were calculated exactly as in Experiment 1.

11. Results

The time course of children's orienting to the named target picture in each of the three prenominal noun conditions is shown in Fig. 2. As in Experiment 1, 24-month-olds rapidly

[3]The target words from Experiment 1 were used as prenominal nouns in Experiment 2 because they are among the earliest words produced by English-learning children and thus would be unambiguously identifiable as object words, consistent with the logic of this control experiment. The target words used in Experiment 2 were also well-known words produced by all the participant children, although perhaps learned somewhat later than those in Experiment 1. If for this reason they were somewhat more difficult to understand, this would affect performance in all three conditions in Experiment 2 and could not account for the differences found among the conditions.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

and reliably identified the correct picture in response to control sentences with no prenominal word. However, unlike in Experiment 1 where the presence of a prenominal adjective had no effect on children's success in identifying the target noun that followed in most conditions, performance in both prenominal noun conditions of Experiment 2 was substantially less efficient than performance on control sentences.

11.1. Accuracy

The mean proportions of looking time to the target picture during the 333–1800 ms window after target noun onset were calculated for each child in each of the three conditions. When these accuracy scores were entered into a one-way ANOVA, the main effect of prenominal word type (accented noun, deaccented noun, or none) was significant, $F(2, 52)=10.40$, $P<0.001$. Planned comparisons revealed that 24-month-olds were significantly less accurate in identifying target words preceded either by an accented prenominal noun (M=53.8%, SD=18.6%), $t(26)=4.29$, $P<0.001$, or by a deaccented prenominal noun (M=57.4%, SD=14.0%), $t(26)=4.31$, $P<0.001$, as compared to the control or 'none' trials (M=70.6%, SD=11.9%). That is, when children heard a familiar noun preceding the target word, they were less likely overall to look at the picture labeled by the target noun than when no prenominal word was present, showing that accuracy was indeed disrupted by the prenominal noun.

11.2. Reaction time

Although the goal was to compare mean reaction time scores across conditions as in Experiment 1, there were unfortunately too few children with stable reaction time means in all three conditions of Experiment 2 for two reasons. First, the total number of trials with prenominal words was reduced from eight to six for each sentence type, as explained above. A second, and less predictable constraint resulted from the fact that the reaction time measure is based solely on distracter-initial trials, when the child must shift to the target picture to respond correctly. Because the number of distracter-initial trials depends on where the child happens to be looking at the target-word onset, it varied from 0 to 6 trials per condition in this experiment, with a mean of 2.7. With so few response latencies per child for each trial type, and the necessity of having latencies in all three conditions in order to include a child's mean in any condition, it was not possible to make meaningful comparisons of reaction times between conditions.

11.3. Shift tendency

In the next analysis we asked how likely children were to show a false alarm response by shifting gaze from one picture to the other as they encountered a prenominal noun. The proportion of trials on which shifts occurred prematurely in response to deaccented and accented prenominal nouns was calculated for each child. When encountering an accented prenominal noun before the target word, children shifted prematurely significantly more often (M=60.7%, SD=30.4%) than in response to a deaccented prenominal noun (M=43.6%, SD=27.2%), $t(31)=-2.17$, $P<0.05$.

Were 24-month-olds less likely to shift prematurely when hearing a prenominal adjective before the target word than when hearing a prenominal noun? A second analysis compared shift tendency data across the first two experiments. In Experiment 1, when children encountered accented adjectives preceding the target word, the rate of false alarm shifting during the prenominal word (M=35.1%, SD=22.8%) was significantly lower overall than it was for the children in Experiment 2 (M=52.1%, SD=29.9%), $t(180)=-4.19$, $P<0.001$. Thus, behavior indicating false alarms was less common when the prenominal word was an adjective than when the prenominal word was clearly interpretable as a noun.

12. Discussion: Experiment 2

Experiment 2 was designed as a control study to provide a point of comparison for understanding the findings of the previous experiment. In Experiment 1, we found that when listening to a sentence such as *Where's the glib BUNNY?* 24-month-olds apparently disregarded the prenominal adjective as they quickly and accurately identified the target noun that followed. That is, they avoided a false alarm response of shifting gaze prematurely during the adjective, and the potentially disruptive effect of misinterpretation. Because they typically waited to hear the target word before shifting gaze to the target picture, children responded as if they could tell that the target word was still upcoming, and thus that the prenominal adjective was not a potential noun. One concern with this interpretation is that children might not have had enough time to respond with a false alarm in the deaccented adjective conditions, given that these words were on average less than 400 ms in duration. In Experiment 2, the prenominal adjectives used in Experiment 1 were replaced with prenominal nouns matched in duration to detect whether children had time to respond to a very short content word preceding the target noun. This manipulation also enabled us to assess the nature and extent of disruption that would occur if the child did indeed interpret the prenominal word as an object label. An additional prediction was that accented prenominal nouns would lead to even greater disruption than deaccented prenominal nouns by providing a misleading prosodic cue typically associated with sentence-final content words.

When 24-month-olds in Experiment 2 heard a familiar noun in prenominal position, recognition of the subsequent target noun was significantly disrupted. Even though the object denoted by the prenominal noun was not pictured, children were more likely to shift gaze as it was heard. This pattern of responding is consistent with the Swingley and Fernald (2002) finding that 24-month-olds will search for a matching picture when they hear a familiar or unfamiliar object name in sentence-final position even when no matching picture is available. Thus, the findings of Experiment 2 show that there was indeed enough time for the child to make a nominal interpretation and response during the prenominal word. Therefore, the efficiency of target word recognition in Experiment 1 along with the relatively low level of premature shifting during prenominal adjectives cannot simply be attributed to time constraints.

These results also provide us with a model of how children might respond if they did interpret an unknown prenominal adjective as an object name, because the prenominal words in Experiment 2 were, in fact, familiar nouns. Two-year-olds appear to be able to listen through words that they can in some way identify as *not* nouns, avoiding misinterpretation that would interfere with successful processing of subsequent target nouns.

Moreover, parallel to the findings in Experiment 1, accented nouns were more disruptive than deaccented nouns in prenominal position. In both the deaccented and accented conditions, children had unambiguous lexical information that the prenominal word was an object name. However, in the deaccented condition, when the prenominal noun was spoken with an intonation pattern typically used with a deaccented adjective, this prosodic information contradicted the lexical cue supporting an object-label interpretation. In this case, children were somewhat less likely to false alarm by shifting prematurely. In contrast, in the accented condition, the prosodic emphasis on the prenominal noun was more consistent with the possibility that it was a sentence-final noun, rather than the penultimate word in a sentence. Thus, prosodic cues converged with lexical cues to suggest that the prenominal noun could potentially be a name for one of the pictures.

13. Experiment 3

In Experiments 1 and 2, when accent was placed on the prenominal word, 24-month-olds frequently shifted gaze prematurely before hearing the crucial target word that would enable them to select the appropriate picture. In contrast, when children encountered deaccented adjectives in Experiment 1, they were more likely to wait for the target noun before initiating a shift, as if anticipating that the object label was still to come. The deaccented prenominal adjectives in Experiment 1 had the prosodic characteristics of unfocused penultimate words, including pitch continuity and relatively shorter duration, while the accented versions had prosodic characteristics more closely resembling final nouns. In principle, the young listeners in this study could have relied on such prosodic cues to anticipate correctly that the deaccented adjectives in Experiment 1 would be followed by another more informative word at the end of the sentence.

The next experiment was designed to assess the plausibility of this prosody-based explanation for children's success in listening through prenominal adjectives. In Experiment 3, we asked whether adults are able to rely exclusively on prosodic cues to distinguish penultimate from sentence-final words in a listening context analogous to that encountered by the children in Experiments 1 and 2. Research on resolution of closure ambiguities shows that adults can use intonation to determine whether an ambiguous word is sentence-final or whether the sentence will continue beyond it. For example, Grosjean and Hirt (1996) investigated responses to sentences such as *Earlier my sister took a dip* vs. *Earlier my sister took a dip in the pool*, which were gated and presented incrementally in successively longer segments. As adults heard a target word (e.g. *dip*), they judged whether 0, 3, or 6 more words would follow, relying on intonation to predict whether or not the target word was the final word in the sentence. In our third experiment, we asked whether prosodic cues could also be used to distinguish between homophones acting as either sentence-final nouns or penultimate adjectives. Adults' success in using prosody to make such a distinction would support the hypothesis that young listeners are also able to use prosodic information when encountering ambiguous prenominal adjectives in continuous speech.

In Experiment 3, we used a version of the gating paradigm established by Grosjean and Hirt (1996) to incrementally present sentences containing adjective/noun ambiguities to adult listeners. To create such ambiguity we used English homophones that commonly function as both nouns and adjectives. In the noun-target version of each homophone pair, the target word functioned as a noun in final position (e.g. *The little boy had a* COLD). In the adjective-target version, the homophonous target word functioned as either a deaccented or accented prenominal adjective (e.g. *The little boy had a* cold *NOSE*, or *The little boy had a* COLD *nose*). For each stimulus sentence, markers were placed at key points within the potentially ambiguous target homophone. Sentences were then truncated at each marker to create four increasingly informative sentence segments that were presented sequentially to participants. As listeners heard the sentence unfold in successive steps, they were asked at each step to predict how the sentence would continue. The question of interest was at what point adults would be able to identify the target homonym unambiguously as either a prenominal adjective or a sentence-final noun.

