Accent and reference resolution in spoken-language comprehension

Delphine Dahan,[a],* Michael K. Tanenhaus,[b] and Craig G. Chambers[c]

[a] *Max Planck Institute for Psycholinguistics, P.O. Box 310, 6500 AH Nijmegen, The Netherlands*
[b] *University of Rochester, Rochester, NY, USA*
[c] *University of Calgary, Calgary, Alta., Canada*

Abstract

The role of accent in reference resolution was investigated by monitoring eye fixations to lexical competitors (e.g., candy and candle) as participants followed prerecorded instructions to move objects above or below fixed geometric shapes using a computer mouse. In Experiment 1, the first utterance instructed participants to move one object above or below a shape (e.g., "Put the candle/candy below the triangle") and the second utterance contained an accented or deaccented definite noun phrase which referred to the same object or introduced a new entity (e.g., "Now put the CANDLE above the square" vs. "Now put the candle ABOVE THE SQUARE"). Fixations to the competitor (e.g., candy) demonstrated a bias to interpret deaccented nouns as anaphoric and accented nouns as nonanaphoric. Experiment 2 used only accented nouns in the second instruction, varying whether the referent of this second instruction was the Theme of the first instruction (e.g., "Put the candle below the triangle") or the Goal of the first instruction (e.g., "Put the necklace below the candle"). Participants preferred to interpret accented noun phrases as referring to a previously mentioned nonfocused entity (the Goal) rather than as introducing a new unmentioned entity. © 2002 Elsevier Science (USA). All rights reserved.

As the propositional content of a sentence or utterance is extracted, new events and entities are introduced into a discourse model and reference is made to those that have already been introduced. Many of the referential links between a sentence and the discourse model occur through *anaphoric expressions*—referring expressions whose interpretation depends upon linking the anaphor to a previously introduced event or entity. There is an extensive literature examining how readers interpret anaphoric expressions in text (see Garnham, 2001). A variety of constraints have been shown to influence the interpretation of anaphoric expressions, including the linguistic form of the anaphoric expression (e.g., whether it is a pronoun or a more fully specified definite noun phrase), properties of the linguistic antecedent that reflect information structure (such as whether the antecedent is topicalized), and constraints based on pragmatic inferences (e.g., Garrod, Freudenthal, & Boyle, 1994; Vonk, Hustinx, & Simons, 1992). In spoken language, the prosody of an utterance,

---

* Corresponding author. Fax: +31-24-352-1213.
E-mail address: delphine.dahan@mpi.nl (D. Dahan).

in particular the distribution of pitch accents, is also likely to affect both the salience of discourse entities and the processing of anaphoric expressions (see Cutler, Dahan, & van Donselaar, 1997, and references therein). The present research investigated how prosody interacts with other linguistic factors in constraining reference resolution.

In any utterance, one or more words receive a pitch accent, typically characterized in English by a large pitch excursion on the lexically stressed syllable of the word and often by lengthening of that syllable (e.g., Ladd, 1996). The distribution of accents has often been explained by appealing to the information status of the expressions in the utterance (e.g., Bolinger, 1972; Selkirk, 1984; for a review, see Ladd, 1996). For example, a word that is accented is often assumed to introduce *new* information into the discourse, as opposed to information that is already established or *given*.[1] Experimental studies have shown that listeners are sensitive to the mapping between accent and information structure. For example, listeners are more likely to judge an accent on a *given* noun phrase as inappropriate and are faster and more likely to indicate that a second sentence of a dialogue makes sense when a *new* noun phrase is accented and a *given* noun phrase is deaccented (Birch & Clifton, 1995). Moreover, comprehension time increases when the accent placement is inappropriate for the information-structure requirements of the sentence in its context (Bock & Mazzella, 1983; Nooteboom & Kruyt, 1987; Terken & Nooteboom, 1987; see also van Donselaar & Lentz, 1994).

In studies of reference resolution, it has been proposed that listeners interpret an accented word as introducing a new discourse entity and a deaccented word as anaphoric. Thus, inappropriate accent patterns would interfere with the interpretation of referential expressions, resulting in the observed delayed comprehension (Bard, Cooper, Kowtko, & Brew, 1991; Terken & Nooteboom, 1987). Terken and Nooteboom (1987) offered an explicit algorithm for the use of accent pattern in referent resolution. According to their hypothesis,

a deaccented expression is interpreted within the restricted set of already activated discourse entities, while the interpretation of an accented entity is built mainly on the basis of the information contained in the speech signal. The deaccentuation of a new referent would lead the listener to try to map the expression onto the restricted set of activated referents, which would eventually fail and require time-consuming reanalysis. The accentuation of an already activated referent would lead the listener to construct an interpretation from the bottom-up input, which will take more time than when the expression can be mapped onto the restricted set of activated referents (see Fowler & Housum, 1987, for a similar proposal). Although this hypothesis about how accent influences reference resolution is highly plausible, its support comes from indirect evidence, that is, slower comprehension times when information structure and accentual status are incompatible. The current research was designed to provide more direct evidence by examining which potential discourse referents are considered by listeners as they process accented and deaccented noun phrases.