14. Method

14.1. Participants

Participants were 24 university students from 18 to 26 years of age who were paid $10 for their participation. All participants were native speakers of American English.

14.2. Speech stimuli

We first identified 28 pairs of homophonous target words that could function either as adjectives or nouns, such as *real/reel* and *pale/pail*. An effort was made to construct carrier sentences that were not strongly biased toward either the adjective or the noun interpretation of the target word. The following are three example pairs (see Appendix for the complete set):

1. **a.** The little boy had a cold

   b. The little boy had a cold nose

2. **a.** She hit him on the back

   b. She hit him on the back porch

3. **a.** The fisherman used a reel

   b. The fisherman used a real worm

Speech stimuli were recorded using Peak™ LE 2.62 sound recording and editing software by a female native speaker of American English. Three tokens of each noun-target sentence were recorded. For the adjective-target sentence in each pair, three tokens were recorded in which the target adjective was accented, and three tokens were recorded in which the sentence-final noun was accented with the preceding target adjective deaccented. To evaluate the naturalness of these spoken sentences, 18 adult volunteers listened to randomly ordered subsets of the stimuli, rating each token on a 4-point scale ranging from 0 (quite unnatural) to 3 (quite natural). The token of each sentence rated as most natural was chosen for inclusion in the final stimulus set.

The target homophone was always a monosyllabic consonant-vowel-consonant word. For each sentence, a marker was placed at four phonetically defined points within the target word: (1) after the initial consonant (e.g. *The farmer's wife had a* p|); (2) after the vowel and before the final consonant (e.g. *The farmer's wife had a* peI|); (3) after the final consonant, either before the pause if the target was a noun or before the final word if the target was an adjective (e.g. *The farmer's wife had a* peIl|); (4) at the end of the utterance, ending either with silence (e.g. *The farmer's wife had a pail*), or with the final noun (e.g. *The farmer's wife had a pale face*). Sentences were then truncated at each marker to produce four increasingly informative segments for each sentence. When the target word was an adjective, cuts were made conservatively early to reduce the likelihood of coarticulation effects.

Sentences were arranged in two sets of four lists each. In one set A, A′, B, B′ all prenominal adjectives were deaccented, and in the second set C, C′, D, D′ all prenominal adjectives were accented. List A included 28 sentences arranged in quasi-random order. Half of the sentences contained a noun target, while the other half contained a deaccented adjective target. It was never the case that both members of a particular homophone pair were included in the same list. List A′ was complementary to List A, comprising sentences with the other member of each homophone pair presented in the same order. Lists B and B′ were constructed by reversing the order of lists A and A′ respectively. Lists C, C′, D, and D′ followed the same pattern using accented versions of the adjective targets. Thus, for all participants, half the stimulus sentences had a sentence-final noun as the target word, while the other half had a prenominal adjective as the target word. Half the participants heard sentences in which the adjective target words were deaccented, while for the other half the adjectives were accented.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

14.3. Procedure

The speech stimuli were played from an iMac™ laptop computer over high-quality headphones. Each subject listened to one list of 28 sentences with each sentence unfolding in four successive segments. Participants were given answer sheets on which the 28 sentence pairs were printed in the order in which they would be heard. That is, although each participant heard only one sentence for each homophone pair, both alternatives were presented as options, and the listener was asked to choose between them. After hearing each segment, participants judged which of the two options was most likely the sentence that they were hearing by marking their answer sheet. Asking participants to choose between two options reduced possible frequency-based biases of one interpretation of the homophone over the other, because it was made clear from the start that there was a 50% chance the target word would be used either as an adjective or a noun.

14.4. Accuracy measurement

For each participant, mean accuracy scores were calculated at each of the four segments for each noun-target and adjective-target stimulus sentence.

15. Results

Because subjects were at ceiling after the fourth segment when they had heard the complete stimulus sentence, accuracy scores for this segment were not included in any statistical analysis. Mean accuracy scores were entered into a 3 (segment)×2 (homophone type)×2 (accent level) ANOVA with segment and homophone type as within-subjects factors, and accent level as a between-subjects factor. There was no main effect of accent level and no significant interaction of accent level with segment or homophone type. Regardless of whether the ambiguous target word was accented or deaccented, adults were equally successful in identifying the correct part of speech of the ambiguous homophone.

The omnibus test did, however, reveal main effects of segment, $F(2, 114)=87.90$, $P< 0.001$, and of target word type, $F(1, 114)=26.50$, $P<0.001$, and also a significant interaction between the two, $F(2, 114)=10.60$, $P<0.001$. Follow-up tests revealed that the effect of segment was significant regardless of whether the target word was a noun or an adjective, $F_{adj}(2, 40)=64.85$, $P<0.001$; $F_{noun}(2, 40)=33.32$, $P<0.001$. Thus, adult listeners were increasingly correct at choosing the part of speech of the ambiguous homophone as they heard more informative segments. However, the interaction of homophone type and segment reflected a different in accuracy for nouns and adjectives in the first segment, $F(1, 20)=20.30$, $P<0.001$, although not in the second or third. In other words, after hearing the first segment, participants were significantly more correct in identifying the homophone when it was a noun than when it was an adjective. By the second segment, however, listeners' accuracy was no longer significantly different between adjective and noun versions of the homophonous target words.

In the next analysis, we asked whether accuracy in disambiguating the target word was beyond what we would expect if participants were merely guessing. Because no differences in accuracy were found between accented and deaccented adjective homophones, we collapsed the data across accent level in the following analyses. As shown in Fig. 3, mean accuracies were compared to chance (50%) for each segment in each target type. At the first segment, noun-targets were correctly identified significantly more often than predicted by chance, $t(20)=7.66$, $P<0.001$, although adjective-targets were not, $t(20)=-1.05$, $P>0.10$. By the second segment, after subjects had heard the vowel of the target word, the proportion of correct responses was significantly above chance for both types of target words, $t_{adj}(20)=9.74$, $P<0.001$; $t_{noun}(20)=17.40$, $P< 0.001$, and accuracy increased still further at the

third segment at the end of the vowel, $t_{adj}(20)=11.70$, $P<001$; $t_{noun}(20)=20.80$, $P<0.001$. Thus, adults were above chance in correctly identifying ambiguous homophones as nouns after hearing only the first consonant, while adjectives were correctly identified by the end of the vowel.

16. Discussion: Experiment 3

The main finding of Experiment 3 was that adults were highly successful in identifying the form class of ambiguous homophones functioning either as prenominal adjectives or as sentence-final nouns after hearing only 200–300 ms of the word. Accuracy was above chance even earlier for noun homophones than for adjectives, consistent with results from the closure ambiguity study of Grosjean and Hirt (1996) in which listeners could more quickly determine when a target was indeed the final word in the utterance than when it was not.

Why was disambiguation asymmetrically faster when the target word was in final position? This may be because the emphasis placed on a sentence-final word is more distinctive than the relative de-emphasis of a penultimate word. The accent on a word may simply be easier to identify, and thus listeners may be quicker to identify a word as a final noun than a prenominal adjective. Additionally, prosodic information in the preceding context could cue a listener that an utterance is almost complete before an accented target word even begins. Cutler (1976) showed that adult listeners responded more quickly to a word spliced into the end of a sentence that had originally preceded an accented rather than an unaccented word, indicating that the intonation of the words just before an accented target word could provide predictive information. However, preceding context may not be as helpful in identifying a deaccented penultimate word. Such information distinguishing deaccented penultimate words from accented final words is relevant to our interpretation of how 24-month-olds respond when hearing deaccented prenominal adjectives such as those in Experiment 1. If children, like adults, can use prosodic de-emphasis as a cue that another word is upcoming, then young listeners could effectively avoid misinterpreting a prenominal adjective as the final noun of an utterance.

Another important finding from Experiment 3 is that even when prenominal adjectives were accented, making them potentially harder to disambiguate from nouns, adults still recognized them as adjectives well before the onset of the following noun. The 24-month-olds in the first two experiments responded more slowly to target nouns preceded by accented as compared to deaccented prenominal adjectives, but accent level did not affect adults' performance in Experiment 3. Adults' resistance to misinterpretation of accented prenominal adjectives may indicate a more sensitive use of prosodic patterns that distinguish an accented penultimate adjective from a final noun. However, it is also important to note that our measure of adults' performance in Experiment 3 did not capture real-time processing in the way that we monitored children's comprehension in Experiments 1 and 2. A more fine-grained measure might have revealed that even adults have more subtle difficulty with accented than deaccented prenominal adjectives.

16.1. Acoustic features that affect prosodic cues

What specific acoustic properties could enable listeners to make such efficient use of prosodic cues in this disambiguation task? To explore this question, we analyzed the stimuli in Experiment 3 for differences in duration and pitch contour, the two most common acoustic correlates of accent described in the literature (e.g. Cooper et al., 1985; Crystal & House, 1988). These analyses revealed an acoustic basis for the perceived differences we observed between homophones serving as final nouns and as prenominal adjectives, both accented and deaccented. Fig. 4 depicts one example of a homophonous word in all three

NIH-PA Author Manuscript

NIH-PA Author Manuscript

sentence types, indicating the different acoustic patterns of the target word used in each context. When a homophonous word was used as a noun it had the longest duration (*M*=489 ms; SD=76 ms), followed by the accented adjective form (*M*=350 ms; SD=61 ms), and deaccented adjective form (*M*=286 ms; SD=58 ms). Across the 28 sets of stimulus sentences, 27 followed this same pattern. A one-way ANOVA showed a significant difference among these target word durations $F(2, 81)=70$, $P<0.001$, and pair-wise tests confirmed that each homophone type differed in duration from the others, $t(27)_{\text{accented/noun}}=7.40$, $P<0.001$; $t(27)_{\text{deaccented/accented}}=11.00$, $P<0.001$. Thus, when a homophone such as *cold* was deaccented, it was significantly shorter than when it functioned as an accented adjective, and it was longest in duration when used as a sentence-final noun. As predicted based on our findings with young listeners in Experiments 1 and 2, the durations of accented adjectives fell between those of deaccented adjectives and nouns, making them more similar to nouns acoustically than were their deaccented counterparts.