While it is generally agreed that the distinction between *given* and *new* plays an important role in explaining accent patterns, definitions of what is *given* and what is *new* vary considerably (Clark & Haviland, 1977; Halliday, 1967; Prince, 1981). For example, *given* information is sometimes defined as "information known to the listener" or as "anaphorically recoverable;" *new* information is defined as "information novel to the listener," "textually and situationally nonderivable," but also as "contrary to some predicted or stated alternative." Nevertheless, in most published psycholinguistic studies on this topic, *given* information is typically equated with entities that have been previously introduced into the discourse and which share the same linguistic form and grammatical role as their antecedents (although see Bock & Mazzella, 1983). In contrast, an expression is described as containing *new* information when it introduces an entity that has not been previously mentioned. Thus, the *given/new* distinction is operationalized by contrasting entities that have been previously mentioned in the same grammatical role with unmentioned entities.

However, recent production studies have suggested a more complex mapping between accentuation and information structure. In particular, it appears that accented noun phrases may be used

---

[1] Note that the relationship between accentuation and information structure is more complex when the most informative part of the utterance consists of a multiple-word constituent, such as an entire phrase. However, we limit our discussion to the case of entities that can be referred to with single accented or unaccented words (case of "narrow focus," see Ladd, 1980).

to refer to previously mentioned entities in some cases. For example, while Terken and Hirschberg (1994) argued that deaccentuation of an already mentioned entity is more likely to occur if the entity occupies the same grammatical role in its first and second mentions, analyses of spontaneous dialogue have shown that deaccenting upon repeated mention is rare and is not affected by considerations of sentence structure (Bard & Aylett, 1999). It appears that accent pattern cannot be accurately predicted by a simple distinction between mentioned and unmentioned entities.

The present research focused on the referential interpretation of accented and deaccented definite noun phrases under conditions where the eventual referent was the most salient previously mentioned entity, a less salient mentioned entity, or a new discourse entity. A definite noun phrase will typically contain enough information for its referent to be uniquely identified once the processing of the head noun is complete.[2] However, because the sound pattern of a word unfolds over time, the acoustic/phonetic input is often briefly consistent with multiple lexical candidates. During spoken-word recognition, candidates matching the unfolding input become partially activated and compete for recognition until sufficient information arrives to uniquely identify a single candidate (e.g., Marslen-Wilson, 1987; Zwitserlood, 1989). Nevertheless, listeners may have access to information about the accentual status of a word *before* the word itself has been disambiguated (and sometimes before the word is heard, as shown by Cutler, 1976). For example, a listener might be able to identify that the stressed syllable in candle carries a pitch accent before encountering phonetic information that would distinguish it from its lexical competitor, candy. The current studies use temporary ambiguities like these to test the hypothesis that listeners use accent to help circumscribe referential candidates.

In the present studies, participants followed spoken instructions to move objects displayed on a computer screen. The display contained four objects and four geometric shapes. On critical trials, two of the objects had names that overlapped at onset. The overlap included the initial stressed syllable for polysyllabic names, e.g., candle and candy, or the syllable onset and nu-

cleus for monosyllabic names, e.g., bell and bed. A trial consisted of two consecutive instructions. The first instruction mentioned one of the objects (e.g., "Put the candle/candy below the triangle"), making it *given* information and setting the context for the second, critical instruction. This instruction could refer either to the object mentioned in the first instruction or to a different object (e.g., "Now put the candle above the square"). Thus, the noun phrase in the critical instruction was either used anaphorically, to refer to the object previously mentioned in the first instruction, or nonanaphorically, to refer to a previously unmentioned object. The effect of accentuation was evaluated by varying whether the critical noun phrase received an accent. We used the pattern of eye fixations to the referent of the noun phrase in the critical instruction and its lexical competitor to examine how accent affects reference resolution.

We monitored eye movements because they provide a continuous and fine-grained measure of spoken-language processing in which the response is closely time locked to the input without interrupting the speech stream and without requiring the use of metalinguistic judgments (Cooper, 1974; Tanenhaus, Magnuson, Dahan, & Chambers, 2000; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995, 1996). It is thus well suited for asking questions about real-time interpretation. Recent studies have shown that fixations to potential referents can reveal on-line use of linguistic information during reference resolution (Arnold, Eisenband, Brown-Schmidt, & Trueswell, 2000; Chambers, Tanenhaus, Eberhard, Filip, & Carlson, 2002; Eberhard, Spivey-Knowlton, Sedivy, & Tanenhaus, 1995; Sedivy, Tanenhaus, Chambers, & Carlson, 1999). Eye movements are also sensitive to subtle acoustic/phonetic variation that provides probabilistic information about potential lexical candidates (Allopenna, Magnuson, & Tanenhaus, 1998; Dahan, Magnuson, & Tanenhaus, 2001a; Dahan, Magnuson, Tanenhaus, & Hogan, 2001b). For example, in displays similar to the ones used in the current experiments, fixations to pictures with names that phonetically match the input begin to increase 200–300 ms after the onset of the target word. When the programming time of approximately 200 ms for a saccadic eye movement is taken into account, this result demonstrates that eye movements are sensitive to changes in lexical activation beginning with the onset of the word (Allopenna et al., 1998; Dahan et al., 2001a).

---

[2] For the sake of simplicity, we are excluding noun phrases with postnominal modifiers, e.g., "the candle with a thick wick."

Crucially, fixation patterns are also affected by coarticulatory information in vowels that provide partial cues to the place of articulation of an upcoming consonant, demonstrating that eye movements are sensitive to subtle phonetic information extending over a very brief time window (Dahan et al., 2001b). Since the acoustic changes resulting from a pitch accent are present in vowels (and possibly in the entire stressed syllable, see Turk & Sawusch, 1997), this sensitivity indicates that effects of accent on reference resolution should be detected in the proportion and timing of fixations to the referent and its lexical competitor.

We evaluated the hypothesis that listeners will preferentially interpret a deaccented noun phrase as referring to the most prominent previously mentioned entity and an accented noun phrase as referring to a *new*, unmentioned entity or to a previously mentioned but less prominent entity. In Experiment 1, we varied the accent pattern of the noun phrase in the second instruction and whether it referred to an unmentioned entity or to a previously mentioned entity (e.g., "Put the candle/candy below the triangle. Now put the CANDLE above the square/Now put the candle ABOVE THE SQUARE."[3] Of interest was participants' early interpretation of the accented or deaccented noun phrase, that is, during the narrow time window where the phonemic signal was consistent with both the target (e.g., candle) and its competitor (e.g., candy). We hypothesized that, early in the presentation of the noun, listeners would be more likely to interpret the noun phrase anaphorically when this noun phrase was deaccented than when it was accented. In Experiment 2, the thematic role in which the referent of the second instruction was introduced in the first instruction was varied (Theme or Goal) and we examined the interpretation of an accented expression with regard to the thematic role that its referent played during its previous mention.

Experiment 1

This experiment evaluated the hypothesis that an accented noun phrase tends to be interpreted as

referring to a *new* (defined here as unmentioned) entity and a deaccented noun phrase as referring to a *given* (defined here as mentioned) entity. On each trial, four different moveable objects were displayed, along with four fixed geometric shapes, and participants were instructed to perform two successive actions. The first instruction introduced an object in the role of Theme (e.g., "Put the candle / candy below the triangle"). Following this instruction, the mentioned object (the candle or the candy) would have the status of *given* information, whereas subsequent mention of the other objects (i.e., the unmentioned objects) would convey *new* information. The second instruction contained a noun phrase which referred either to the same object mentioned in the first instruction or to a different object. This noun phrase was either accented or deaccented—in which case the most prominent phrase was the Goal (e.g., "Now put the CANDLE above the square" vs. "Now put the candle ABOVE THE SQUARE"). As mentioned earlier, a full noun can provide enough information to disambiguate its referent. For instance, the sound form of the word candle will eventually uniquely identify its referent, especially in the limited context provided by the four objects in the visual display. However, the point in time where the intended referent can be identified can vary depending on the names of the other objects on the display. For instance, if the picture of a candy is presented along with the referent picture candle, speech information from the target word candle cannot be used to uniquely identify a referent until after the syllable can. Importantly, this ambiguous portion, especially when corresponding to the lexically stressed syllable of the word, can convey pitch-accent information. We used temporary lexical ambiguities like these to test whether the information status (i.e., *given* or *new*) of the lexical-competitor object interacted with the presence or the absence of an accent on the name of the referent in the second instruction. We orthogonally varied (a) whether the competitor object (e.g., candy) had been mentioned in the first instruction (and thus whether it was associated with *given* or *new* information as the second instruction was heard) and (b) whether the name of the referent object in the second instruction (e.g., candle) was accented or deaccented. This yielded the four conditions illustrated in Table 1.

The hypothesis that an accent on a noun phrase is interpreted as introducing a *new* entity in the discourse predicts that a *new* competitor ob-

---

[3] Throughout this article, the use of capitals indicates which phrase the speaker was instructed to emphasize, not which words were accented.

Table 1
Experiment 1: Illustration for the four conditions (discourse context × accent pattern)

Discourse context	Accent pattern	First instruction	Second instruction
Anaphoric (*new* competitor)	Accented target word	Put the candle below the triangle	Now put the CANDLE above the square
Anaphoric (*new* competitor)	Deaccented target word	Put the candle below the triangle	Now put the candle ABOVE THE SQUARE
Nonanaphoric (*given* competitor)	Accented target word	Put the candy below the triangle	Now put the CANDLE above the square
Nonanaphoric (*given* competitor)	Deaccented target word	Put the candy below the triangle	Now put the candle ABOVE THE SQUARE

Note. Capitals indicate which part of the second instruction the speaker was instructed to emphasize. The target word in the second instruction is underlined.

ject will be considered as a possible referent more when the referent name in the second instruction is accented than when it is deaccented. We thus predicted more fixations to the *new* competitor in the anaphoric-accented condition than in the anaphoric-deaccented condition. Conversely, if deaccentuation of a noun phrase biases an anaphoric interpretation, a *given* competitor object is predicted to be considered as a possible referent more when the referent name in the second instruction is deaccented than when it is accented. We thus predicted more fixations to the *given* competitor in the nonanaphoric-deaccented condition than in the nonanaphoric-accented condition. Note that the presence of an accent usually increases the duration of the accent-bearing word; the duration of the speech fragment that matches both the target and its competitor thus extends over a longer period of time. This duration increase alone could account for the first prediction (namely, more fixations to the *new* competitor when the referent name is accented than when it is deaccented), because the speech input disambiguates between the target and the competitor later. However, it could not account for the second prediction, that is, more fixations to the *given* competitor when the referent name is deaccented than when it is accented. On the contrary, the increase in duration should work *against* this prediction. Testing both predictions is thus crucial to establish the role of accent in referential interpretation.