Complementing the acoustic analyses, each stimulus sentence was also transcribed using the ToBI system (Beckman & Elam, 1997; Silverman et al., 1992) to describe the presence and types of pitch accents. As shown in Fig. 4, homophones used as final nouns were characterized by a consistent pitch accent, labeled as a standard H* followed by a low boundary tone (L–L%). When a homophone was used as an accented prenominal adjective, a pitch accent was also consistently present, most often described as an L+H* (commonly indicating contrastive stress). A low boundary tone (L–L%) marked the noun that followed. Finally, when the homophone was a deaccented prenominal adjective, pitch tended to stay level. In this case, no pitch accent occurred until the following noun, which most often received an H* followed by low boundary tones. Of the 28 stimulus sets, 24 clearly followed this pattern. In general, there was a gradation from deaccented adjective to noun in how prominently the target homophone was marked by the acoustic features of vowel lengthening and pitch movement. Here too accented adjectives were phonetically more similar to final nouns than were deaccented adjectives because pitch accent was present to some extent in both cases. However, while the homophone in its accented adjective form was acoustically more noun-like, adults were not misled to interpret it as a noun as children seemed to be, perhaps because of a finer sensitivity to the different *type* of pitch accent in each case.

The results from Experiment 3 demonstrate that discriminating between prenominal adjectives and final nouns is a simple task for adults, even when the preceding linguistic context is identical in both cases and the grammatical form class of the word is ambiguous. In this study, the only basis for discrimination up to the end of the sentence was characteristic acoustic differences between words used in penultimate vs. final position. Since prosodic information was clearly guiding adults' success in this task, it seems plausible that the 2-year-olds in Experiment 1 could also have relied on prosodic cues as they listened through prenominal adjectives rather than interpreting them as nouns.

17. General discussion

Here we investigated how young language learners as well as experienced adults respond in real-time to potentially ambiguous prenominal adjectives. These studies demonstrate that despite syntactic and lexical indeterminacies that could bias a listener towards mistaking a prenominal adjective for a noun, both children and adults were able to avoid disruption in speech comprehension by using prosodic cues in combination with distributional information and lexical knowledge. In Experiment 1, 24-month-olds listened through deaccented adjectives, identifying the subsequent target noun just as quickly and reliably as they did when no adjective was present, even if the prenominal word was unfamiliar. When an unfamiliar adjective was accented, however, there was some tendency for children to

shift prematurely during the prenominal word, as if they inferred that it might be an object name. One interpretation of these findings is that while children were not misled by deaccented adjectives, *accented unfamiliar* adjectives were more likely to bias the young listener toward interpreting the novel word as a label. Such false-alarm responses had a disruptive effect on on-line processing, resulting in slower and less reliable recognition of the target noun. In Experiment 2, prenominal adjectives were replaced with familiar nouns to assess the extent of disruption that would occur if a child unambiguously interpreted a prenominal word as an object label. Prenominal nouns substantially disrupted responses to the target nouns that followed, as might be expected if the children in Experiment 1 had mistakenly responded to prenominal adjectives as potential nouns. Experiment 2 also demonstrated the influence of prosodic cues: When prenominal nouns were not only lexically familiar as object names but were also accented to sound more like final nouns, young listeners were most disrupted. Finally, Experiment 3 explored an adult analogue of the indeterminacy problem faced by 2-year-olds, investigating how adults could resolve the form class ambiguities of homophonous words presented in string-identical sentences. Adults were able to use prosodic information alone to discriminate between adjective/noun homophones before the target word had been entirely heard. This finding lends further support to the hypothesis that prosodic cues can be used to improve the efficiency of processing prenominal words, even when efficiency means listening through certain words in order to respond quickly and accurately to what comes next in the speech stream.

We argue here that our results illuminate the efficiency with which children and adults can process ambiguous words in speech when distributional cues are ambiguous but prosody is informative. However, it is important first to consider alternative interpretations of the finding that for the children in Experiment 1, accented prenominal adjectives caused greater disruption than deaccented adjectives. Because accented adjectives were on average 100 ms longer than their deaccented counterparts, one possibility is that children simply had more time to shift during an accented prenominal word and thus showed a greater disruptive effect in that condition. This concern was addressed in Experiment 2; when prenominal adjectives were replaced by prenominal nouns matched in duration, children were as likely to shift prematurely during the shorter deaccented nouns as during the longer accented nouns. Thus, when children listened through the shorter adjectives in Experiment 1, it was not because they did not have sufficient time to shift.

Another alternative explanation for why children's word recognition was more disrupted by some prenominal words than others is that acoustically de-emphasized words might have been less likely to trigger lexical processing altogether[4]. By this account, it was not the case that 24-month-olds listened through deaccented adjectives by actively using prosody as a source of information, but rather that unfamiliar deaccented words were simply not processed and thus were passively ignored much of the time. This explanation seems unlikely given abundant evidence that word-like units are salient to infants long before they are processed as lexical items. Infants in the first year already identify word-like clusters of syllables in fluent speech and represent them at some level, even in strings of nonsense sounds with minimal prosodic variation (Saffran, Aslin, & Newport, 1996). Swingley (2005) makes the case that words must be represented in some form even earlier, before infants can begin to show language-specific prosodic parsing biases around the age of 8 months. Thus it is implausible that two-year-olds would completely ignore a well-formed content word in a communicatively relevant sentence just because it did not receive prosodic emphasis. Moreover, although English-speaking adults respond more quickly to prosodically

[4]Thanks to an anonymous reviewer for suggesting this possibility.

emphasized words, Cutler and Swinney (1987) found that children up to 5 years of age did not respond differently to accented and deaccented words in a word-monitoring task.

A third alternative explanation for why children's word recognition was more disrupted in the accented adjective conditions in both Experiments 1 and 2 rests on a pragmatic argument. For adults, stress on a prenominal adjective in English often indicates that it is being used contrastively. For example, the request to *Hand me the BLUE book*, implies that a book of another color is also present with which the blue book is being contrasted. Although the accented adjectives used here were uninformative with no contrast implied, this pragmatic anomaly could potentially have been confusing. If so, was it this implicit contrast that motivated children to shift more often? This explanation seems unlikely given several studies showing that children this age are too young to make such use of stress for semantic or pragmatic understanding. Cruttenden (1974) found that the ability to use stress to infer semantic contrast does not fully develop until between 7 and 10 years of age. Holdgrafer and Campbell (1986) found a similar developmental delay in the ability to infer that accent marks new rather than given information, and a recent study by Vogel and Raimy (2002) showed that not until age 11 can children use stress patterns to distinguish between pairs of words used as adjective-noun phrases vs. adjective-noun compounds (e.g. *hot dog* vs. *hotdog*). Thus, it is unlikely that the 2-year-olds in our studies appreciated the pragmatic implications of accenting prenominal adjectives and were disrupted in identifying the target word because they somehow expected a property-based contrast.

A final concern is that the prenominal adjectives we have focused on here may not be broadly representative of the syntactic frames in which young children are actually exposed to adjectives. To address this question, we examined adjective use by five parents interacting with their 1- to 3-year-old children, using transcripts from the following sessions in the CHILDES database (MacWhinney, 2000): Adam 01–20, Eve 01–20, and Sarah 003–040 from the Brown corpora, Peter 01–20 from the Bloom 1977 corpus, and Shem 01–35 from the Clark corpus. This sample comprised a total of 89,857 child-directed utterances. Based on the MacArthur-Bates Communicative Development Inventory norms (Dale & Fenson, 1996), we focused on 12 adjectives that English-learning children typically comprehend early and produce before the age of two according to maternal report: *big, broken, blue, careful, cold, dirty, good, hot, nice, pretty, red,* and *wet*. The question of interest was how often overall parents used these descriptive words, and what proportion of the time each of the 12 adjectives occurred in a prenominal frame (e.g. *See the blue block*) as opposed to a predicate frame (e.g. *This block is blue*). Preliminary analysis revealed that one of the words in this set, *careful*, was almost always spoken either in isolation or in the frame *Be careful!* Because *careful* in these contexts functioned more like an adverb than an adjective, this word was not included in the subsequent analyses.