Method

Participants

Sixteen native speakers of American English were recruited at the University of Rochester and were paid a small amount for their participation.

Materials

Twenty-four pairs of picturable nouns with phonetically similar onsets (e.g., candle–candy) were selected. One member of each pair was assigned the role of target, and the other was the competitor. (Note that this role reflects which item of the pair was target or competitor in the *second* instruction of each trial. In the first instruction, either the target or the competitor played the role of referent picture, see Table 1.) The mean lexical frequencies of targets and competitors, as reported in Francis & Kučera (1982), were identical (39.6 per million). Each of the 24 item pairs was associated with two picturable, distractor nouns, resulting in four pictures for each experimental display. The complete set of materials is presented in Appendix.

In addition to these 24 experimental trials, 28 filler trials were constructed. Recall that for the experimental trials, the target picture in the second instruction was either the same picture as in the first instruction or its phonetic counterpart. In order to prevent participants from developing expectations that pictures with phonetically similar names were likely to be targets in either instruction, 12 of the filler trials consisted of two items that started with similar sounds and two phonetically unrelated items (e.g., peanut, peacock, bottle, and saw). For 6 of these fillers, the target of the first instruction was one of the phonetically related items (e.g., peanut), while the target of the second instruction was one of the two phonetically unrelated items (e.g., bottle). For the 6 other fillers, neither of the phonetically related items was the target on either instruction. Finally, 16 more fillers composed of four phonetically unrelated items were constructed. Overall, in half of the 28 fillers, the same object was referred to in the first and second instructions (anaphoric), while in the other half, two different objects were referred to (nonanaphoric). Four of the fillers

(2 anaphoric, 2 nonanaphoric) were presented at the beginning of the session to familiarize participants with the task and the procedure.

The 208 pictures [(24 experimental trials + 28 filler trials) × 4 pictures] were selected from the Snodgrass & Vanderwart (1980) and the Cycowicz, Friedman, Rothstein, & Snodgrass, 1997) picture sets, as well as from children's picture dictionaries and a commercially available clip-art database. All were black and white line drawings.

The spoken instructions were recorded by a male native speaker of American English in a sound-proof room, sampling at 22,050 Hz. These instructions were read from printed materials. The sentence fragment to be accented in the second instruction was indicated by the use of capitals. The first instructions were followed by a semicolon, which was intended to trigger the production of an intonational rise (i.e., a high boundary tone). Pierrehumbert and Hirschberg (1990) have suggested that boundary tones convey information about relationships among intonational phrases; a high boundary tone indicates that the current phrase is to be interpreted with respect to the following phrase. The presence of a high boundary tone at the end of the first instruction was intended to maximize the likelihood that listeners would interpret the second instruction with respect to the context established in the first instruction. For each experimental trial, four utterances were recorded: two utterances to be used as first instructions, referring to either the target or the competitor of each pair, and two used as second instructions, with an accented or deaccent target word. These four instructions were paired in four different ways when presented to the participants (see Table 1). In order to prevent the speaker from producing potentially nonfelicitous accent patterns (e.g., deaccenting the target word when it was not present in the preceding context) or adopting a contrastive interpretation of the second instruction when the first instruction referred to an object with a similarly sounding name (e.g., bed vs. bell), some of the four utterances were paired with an utterance that referred to neither of the two items. An example of the recording script for one experimental trial is shown in example 1 (target, candle; competitor, candy; the four utterances used in the experiment are underlined):

(Example 1)

Put the candle below the triangle; now put the candle ABOVE THE SQUARE.

Put the book below the triangle; now put the CANDLE above the square.

Put the candy below the triangle; now put the BOOK above the square.

The accentual pattern for the 28 fillers (i.e., accented or deaccented noun phrase on the Theme) was counterbalanced across the anaphoric and nonanaphoric fillers.

Each utterance was edited and its prosody transcribed using the ToBI labeling system (Beckman & Hirschberg, 1994; Silverman et al., 1992; for a brief summary of this coding system, see Carlson, Clifton, & Frazier, 2001). The first instructions were characterized by the prosodic pattern described in (1) and (2), in 23 and 15 of the 48 instructions, respectively. The prosody of the remaining 10 instructions differed slightly. Five were identical to (1) but with a simple H* pitch accent on the Theme; 3 were identical to (2) but with a L+H* pitch accent on the Theme; finally 2 instructions were characterized by a L* pitch accent on the Theme, followed by a H − H%. Crucially, all of the first instructions ended with a high boundary tone, as we had intended.

(1) Put the candy   below the diamond
      H*   L + H*   H* H − H%
               L − L%
(2) Put the sandal   above the circle
      H*   H* L − H%   H* H − H%

The prosody of the second instruction varied dramatically between the accented and deaccented conditions. In the accented condition, the prosody of most of the instructions was transcribed as (3) (11 of 24 cases) or (4) (6 cases), depending on the type of pitch accent on the Theme. For the remaining 7 cases, the boundary tone following the Theme was low (L − L%), occurring after a simple H* (4 cases) or a more complex pitch accent (L + H*, 3 cases). Finally, the pitch accent on the goal was transcribed either as H + !H* (as in the examples below), if the portion of speech preceding the accented syllable of the Goal was high- pitched, or as a simple H* if that portion was not high-pitched. Beyond these differences, the transcriptions indicate that all of the second instructions in the accented condition received a well-marked pitch accent on the stressed syllable of the Theme (with the prominence sometimes further enhanced by a sharp rise from a valley), no pitch accent on the preposition (below or above), and a relatively reduced prominence on the Goal.

(3) Now put the BELL   above the triangle
    H*   L + H*   H + !H*
                  L − H%   L − L%

(4) Now put the SANDAL below the square
    H*   H*   H + !H*
            L − H%   L − L%

In the deaccented condition, the second instruction received the prosodic pattern illustrated in (5) (12 of 24 cases), or that illustrated in (6) (9 cases), or that in (7) (3 cases). Thus, all instructions in this condition were characterized by a reduced prominence on the Theme (due to a high pitch preceding a downstepped pitch accent) and a highly prominent Goal (due to an intonational boundary preceding the preposition *above* or *below* and a sharp rise leading to a high pitch accent).

(5) Now put the bell
    L + H*   H + !H* H − L%
    ABOVE THE TRIANGLE
      L + H*   H + !H* L − L%

(6) Now put the candle
    L* + H   H + !H* H − H%
    BELOW THE CIRCLE
      L + H*   H + !H* L − L%

(7) Now put the chain
    H*   H + !H* H − H%
    BELOW THE DIAMOND
      L + H*   H + !H* L − L%

Additional acoustic analyses were performed on the second instructions to confirm the contrast between the accented and deaccented versions. These analyses (presented in Table 2) consisted of measuring F0 peaks on stressed vowels at various points in the instruction, as well as a few durational measurements. Table 2 presents those values averaged across items in both conditions. We also computed the mean difference between the accented and deaccented conditions and the standard deviation of these differences across items; the latter provides a dispersion value of the differences that can be used to estimate statistical significance. These analyses revealed that the Theme benefited from a large pitch excursion and lengthening, occasionally followed by a long pause, in the accented condition compared to the deaccented condition. Conversely, the preposition was characterized by a higher pitch in the deaccented condition than in the accented condition.

Procedure

Participants were seated at a comfortable distance from the computer screen (for details, see Dahan, Swingley, Tanenhaus, & Magnuson, 2000). Eye gaze was monitored using an *Applied Science Laboratories* head-mounted eye tracker (Model E 5000). A scene camera was aligned with the participant's line of sight. A calibration procedure allowed software to superimpose crosshairs showing the point of gaze on a HI-8 videotape recording of the scene. The visual scene was recorded at a rate of 30 frames per second, and each frame was stamped with a time code. Auditory stimuli were played to the participant through headphones and simultaneously to the HI-8 VCR, providing an audio record of each trial. The structure of each trial was as follows: First, a 5×5 grid with a centered cross appeared on the screen, and participants were told to click on the cross. This allowed the experimenter to check that calibration accuracy was acceptable, as participants briefly fixated on the cross before clicking on it. Four line drawings and four colored geometric shapes appeared on specific cells of the grid. Participants' eye movements were not constrained in any way. As the pictures appeared on the screen, the first spoken instruction began,

Table 2
Experiment 1: Mean F0 peak on stressed vowels (in Hz) and duration (in ms) of various parts of the second instruction in the accented and deaccented conditions, as well as the average and standard deviation of the difference between the accented and deaccented conditions

	F0 peak				Duration	
	"Now"	Theme	Preposition	Goal	Theme noun	Pause following Theme (if any)
Mean accented	141	151	104	105	539	138 ($N = 8$)
Mean deaccented	150	121	114	109	330	35 ($N = 1$)
Mean accented–deaccented	−9	30	−10	−4	209	
Standard deviation accented–deaccented	17	14	8	12	66	

instructing participants to move one of the four pictured objects above or below one of the four geometric shapes (e.g., ''Put the candle below the triangle''). As soon as the action was performed, the second instruction was played (e.g., ''Now put the candy above the square''). Once the participant completed the instruction, the next trial began. The positions of the geometric shapes did not vary across trials. The position of each pictured object was randomized for each participant and each trial.

Four lists were constructed by varying in which of the four conditions the experimental items were presented (i.e., anaphoric-accented, anaphoric-deaccented; nonanaphoric-accented, and nonanaphoric-deaccented). Within each list, six experimental items were assigned to each condition. Four participants were randomly assigned to each list. For each list, three random orders were created and participants were randomly assigned to each order.