Parents used prenominal frames in 52% of the 3,067 occurrences of these common adjectives in speech to children. However, the tendency for adults to use either a prenominal or predicate frame was not consistent across the eleven adjectives analyzed. Table 3 reports the mean proportion of occurrences of each adjective in each syntactic frame. We found that *big, blue,* and *red* were most often used attributively in prenominal position, while *hot, cold, wet*, and *broken* occurred primarily as predicates in sentence-final position. The remaining four adjectives—*pretty, good, nice*, and *dirty*—were more or less equally likely to occur in either frame. It is interesting to note that this pattern of results confirms observations made by Nelson (1976) in one of the earliest studies of adjective use in speech to children. Nelson concluded that adjectives like *big* and *blue* that are most likely to appear in prenominal position typically describe more inherent, permanent properties of objects and function to subcategorize the nouns they modify, while adjectives like *broken* and *hot* that occur more often in predicate frames tend to denote temporary or unusual changes of state (see also

Prasada & Cummins, 2000). The correlation of semantic factors to some extent with choice of syntactic frame may provide useful information about adjectives as children learn about their meanings and grammatical functions. However, the main point of this transcript analysis was to demonstrate that when young children hear adjectives in speech directed to them, about half the time overall these words occur in prenominal frames. Thus the problem of interpreting an unfamiliar adjective preceding a familiar noun is indeed a challenge regularly confronted by the 2-year-olds who were the focus of this research

17.1. Prosodic cues in sentence processing

Prosodic cues facilitate spoken language understanding in many ways. Cutler and Foss (1977) found that adults recognize and respond faster to accented vs. deaccented words in speech regardless of their syntactic function. Conversely, incorrect lexical stress can disrupt processing for adults (e.g. Birch & Clifton, 2002; Buxton, 1983). In addition, several studies show that the prosodic contour of the entire phrase or utterance can affect listeners' processing of individual words (e.g. Cutler, 1976; Grosjean & Hirt, 1996), and thus prosody can be used to some degree to anticipate how a sentence may continue, consistent with our findings with both adults and children.

For adults, the effect of prosody on speech comprehension is also tied to semantic and pragmatic understanding. Because speakers use accent to convey focus, adult listeners interpret this intonational cue in on-line comprehension as indicative of new or contrastive information. In an eye-tracking study, Dahan, Tanenhaus, and Chambers (2002) found that adults' speed of reference resolution was influenced by accenting or deaccenting initial phonemes of phonetically overlapping words (e.g. *bed* or *bell*) based on which word had already been referenced. If a participant had just heard, *Put the bed below the triangle,* and then, *Now, put the BE*… in an accented condition vs. *Now, put the be*… without accent on the target word, participants were more likely to look to the bell, demonstrating anticipation that the accented word would mark *new* information instead of a referent already mentioned. Snedeker and Trueswell (2003) also found that adult listeners can use prosodic information on-line to resolve syntactic ambiguities in sentences such as *Tickle the frog with the feather* in which the prepositional phrase *with the feather* can be interpreted as a modifier of the verb as in *Tickle the frog by using the feather*, or as a modifier of the noun *frog* as in *Tickle the frog that has the feather*. Listeners were able to resolve ambiguity in this situation using pause-break information that prosodically parsed the sentences as in *[Tickle][the frog with the feather]* vs. *[Tickle the frog][with the feather]*, corresponding to each of the two possible interpretations.

While the adult listener can use prosody to resolve temporary ambiguities in real-time language processing, less is known about how young children integrate prosodic cues in comprehension. On the one hand, language-specific intonation patterns help 14-months-old infants in the initial stages of parsing speech into units that correspond to words and phrases (Cutler, Dahan, & van Donselaar, 1997; Jusczyk, Cutler, & Redanz, 1993; Jusczyk & Kemler Nelson, 1996). However, other aspects of prosodic processing that might aid comprehension are much slower to develop. Although adults recognize stressed words more rapidly than unstressed words (Cutler & Foss, 1977), children do not show this advantage until after 6 years of age (Cutler & Swinney, 1987), nor are children as sensitive as adults to the use of accent to indicate new vs. given information or prosodic contrast (e.g. Cruttenden, 1985). The present findings reveal a new dimension of children's sensitivity to prosodic information in sentence comprehension. The 2-year-olds in Experiment 1 were able to take advantage of prosodic cues signaling that deaccented prenominal adjectives in continuous speech should *not* be considered as novel noun object labels. The task of the child in this situation was not to make any inference about whether accented or deaccented words conveyed new vs. given information, but simply to do what adults seem to do when

NIH-PA Author Manuscript

NIH-PA Author Manuscript

encountering closure ambiguities (e.g. Grosjean & Hirt, 1996). If as children heard the prenominal adjective, they were able to discern from the prosody that there was still more speech to come, then they could wait for the target noun to be spoken and thus avoid making a premature response. In the majority of phrases containing non-contrastive prenominal adjectives, accent falls on the final noun, while the adjective tends to be deaccented in comparison. This prosodic difference between prenominal adjectives and final nouns, in conjunction with information about the relative position of these words within a phrase, could provide valuable cues guiding children's online processing in situations of potential ambiguity.

17.2. The use of multiple cues in early language processing

Recent research on speech processing based on computational models demonstrates that the simultaneous use of multiple cues based on morphology, stress and boundary information can benefit the listener more than any single source of information (Christiansen et al., 1998; Seidenberg & MacDonald, 2001). Our findings indicate that even very young language learners are able to integrate multiple cues in interpreting ambiguities in spoken language. Children's tendency to respond to a prenominal word as if it were potentially the name for a target object was influenced by at least three kinds of information: word position, lexical familiarity, and prosodic pattern. Knowledge of the distributional regularities associated with prenominal adjectives—i.e. that a word preceded by a determiner such as *the* and immediately followed by an object name is unlikely to be an object name as well—is in principle sufficient to enable the child to avoid misinterpretation. That is, in the situation confronting children in Experiments 1 and 2, the prenominal word was only briefly ambiguous, given that a target noun matching one of the pictures appeared at the end of every stimulus sentence. Thus, if children on every trial had simply *waited* to see if the ambiguous word was followed by a familiar object name before searching for the matching object, their performance would have been equally good in all conditions. However, the distributional cue of word position was not the only factor guiding children's on-line interpretation when encountering an ambiguous word, and errors in some conditions were disruptive. This was presumably because 2-year-olds, like adults, have a countervailing tendency to place their bets early, responding as soon as it seems likely that they have sufficient information to make the correct choice. We propose that it was the interplay of two additional factors—lexical familiarity and prosodic pattern—that accounted for these differences in performance across the six conditions in Experiments 1 and 2, as both cues converged to different degrees in support of a noun interpretation for the prenominal word.

To illustrate this point, Fig. 5 combines results from Experiments 1 and 2, showing overall accuracy in each condition. The ordering of the six conditions reflects the extent to which the prenominal word disrupted efficient processing of the final target noun, with increasing disruption from left to right. This left-to-right ordering also reflects the extent to which the prenominal words used in each condition were prosodically and lexically consistent with a noun interpretation. Accuracy in target word recognition was highest in the leftmost condition. Here the target word was preceded by a common adjective such as *pretty*, known to all the children and thus unlikely to be interpreted as an object name. This familiar adjective was also deaccented, a prosodic pattern uncharacteristic of final nouns. Thus children's peak performance in this condition could be accounted for by the fact that neither lexical nor prosodic characteristics of the prenominal word were likely to bias the child toward misinterpreting the adjective as a potential object word. At the other extreme, the bar on the far right shows that children experienced the greatest disruption when the final target noun was preceded by an accented noun in Experiment 2. In this case, because the prenominal word was a known object name emphasized by a pitch accent, both lexical and prosodic cues biased children to mistake the penultimate word for the final target noun,

resulting in frequent false-alarm responses. Moving from left to right across the six conditions in Experiments 1 and 2, children's success in reliably identifying the final target noun declined as they encountered prenominal words that were progressively more similar to nouns and thus were increasingly likely to lead to premature shifting at the expense of overall accuracy. We would not have expected such gradations in performance if children were not taking both prosodic and lexical cues into account in evaluating the status of the prenominal word as it unfolded. These results suggest that even very young language learners are capable of integrating multiple sources of linguistic and contextual information as they interpret continuous speech.

17.3. The development of efficiency in sentence processing, or learning to 'listen through'

Children's tendency to hold off in trying to identify a referent object when encountering an uninformative and unstressed prenominal adjective turns out to be an efficient strategy for interpreting a sentence with the target noun in final position. However, this is not the aspect of efficiency that is typically the focus of research on speech processing by adults, where the ability to respond rapidly is emphasized more than the ability to postpone a response when appropriate. For example, Allopenna et al. (1998) found that subjects' saccade latencies to the appropriate object were on average 100 ms faster when the target object did not have a competitor in the display sharing the same sound. This type of finding is usually interpreted in terms of how quickly a response occurs relative to the point of disambiguation. Similarly, studies of referential ambiguity highlight how rapidly adults make use of semantic and contextual information as they interpret spoken sentences incrementally (Eberhard et al., 1995; Sedivy et al., 1999). However, when adults respond more rapidly to *candy* or *plain red square* in unambiguous contexts than in situations where more than one object is initially consistent with the unfolding word or phrase, they are actually demonstrating *two* types of efficiency. Not only are they using disambiguating information as early as possible, they are also efficiently postponing the search for a referent until they have the information necessary. Thus studies of real-time speech processing show that adults not only know when to respond, but also when *not* to respond.

For children as well, learning when they should not respond is as important to success in comprehension as learning when they should respond. For example, children also show restraint when faced with competitors that share the same initial sound as the target word. Swingley et al. (1999) showed that 24-month-olds were slower to respond to *dog* when looking at pictures of a dog and a doll than when looking at pictures of a dog and a tree. Thus in a situation of lexical ambiguity, children delayed their search for the appropriate referent until they had heard enough of the target word to choose correctly. In the current study, our findings show that even when encountering unfamiliar words, young children are developing this kind of efficiency by using prosodic patterns as cues indicating when to respond and when to wait.