The data were collected from the videotape records using an editing VCR with frame-by-frame controls and synchronized video and audio channels. Coders used the crosshairs generated by the eye tracker to establish, for each experimental trial, which of the four pictures or the cross was fixated at each time frame (see Dahan et al. (2000), for full details on the coding procedure).

Results

For three participants, a few trials were missing because of technical failure or track loss (11 trials, in total). In addition, a few trials were excluded from the analyses because some participants did not fixate the target picture in the second instruction while clicking on it and moving it with the computer mouse (9 trials) or because they moved the wrong object in the first instruction (2 trials). Missing trials represented 5.7% of the data (22 trials of 384). In order to give all participants' data the same weight in the analyses, fixation values for the missing trials were estimated by using the participants' average proportions over the remaining trials.

Fig. 1 presents the proportions (averaged across participants) of fixations to the target and the competitor in 33-ms time slices from 0 to 1500 ms beginning with the onset of the target word in the second instruction (including the closure for initial voiceless stop consonants). It is important to emphasize that the minimum latency to plan and launch a saccade is estimated to be
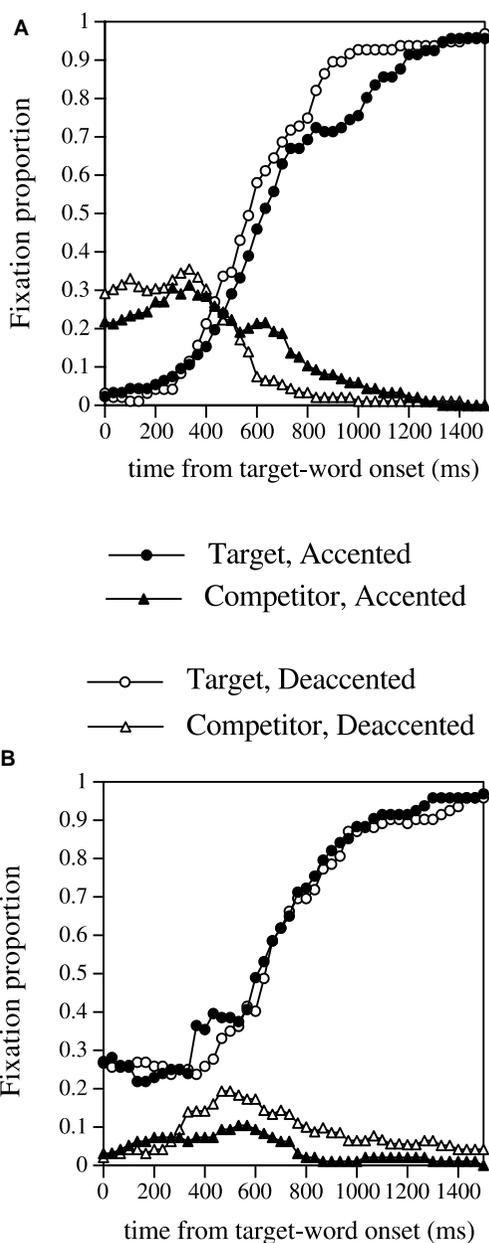

Fig. 1. Experiment 1. Fixation proportions over time to the target and the competitor pictures from the onset of the target word, as a function of the accent pattern (accented or deaccented), in the anaphoric condition (A) and in the nonanaphoric condition (B).

between 150 and 180 ms in simple tasks (e.g., Fischer, 1992; Saslow, 1967), and intersaccadic intervals in tasks like visual search fall in the range of 200–300 ms (e.g., Viviani, 1990). Consequently in the present data, fixations realized within the

first 200 ms of the target word are likely to have been driven by speech input that *preceded* the onset of the target word and not from the target word itself. We thus did not expect to see effects of processing the target word before approximately 200–300 ms after word onset (see Allopenna et al., 1998; Dahan et al., 2001a; for results supporting such an estimate).

Fig. 1A presents the fixation proportions to the target and its competitor in the anaphoric discourse context, where both the first and second instructions referred to the same object (the target) and thus where the competitor object corresponded to a *new* entity. Fig. 1B presents the fixation proportions to the target and its competitor in the nonanaphoric discourse context, where the first instruction referred to the competitor and the second instruction to the target. In this condition, the competitor object constitutes a *given* entity. In both graphs, fixations to target and competitor are indicated with filled symbols when the target word carried an accent and with empty symbols when the target word was deaccented. Inspection of the graphs reveals an effect of accentuation pattern on competitor fixations beginning approximately 300 ms after the onset of the target word. In the anaphoric condition (Fig. 1A), the proportion of fixations to the (*new*) competitor decreased at a slower rate in the accented condition than in the deaccented condition. Conversely, in the nonanaphoric condition, the proportion of fixations to the (*given*) competitor increased more sharply and remained higher in the deaccented condition than in the accented condition. Both effects are consistent with the hypotheses that a *new* entity is considered the potential referent of a noun phrase more often when the noun phrase is accented than deaccented and that a *given* entity is considered the potential referent of a noun phrase more often when the noun phrase is deaccented than accented.