It is clear that both of these aspects of efficiency in interpreting continuous speech—the ability to respond very rapidly when appropriate, as well as the ability to hold back in cases of ambiguity—could be of considerable importance to the young language learner. Even for experienced adults, understanding rapidly spoken sentences depends on incremental processing from moment to moment, rather than waiting until the end of a sentence to make sense of what was said (e.g. Marslen-Wilson & Tyler, 1980; Tanenhaus et al., 1995). The inexperienced infant just beginning to learn language faces considerably greater challenges in interpreting speech in which so many words are unfamiliar. For example, the child needs to pay attention to each word as it comes along and to remember and relate non-adjacent words in the sequence, a skill essential for gradually mastering the long-term dependencies that characterize grammatical relations. Thus children who can recognize and interpret known words quickly and accurately are likely to have more cognitive resources available

for attending to other less familiar words in the sequence. Such efficiency could lead to greater success in learning new words as well as in keeping track of distributional information about structural relations among the words. In contrast, the child who is slower to respond and more prone to false alarms when encountering unfamiliar words will more often make interpretative errors that have to be corrected. Although such errors may only involve a 200–300 ms delay, disruptions of this sort could be quite costly, given how quickly speech proceeds and how frequently 2-year-olds encounter unfamiliar words. While the costs of such inefficiency in speech processing by young listeners could be reflected in adverse effects on language learning at various levels, one potential consequence is especially relevant to the present study. To the extent that children make frequent errors by responding to unfamiliar prenominal adjectives as potential object names, it will obviously be more difficult for them to begin to learn the meaning of the novel adjective. Only when they are able to listen through the ambiguous word, waiting until they have heard the target noun that follows, can they make sense of the sentence. And only then will they begin to understand the relation between the unfamiliar modifying word and the subsequent word in the sequence, i.e. the name of the familiar object that the novel word pertains to.

17.4. Adjectives as a grammatical category

Learning to listen through an ambiguous adjective may enable the child to get at the meaning of the novel word, contributing to the eventual appreciation of adjectives as a syntactic form class. The question of how young children analyze the individual utterances they hear and map words into appropriate syntactic categories has been debated from nativist and empiricist perspectives. These viewpoints disagree fundamentally on whether or not grammatical categories are innately specified, but both require mechanisms that guide the young learner at the outset. According to Pinker (1984), children make their initial classifications of words using core semantic knowledge by mapping perceptually accessible categories such as object and action onto the grammatical categories noun and verb that are universal across languages. The fact that concrete objects are classified in a noun category with denotations that are both common and stable cross-linguistically makes this semantics-first approach plausible for the case of nouns (Gentner, 1982). However, verbs are inherently relational and their meanings cannot be discerned as easily just by observing actions and events in the world (Genter & Boroditsky, 2001; Gleitman & Gleitman, 1992; Snedeker, 2000). And for adjectives, the semantic bootstrapping approach is a non-starter. Given that adjectives are notoriously mutable in meaning and dependent on the nouns they modify, it seems quite implausible that young children could rely on direct perception of core meanings for relational notions like *good* and *soft* as the primary basis for an initial classification of adjectives as a coherent grammatical class.

From the empiricist perspective, the learner must discover syntactic categories in the ambient language using information in the linguistic input and communicative context. Recent analyses of distributional regularities in child-directed speech have shown how patterns of lexical co-occurrence could potentially be used by the child to infer the initial classification of words. Models using highly local distributional information consisting of the words immediately preceding and following particular target words have been quite successful in classifying target words appropriately as nouns, with varying degrees of success for words in other syntactic categories (Mintz et al., 2002; Redington, Chater, & Finch, 1998). Not surprisingly, adjectives are classified less reliably than nouns and verbs in these models because they occur much less frequently and their co-occurrence patterns are similar to those of nouns. An unknown word preceded immediately by a determiner such as *the* or *a* could turn out to be an adjective, but a model that takes only the preceding words into account would misclassify it as a noun. Mintz (2003) recently improved upon earlier classification rates by using frequent frames in child-directed speech rather than the bigram

statistics used in previous studies. A frame was defined as two jointly occurring words with an intervening target word such as *do__want* and *put__in*, consisting largely of closed class items. However, of the 45 most frequent frames used in this analysis, *the__one* was the only context in which an adjective could occur.

Although a novel word appearing in the frame *the__one* would be uniquely identified as an adjective if a computer made the classification, children do not fare as well. As discussed in the introduction, numerous studies have found that 2-year-olds fail to correctly interpret the novel adjective in phrases like *the TIV one*, preferring a noun interpretation instead (e.g. Taylor & Gelman, 1988). This is a paradoxical finding given that children at this age use adjectives productively in their own speech. We suggested earlier that even 2-year-olds would be likely to parse this phrase as *the TIVwun*, processing it is a strong/weak bisyllabic word that refers to an object. This explanation is supported by our present findings, given that distributional, prosodic, and lexical cues in the phrase *the TIV one* are all consistent with the preferred noun interpretation. But when Mintz and Gleitman (2002) substituted familiar object names for the pronoun *one (e.g. Show me a STOOF horse, Show me a ZAV car)*, 2-year-olds succeeded in identifying *STOOF* correctly as a property term more often than chance. However, their performance was far from perfect and they still made a noun interpretation on many of the trials, again consistent with our finding that encountering an accented unfamiliar word following a determiner will motivate children to search for an object referent, at least some of the time.

Considering our results in the context of these and other recent findings, we can speculate on how young language learners might integrate distributional information with prosodic cues and rudimentary lexical knowledge as they begin to make sense of ambiguous adjectives. The Mintz and Gleitman (2002) results suggest that 2-year-olds are more or less successful in identifying a novel word as a property term rather than an object name when it occurs in the frame *a/the__[noun]*. Note that *[noun]* in this case may not yet represent an abstract grammatical category, but rather a collection of familiar object names that happen to be nouns (Tomasello, 2001). Although nouns may cohere as a grammatical form class well before verbs and other relational terms (e.g. Bloom, 2000; Genter & Boroditsky, 2001), our proposal does not rest on this assumption. If children analyzed a novel word in the frame *a/the__[familiar noun]* the way a computer does in a distributional analysis of frequent frames (e.g. Mintz, 2003), the entire sequence of words would be taken into account before classifying the word as an adjective and there would be no temporary ambiguity. But, as we have shown here, speech is processed incrementally even by 24-month-olds and they don't necessarily wait to hear the final word, especially if there is probabilistic evidence that the novel word could be an object label. This is where prosody comes in. By keeping track of prosodic patterns that co-occur with distributional information, the child can begin to learn something about the kinds of words that co-occur with familiar object names. Although the frame *Where's the__* strongly suggests that an object name is coming up, the deaccented prosodic pattern of the following word contradicts this distributional cue, indicating that the object name is not the next word in the sequence. By listening through the deaccented word following the determiner, the child will be able to access the meaning of the familiar target noun as efficiently as possible. With prosodic cues enabling the young listener to wait in anticipation of the final element in the noun phrase, it may be possible to use the distributional cue of *prenominal* occurrence to identify a class of words that are *not* nouns. Such a transitional category of words that are 'not nouns' (but are followed by nouns) is obviously underspecified, but may initially direct the child's attention toward interpreting a prenominal word in relation to the adjacent object word. Because adjectives are fundamentally relational, it is only in the context of the relevant object word that the meaning of an adjective can be understood at all. It may even be that the typical prosodic pattern of a deaccented prenominal adjective followed by an accented head noun serves,

almost iconically, to emphasize the subordinate relation of the modifier word and its semantic dependency on the object word that follows.

The idea that children integrate distributional information with lexical knowledge in building grammatical categories, first articulated by Maratsos and Chalkey (1980), was recently elaborated by Mintz (2003) in his notion of 'distributional bootstrapping'. Mintz proposed that young language learners initially group nouns using only distributional information, and then readily discover the semantic coherence of the words in this class based on the concrete nouns that are its members. Once nouns have been identified, the positions of nouns in utterances can then be used in subsequent distributional analyses that will eventually lead to the induction of verbs as a grammatical category. Our proposal is related to this idea, in that children use frames beginning with familiar articles such as *the* and *a* and ending in a familiar object name to identify a class of prenominal words. However, our view differs from that described by Mintz in two important respects. First, we make no assumptions about pre-given syntactic categories, which are central to the notion of distributional bootstrapping; our view is equally consistent with the item-based framework of Tomasello and colleagues (Cameron-Faulkner, Lieven, & Tomasello, 2003; Tomasello, 2001). And second, we propose that children use prosodic cues in conjunction with distributional information in their initial identification of prenominal adjectives. Although researchers advocating distributional approaches have focused on information available in patterns of lexical co-occurrence, they have also emphasized the potential importance of other converging sources of information such as prosody that could be used in conjunction with distributional information in grouping new lexical items in appropriate categories (Mintz, 2003; Mintz et al., 2002; Redington et al., 1998). Our findings provide the first experimental evidence that 2-year-olds respond differentially to adjectives in common syntactic frames based on converging lexical and prosodic cues predicting whether or not the ambiguous word is likely to be an object name.