Inspection of the fixation proportions at the precise onset of the target word (i.e., at the point before eye movements could reflect the effect of information from the target word itself) revealed a bias for fixating *new* pictures, that is, pictures that were *not* referred to in the first instruction. In the anaphoric condition (Fig. 1A), the target was referred to in the first instruction, thus acquiring the status of *given* entity, whereas the competitor was not referred to in the first instruction and would thus convey *new* information. In this condition, fixations to the target picture were very rare at target onset; conversely, between 20 and 30% of

the participants fixated the competitor picture. In the nonanaphoric condition (Fig. 1B), the target was a *new* entity, whereas the competitor was *given*. The competitor picture was rarely fixated at target onset, while the target picture was fixated by about 28% of the participants.

Statistical analyses were conducted by calculating, for each condition, the proportion of fixations to the competitor over the time windows extending from 0 to 300 ms and 300 to 1000 ms after the onset of the target word. Two-way (discourse context × accent pattern) ANOVAs were then conducted separately for the scores from each time window. Over the 0–300-ms time window, the analyses revealed a main effect of discourse context on competitor fixations ($F_1(1, 15) = 29.68$, $p < .0001$, $MSE = .029$; $F_2(1, 23) = 33.82$, $p < .0001$, $MSE = .0347$), with more fixations to the competitor when it constituted *new* rather than *given* information, no effect of accent, and no interaction between the two factors. This result confirms the bias due to the *given* vs. *new* status of the displayed pictures discussed previously. Over the window extending from 300 to 1000 ms after target onset, the mean proportion of fixations to the competitor in the anaphoric condition was significantly higher when the target word was accented than when it was deaccented (.17 vs. .12). Conversely, in the nonanaphoric condition, the mean fixation proportion to the competitor was significantly higher when the target word was deaccented than when it was accented (.13 vs. .06). This pattern was reflected in a reliable interaction between accent pattern and discourse context ($F_1(1, 15) = 14.18$, $p < .005$, $MSE = .005$; $F_2(1, 23) = 8.42$, $p < .01$, $MSE = .014$) and significant (or marginally significant) pairwise comparisons between accent conditions for each discourse condition independently ($t_1(15) = 2.28$, $p < .05$, $t_2(23) = 1.93$, $p = .06$, for the anaphoric condition; $t_1(15) = 2.74$, $p < .05$, $t_2(23) = 2.77$, $p < .05$, for the nonanaphoric condition). No significant main effect of discourse context or accent pattern was found.

As Fig. 1 shows, the effect of accent on fixations to the competitor occurs later in the anaphoric condition (Fig. 1A) than in the nonanaphoric condition (Fig. 1B). This might be due to the fact that, in the anaphoric condition, the deaccented condition showed a (nonsignificant) initial advantage over the accented condition, by contrast with the final pattern. Because the fixation-proportion values at time t are not fully independent from those at time $t + 1$ or at time $t - 1$

(a fixation usually extends for longer than one video frame, 33.3 ms), one would not expect to see such an initial bias reverse instantaneously. This might have delayed where in time the fixation proportion to the competitor in the accented condition became significantly higher than the fixation proportion in the deaccented condition. A baseline difference between the accented and deaccented conditions was not present in the nonanaphoric condition, thereby allowing a difference between the two conditions to be detected soon after 300 ms.

In contrast to the pattern of results obtained for fixations to the competitor, the pattern for the targets was less straightforward. Mean target fixations are presented in Table 3. Over the 0–300 ms time window, analyses revealed a main effect of discourse context on target fixations ($F_1(1, 15) = 34.95$, $p < .0001$, $MSE = .02$; $F_2(1, 23) = 42.21$, $p < .0001$, $MSE = .0256$), with more fixations to the target when it constituted *new* than *given* information, no effect of accent, and no interaction between the two factors. This confirms the bias due to the *given* vs. *new* status of the target picture depending upon the discourse context. Over the 300–1000 ms time window, the proportion of participants fixating the target referent rose faster in the anaphoric-deaccented condition than in the anaphoric-accented condition (.59 vs. .50), as predicted by the hypothesis that a deaccented noun phrase is interpreted anaphorically. However, there was only very weak evidence for an advantage in fixating nonanaphoric targets when the noun phrase was accented compared to when the noun phrase was deaccented (.58 vs. .55). This pattern is reflected in a significant (or marginally significant) interaction between accent pattern and discourse context ($F_1(1, 15) = 5.83$, $p < .05$, $MSE = .011$; $F_2(1, 23) = 3.16$, $p = .09$, $MSE = .0369$) and a significant pairwise comparison between accent conditions in the anaphoric condition, but a nonsignificant pairwise comparison in

the nonanaphoric condition (for the anaphoric condition, $t_1(15) = 2.74$, $p < .05$, $t_2(23) = 2.78$, $p < .05$; for the nonanaphoric condition, $t_1(15) = .89$, $t_2(23) = .64$). No significant main effect of discourse context or accent pattern was found. Thus, it appears that the anaphoric interpretation of the referent in the second instruction was facilitated when the noun phrase was deaccented compared to when it was accented, but the nonanaphoric interpretation of the referent was not facilitated when the noun phrase was accented compared to deaccented. We will return to this asymmetry in the results under General discussion after presenting the findings from Experiment 2, which help to clarify the pattern of results for the target in the current experiment. To anticipate, we suggest that intonational and lexical factors contributed to favoring the interpretation of the accented noun phrase in the second instruction as referring to a previously mentioned but nonfocal entity.