17.5. Cross-linguistic implications

Developing the ability to listen through ambiguous adjectives may be valuable to children learning languages like English in which adjectives abound and can appear before the nouns they modify. But the particular challenge of identifying prenominal adjectives is clearly not a universal problem, given that some languages have no grammatical category of adjectives at all or only a small, closed class of modifier words (Dixon, 1982). Moreover, in the many other languages that do have a substantial adjective class, cross-linguistic differences in morphology and word order inevitably influence the difficulty of the processing task for the learner. In languages with extensive inflectional morphology, adjectives may be marked in ways that facilitate identification. For example, the variable case endings on German prenominal adjectives may reduce ambiguity by distinguishing them to some extent from nouns. Cross-linguistic variability in the ordering of constituents and of words within constituents could also influence the nature and extent of potential ambiguity. Although adjectives are typically adjacent to the words they modify within a noun phrase, in many languages they follow rather than precede the noun. In Catalan, for example, the adjective *cansat* ('tired') would come after the noun *home* ('man') in *un home cansat*. Given that the noun comes first and cannot be immediately followed by another noun, it seems unlikely that a child hearing an unfamiliar adjective in postnominal position would be tempted to interpret the novel word as a potential object name. However, postnominal adjectives can be ambiguous in other ways, in that the same string of words can have different interpretations depending on whether a prepositional phrase is attached to the noun phrase, e.g. *un home, cansat de Paris* ('a man tired of Paris') or the adjective phrase, e.g. *un home cansat, de Paris* ('a tired man from Paris'). Although the nature of the ambiguity is different in Catalan and English, perceptual studies show that prosodic cues are also used to interpret ambiguous

postnominal adjectives in Catalan (Prieto, 1997), as we found with prenominal adjectives in English.

Although the structural characteristics of adjectives vary considerably across languages, the broad implications of our findings go beyond the particular English construction explored in this research. Because rapid processing of sequential information is fundamental to speech comprehension, children learning languages very different from English must confront comparable processing problems, and mechanisms of the sort we describe here could also prove useful to them in resolving temporary ambiguities of other kinds. In exploring how 2-year-old children learn to listen through prenominal adjectives in English, we have focused on one specific example of a more general aspect of linguistic competence essential to understanding speech in any language, the ability to integrate distributional, lexical, prosodic, and other available sources of information in making sense of words that are known, without costly interference from new words in the sentence that are not yet known.

Acknowledgments

This research was supported by a grant from the National Institutes of Health to Anne Fernald (HD 42235) and an Anderson Fund Graduate Research Grant to Kristen Thorpe. Our thanks to Renate Zangl, Virginia Marchman, Alycia Cummings, and the research assistants at the Center for Infant Studies at Stanford for their many contributions, and to the parents and infants who participated in this study.

References

Allopenna PD, Magnuson JS, Tanenhaus MK. Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. Journal of Memory & Language. 1998; 38:419–439.

Beckman, ME.; Elam, GA. Guidelines for ToBI labeling version 3.0. Columbus, OH: The Ohio State University Research Foundation; 1997.

Bever, TG. cognitive basis for linguistic structures. In: Hayes, JR., editor. Cognition and the development of language. New York: Wiley; 1970.

Birch S, Clifton C. Effects of varying focus and accenting of adjuncts on the comprehension of utterances. Journal of Memory & Language. 2002; 47:571–588.

Bloom, P. How children learn the meaning of words. Cambridge, MA: The MIT Press; 2000.

Bolinger, D. Intonation and its parts: Melody in spoken English. Stanford, CA: Stanford University Press; 1986.

Buxton, H. Temporal predictability in the perception of English speech. In: Cutler, A.; Ladd, DR., editors. Prosody: Models and measurements. Heidelberg: Springer; 1983. p. 111-121.p. 111-121.

Cameron-Faulkner T, Lieven E, Tomasello M. A construction based analysis of child directed speech. Cognitive Science. 2003; 27:843–873.

Christiansen MH, Allen J, Seidenberg MS. Learning to segment speech using multiple cues: A connectionist model. Language & Cognitive Processes. 1998; 13:221–268.

Cooper WE, Eady SJ, Mueler PR. Acoustical aspects of contrastive stress in question–answer context. Journal of the Acoustic Society of America. 1985; 77:2142–2156.

Crain, S.; Steedman, MJ. On not being led up the garden path: The use of context by the psychological parser. In: Dowty, D.; Karttunen, L.; Zwicky, A., editors. Natural language parsing. Cambridge: Cambridge University Press; 1985.

Cruttenden A. An experiment involving comprehension of intonation in children from 7 to 10. Journal of Child Language. 1974; 1:221–231.

Cruttenden A. Intonation comprehension in ten-year-olds. Journal of Child Language. 1985; 12:643–661.

Crystal TH, House AS. Segmental durations in connected-speech signals: Syllabic stress. Journal of the Acoustical Society of America. 1988; 83:1574–1585.

Cutler A. Phoneme-monitoring reaction time as a function of preceding intonation contour. Perception & Psychophysics. 1976; 20:55–60.

Cutler A, Carter DM. The predominance of strong initial syllables in the English vocabulary. Computer Speech and Language. 1987; 2:133–142.

Cutler A, Dahan D, van Donselaar W. Prosody in the comprehension of spoken language: A literature review. Language & Speech. 1997; 40:141–201. [PubMed: 9509577]

Cutler A, Foss DJ. On the role of sentence stress in sentence processing. Language & Speech. 1977; 20:1–10. [PubMed: 592948]

Cutler A, Swinney DA. Prosody and the development of comprehension. Journal of Child Language. 1987; 14:145–167. [PubMed: 3558521]

Dale PS, Fenson L. Lexical development norms for young children. Behavior Research Methods, Instruments, & Computers. 1996; 28:125–127.

Dahan D, Tanenhaus MK, Chambers CG. Accent and reference resolution in spoken-language comprehension. Journal of Memory & Language. 2002; 47:292–314.

Dixon, RMW. Where have all the adjectives gone? and other essays in semantics and syntax. Berlin: Mouton; 1982.

Eberhard KM, Spivey-Knowlton MJ, Sedivy JC, Tanenhaus MK. Eye movements as a window into real-time spoken language comprehension in natural contexts. Journal of Psycholinguistic Research. 1995; 24:409–436. [PubMed: 8531168]

Fenson, L.; Dale, PS.; Reznick, JS.; Bates, E.; Thal, D.; Pethick, S. Variability in early communicative development Monographs of the Society for Research in Child Development. Vol. 173. 1994. p. 59

Fernald A, Perfors AF, Marchman VA. Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the second year. Developmental Psychology. (in press).

Fernald A, Pinto J, Swingley D, Weinberg A, McRoberts G. Rapid gains in speed of verbal processing by infants in the 2nd year. Psychological Science. 1998; 9:72–75.

Fernald A, Swingley D, Pinto J. When half a word is enough: Infants can recognize spoken words using partial acoustic-phonetic information. Child Development. 2001; 72:1003–1015. [PubMed: 11480931]

Frazier L, Rayner K. Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. Cognitive Psychology. 1982; 14:178–210.

Gasser M, Smith LB. Learning nouns and adjectives: A connectionist account. Language and Cognitive Processes. 1998; 13:269–306.

Gentner, D. Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In: Kuczaj, SA., editor. Language development: Language, thought, and culture. Vol. 2. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc; 1982.

Genter, D.; Boroditsky, L. Individuation, relativity and early word learning. In: Bowerman, M.; Levinson, S., editors. Language acquisition and conceptual development. Cambridge, UK: Cambridge University Press; 2001.

Gleitman LR, Gleitman H. A picture is worth a thousand words, but that's the problem: The role of syntax in vocabulary acquisition. Current Directions in Psychological Science. 1992; 1:31–35.

Grosjean F. How long is the sentence? Prediction and prosody in the on-line processing of language. Linguistics. 1983; 21:501–529.

Grosjean F, Hirt C. Using prosody to predict the end of sentences in English and French: Normal and brain-damaged subjects. Language & Cognitive Processes. 1996; 11:107–134.

Hall DG. Semantic constraints on word learning: Proper names and adjectives. Child Development. 1994; 65:1299–1317. [PubMed: 7982351]

Hall DG, Waxman SR, Hurwitz WM. How two- and four-year-old children interpret adjectives and count nouns. Child Development. 1993; 64:1651–1661.

Holdgrafer G, Campbell TF. Children's comprehension of intonation as a marker for discourse topic collaboration. Applied Psycholinguistics. 1986; 7:373–384.

Jusczyk PW, Cutler A, Redanz NJ. Infants' preference for the predominant stress patterns of English words. Child Development. 1993; 64:675–687. [PubMed: 8339688]

Jusczyk, PW.; Kemler Nelson, DG. Syntactic units, prosody, and psychological reality during infancy. In: Morgan, JL.; Demuth, K., editors. Signal to syntax: Bootstrapping from speech to grammar in early acquisition. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc; 1996. p. 389-408.

Klibanoff RS, Waxman SR. Basic level object categories support the acquisition of novel adjectives: Evidence from preschool-aged children. Child Development. 2000; 71:649–659. [PubMed: 10953931]

MacWhinney, B. The database. 3. Vol. 2. Mahwah, NJ: Lawrence Erlbaum Associates; 2000. The CHILDES project. Tools for analyzing talk.

Maratsos, M.; Chalkey, MA. The internal language of children's syntax: The ontogenesis and representation of syntactic categories. In: Nelson, K., editor. Children's language. New York: Gardner Press; 1980. p. 2

Markman, EM. Categorization and naming in children. Cambridge, MA: The MIT Press; 1989.

Marslen-Wilson W, Tyler LK. The temporal structure of spoken language understanding. Cognition. 1980; 8:1–71. [PubMed: 7363578]

Mintz TH. Frequent frames as a cue for grammatical categories in child directed speech. Cognition. 2003; 90:91–117. [PubMed: 14597271]

Mintz TH, Gleitman LR. Adjectives really do modify nouns: The incremental and restricted nature of early adjective acquisition. Cognition. 2002; 84:267–293. [PubMed: 12044736]

Mintz TH, Newport EL, Bever TG. The distributional structure of grammatical categories in speech to young children. Cognitive Science. 2002; 26:393–424.

Nelson K. Some attributes of adjectives used by young children. Cognition. 1976; 4:13–30.

O'Shaughnessy D. Linguistic features in fundamental frequency patterns. Journal of Phonetics. 1979; 7:119–145.

Pinker, S. Language learnability and language development. Cambridge, MA: Harvard University Press; 1984.

Prasada, S.; Cummins, M. Structural constraints on the interpretation of novel count nouns Proceedings of the 25th annual proceedings of the Boston University conference on language development; Somerville, MA: Cascadilla Press; 2000.

Prieto, P. Prosodic manifestation of syntactic structure in Catalan. In: Martinez-Gil, F.; Morales-Front, A., editors. Issues in the phonology and morphology of the major Iberian languages. Washington, DC: Georgetown University Press; 1997.

Redington M, Chater N, Finch S. Distributional information: a powerful cue for acquiring syntactic categories. Cognitive Science. 1998; 22:435–469.

Saffran J, Newport E, Aslin R. Word segmentation: The role of distributional cues. Journal of Memory and Language. 1996; 35:606–621.

Sedivy JC, Tanenhaus MK, Chambers CG, Carlson GN. Achieving incremental semantic interpretation through contextual representation. Cognition. 1999; 71:109–147. [PubMed: 10444906]

Seidenberg MS, MacDonald MC. A probabilistic constraints approach to language acquisition and processing. Cognitive Science. 2001; 23:569–588.

Silverman K, Beckman M, Pitrelli J, Ostendorf M, Wightman C, Price P, et al. TOBI: A standard for labeling English prosody. Proceedings of the 1992 International Conference on Spoken Language Processing. 1992; 2:867–870.

Snedeker JC. Information sources for early word learning. Dissertation Abstracts International: Section B: The Sciences & Engineering. 2000; 60:12B.

Snedeker J, Trueswell J. Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. Journal of Memory & Language. 2003; 48:103–130.

Spivey MJ, Tanenhaus MK, Eberhard KM, Sedivy JC. Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. Cognitive Psychology. 2002; 45:447–481. [PubMed: 12480476]

Swingley D. Statistical clustering and the contents of the infant vocabulary. Cognitive Psychology. 2005; 50:86–132. [PubMed: 15556130]

Swingley D, Fernald A. Recognition of words referring to present and absent objects by 24-month-olds. Journal of Memory & Language. 2002; 46:39–56.

Swingley D, Pinto J, Fernald A. Continuous processing in word recognition at 24 months. Cognition. 1999; 71:73–108. [PubMed: 10444905]

Tanenhaus M, Magnuson J, Dahan D, Chambers C. Eye movements and lexical access in spoken language comprehension: Evaluating a linking hypothesis between fixations and linguistic processing. Journal of Psycholinguistic Research. 2000; 29(6):557–580. [PubMed: 11196063]

Tanenhaus M, Spivey-Knowlton M, Eberhard K, Sedivy J. Eye movements as a window into real-time spoken language comprehension in natural contexts. Journal of Psycholinguistic Research. 1995; 24:409–436. [PubMed: 8531168]

Taylor M, Gelman SA. Adjectives and nouns: Children's strategies for learning new words. Child Development. 1988; 59:411–419.

Tomasello, M. The item-based nature of children's early syntactic development. In: Tomasello, M.; Bates, E., editors. Language development: The essential readings. Malden, MA: USishers; 2001. p. 169-186.

Trueswell, JC.; Tanenhaus, MK. Toward a lexicalist framework of constraint-based syntactic ambiguity resolution. In: Clifton, C.; Frazier, L., editors. Perspectives on sentence processing. Hillsdale, NJ: Englandiates, Inc; 1994.

Vogel I, Raimy E. The acquisition of compound vs. phrasal stress: the role of prosodic constituents. Journal of Child Language. 2002; 29:225–250. [PubMed: 12109370]

Warren P, Grabe E, Nolan F. Prosody, phonology, and parsing in closure ambiguities. Language and Cognitive Processes. 1995; 10:457–486.

Waxman SR, Kosowski TD. Nouns mark category relations: Toddlers' and preschoolers' word-learning biases. Child Development. 1990; 61:1461–1473. [PubMed: 2245738]

Waxman SR, Markow DB. Object properties and object kind: Twenty-one-month-old infants' extension of novel adjectives. Child Development. 1998; 69:1313–1329. [PubMed: 9839418]

Zangl R, Klarman L, Thal D, Fernald A, Bates E. Dynamics of word comprehension in infancy: Developments in timing, accuracy, and resistance to acoustic degradation. Journal of Cognition and Development. 2005; 6:179–208.

Appendix

A.1. List of Stimulus Sentences in Experiment 3

1. The little boy had a cold

   The little boy had a cold nose

2. The golfer was on the green

   The golfer was on the green cart

3. The man paid his fare

   The man paid his fair share

4. The burglar stole all the gold

   The burglar stole all the gold jewelry

5. She was surprised there was still a hole

   She was surprised there was still a whole pie

6. The girl was frustrated with her horse

   The girl was frustrated with her hoarse voice

7. She really enjoyed the week

She really enjoyed the weak tea

8. He didn't have the right

He didn't have the right answer

9. The farmer's wife had a pail

The farmer's wife had a pale face

10. The magician appeared out of the blue

The magician appeared out of the blue door

11. He disapproved of the course

He disapproved of the coarse language

12. We could meet them in the front

We could meet them in the front hall

13. She hit him on the back

She hit him on the back porch

14. They were hoping to find a flat

They were hoping to find a flat trail

15. Stop when you get to the square

Stop when you get to the square sign

16. She spilled her drink on the plane

She spilled her drink on the plain tablecloth

17. The woman spotted a deer

The woman spotted a dear friend

18. The lady thought she saw a bear

The lady thought she saw a bare bottom

19. He was worried about being in the red

He was worried about being in the red car

20. The builder had a board

The builder had a bored expression

21. The living room needed a light

The living room needed a light cleaning

22. The fisherman used a reel

The fisherman used a real worm

23. She put the documents in a safe

She put the documents in a safe place

24. The woman went on a fast

The woman went on a fast train

NIH-PA Author Manuscript
NIH-PA Author Manuscript

25. Her father picked up the mail

Her father picked up the male hamster

26. She couldn’t stand the fat

She couldn’t stand the fat boy

27. The farmer inspected the fowl

The farmer inspected the foul odor

28. The old man had a limp

The old man had a limp arm

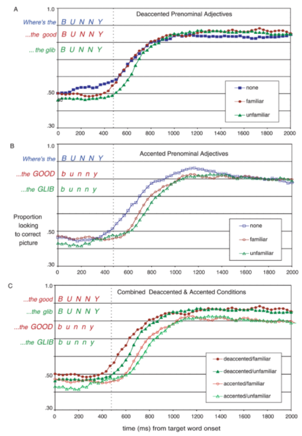

Fig. 1.

The time course of children’s fixations to target pictures in response to target words preceded by deaccented (A) and accented (B) prenominal adjectives in Experiment 1. Curves show changes over time in proportion of total looking to correct picture for unfamiliar adjectives (triangles), familiar adjectives (circles), and control sentences with no prenominal word (squares). Critical trials from deaccented and accented conditions are combined in (C) to enable comparison. Vertical lines mark the mean duration of target nouns. Accented words are indicated in uppercase letters.

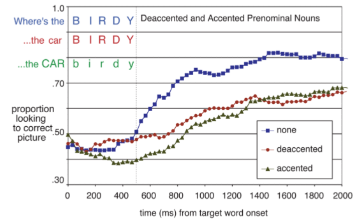

Fig. 2.

The time course of children's looking to target pictures in response to target words preceded by deaccented and accented prenominal nouns in Experiment 2. Curves show changes over time in proportion of total looking to correct picture for accented prenominal nouns (triangles), deaccented prenominal nouns (circles), and control sentences with no prenominal word (squares).

NIH-PA Author Manuscript
NIH-PA Author Manuscript

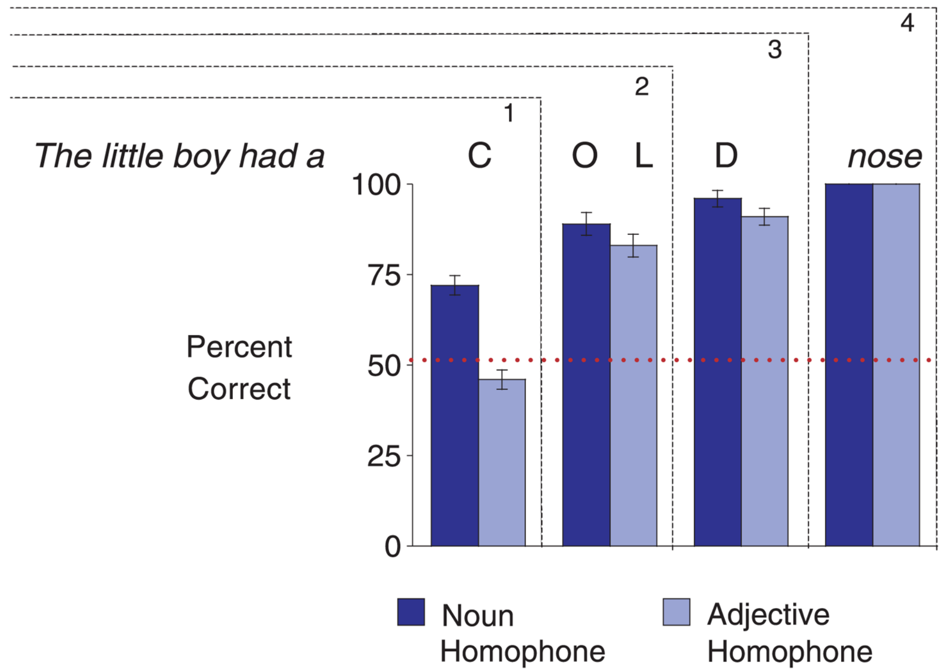

Fig. 3.

Mean accuracy by adults in identifying form class of homophonous target words in Experiment 3. Bars show percent correct judgments for noun and adjective homophones at each segment as the sentence unfolds.

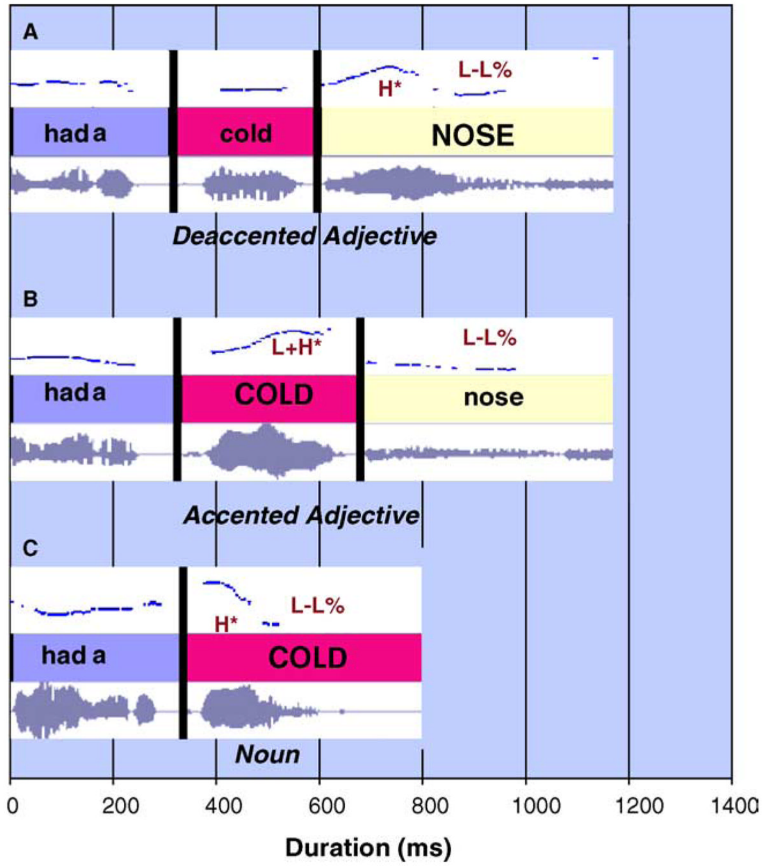

Fig. 4.
Acoustic properties of homophonous target words from sample stimulus sentence in Experiment 3 (‘The little boy had a cold nose’), showing duration, pitch contour, and amplitude envelope for sentence-final words when the homophone was a deaccented adjective (A), accented adjective (B), and noun (C). The accented word in each sentence is indicated in uppercase letters.

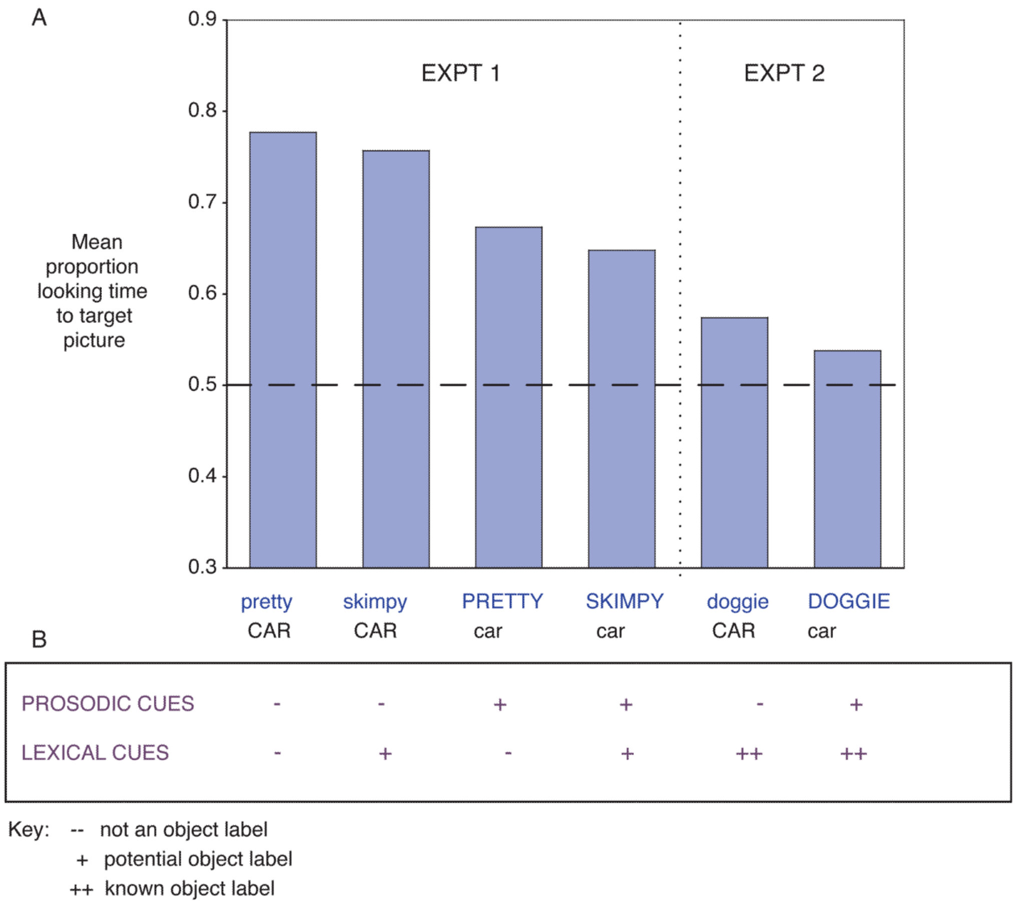

Fig. 5.
(A) Mean proportion of correct looking by 24-month-olds to the target pictures named by sentence-final nouns in each of the six conditions in Experiments 1 and 2. Accented prenominal words are indicated in uppercase letters. (B) Schematic indicating relative strength of prosodic and lexical cues potentially biasing the child towards interpreting the prenominal word as an object name.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 1

Mean accuracy scores[a] for 24-month-olds in Experiment 1 responding to target words preceded by familiar or unfamiliar adjectives that were accented or deaccented

Adjective type	**Accent level**			
	Deaccented		**Accented**	
	M	**(SD)**	***M***	**(SD)**
None[b]	0.75	(0.12)	0.73	(0.12)
Familiar	0.78	(0.13)	0.67	(0.15)
Unfamiliar	0.76	(0.11)	0.65*	(0.14)

*$P<0.05$ comparison to control.

[a]Mean accuracy=mean proportion of time looking to correct picture from 333 to 1800 ms after target word onset.

[b]Control condition.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 2

Mean reaction times (in ms) for 24-month-olds in Experiment 1 to look from distracter to correct target picture, upon hearing the target word preceded by familiar or unfamiliar adjectives that were accented or deaccented

Adjective type	Accent level			
	Deaccented		Accented	
	M	*(SD)*	*M*	*(SD)*
None[a]	654	(129)	655	(139)
Familiar	653	(182)	721*	(154)
Unfamiliar	728	(184)	752*	(175)

* $P<0.05$ comparison to control.

[a] Control condition.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 3

Total number of tokens counted, and mean proportion of adjective use in either prenominal or predicative position from five CHILDES corpora by adjective

Adjective	Total number of tokens	Syntactic frame			
		Prenominal		Predicative	
		M	**(SD)**	***M***	**(SD)**
Big	821	0.73	(0.06)	0.27	(0.06)
Blue	138	0.71	(0.10)	0.29	(0.10)
Red	207	0.67	(0.16)	0.33	(0.16)
Pretty	96	0.53	(0.22)	0.47	(0.22)
Good	740	0.52	(0.10)	0.48	(0.10)
Nice	423	0.51	(0.14)	0.49	(0.14)
Dirty	110	0.42	(0.12)	0.58	(0.12)
Hot	147	0.18	(0.12)	0.82	(0.12)
Cold	109	0.14	(0.12)	0.86	(0.12)
Wet	96	0.14	(0.24)	0.86	(0.24)
Broken	180	0.11	(0.11)	0.89	(0.11)