We also conducted statistical analyses to confirm the apparent bias in the first 200–300 ms after target onset for pictures associated with *new* entities. These analyses included fixations to the targets, to the competitors, and to the unrelated distractors (which always represented *new* entities). Table 4 presents the proportion of fixations on the various objects in the display (i.e., target, competitor, and distractor pictures) for each condition, averaged across the first 200 ms of the target word. For clarity, the proportion values are underlined when the corresponding objects constituted *new* information. As the table shows, there was an initial bias to fixate a picture associated with a *new* entity. A three-way (discourse context × accent pattern × picture type) ANOVA revealed a significant interaction between the discourse context and picture ($F_1(2, 30) = 33.4$, $p < .0001$, $MSE = .0244$), with no other main effect or interaction. This confirms the influence of the discourse status of the pictures on fixations

Table 3
Proportion of fixations to the target and competitor pictures over the time windows of 0–300 and 300–1000 ms as a function of the discourse context (anaphoric or nonanaphoric) and the accent pattern of the target word (accented or deaccented)

| | 0–300 ms | | 300–1000 ms | |
	Competitor	Target	Competitor	Target
Anaphoric accented	.25	.05	.17	.50
Anaphoric deaccented	.31	.03	.12	.59
Nonanaphoric accented	.06	.25	.06	.58
Nonanaphoric deaccented	.04	.26	.13	.55

Table 4
Proportion of fixations to the target picture, its competitor, and the averaged distractor pictures, over the first 200 ms after target onset, as a function of the discourse context (anaphoric or nonanaphoric) and the accent pattern of the target word (accented or deaccented)

	Target	Competitor	Distractors
Anaphoric accented	.04	.23	.18
Anaphoric deaccented	.02	.31	.15
Nonanaphoric accented	.25	.05	.14
Nonanaphoric deaccented	.26	.03	.17

Note. The fixation proportions on *new* pictures are underlined.

occurring in the initial portion. A more sensitive analysis was then conducted using a second three-way [discourse status (*given* vs. *new*) × accent pattern × picture type] ANOVA on fixations to the target and competitor (fixations to the distractors could not be included because these pictures always carried *new* information). This analysis revealed a main effect of discourse status ($F_1(1, 15) = 55.9$, $p < .00001$, $MSE = .0292$), which did not interact with the status of the picture as a competitor or target.

Discussion

The results demonstrated clear effects of accent on reference resolution. Eye movements to the competitor (which phonetically overlapped with the target word at onset) revealed that people were more likely to fixate a *new* entity when the target word was accented than deaccented and more likely to fixate a *given* entity when the target word was deaccented than accented. An effect of accent was found for the discourse-context condition in which the competitor was associated with *new* information (i.e., anaphoric condition) and the discourse-context condition in which the competitor was associated with *given* information (i.e., nonanaphoric condition). As mentioned earlier, the finding that more fixations to the *new* competitor picture occurred when the target word was accented than when it was deaccented could be accounted for by the lengthening of the accented target word. Indeed, the portion of target word that is ambiguous between the target and competitor extends over a longer period of time, thereby allowing the competitor to remain active for longer. However, this argument cannot account for the finding that more fixations to the *given* competitor picture were found when the target word was deaccented than when it was accented. In fact, the magnitude of the predicted effects was quite similar in the two discourse-

context conditions. Thus, the effects of accent on competitor fixations can most easily be accounted for by the hypothesis that accented and deaccented noun phrases direct listeners' attention to different referential domains.

We have been assuming that eye movements to the competitor picture reflect partial hypotheses about the referent of the accented or deaccented noun phrase before the form of the noun could fully disambiguate it. Thus, accent appears to affect referential interpretation of noun phrases *immediately*, even before the complete sound pattern of the noun has been heard and processed. An alternative explanation is that listeners might have anticipated the anaphoric or nonanaphoric interpretation of the noun phrase well before hearing the target word. Previous research has shown that listeners can use prosodic contours to anticipate an upcoming accented word (e.g., Cutler, 1976). Thus, it might be argued that the effects of accent are due to listeners using information on the preceding prosodic context rather than the accent on the noun phrase. However, the data pattern we observed makes such an interpretation unlikely. Effects of accent were *not* observed on fixations to the competitor until about 300 ms after noun onset, which is precisely where competition effects have been observed between onset-overlapping words in previous research using the same paradigm (e.g., Allopenna et al., 1998; Dahan et al., 2001a). Thus, accentuation pattern and phonemic information seem to act simultaneously to constrain the referential interpretation of the noun phrase.

In addition to effects of accent, the results showed that listeners were more likely to fixate *new* pictures (i.e., pictures not previously mentioned in the discourse) at the beginning of the noun in the second instruction. One possible explanation is that listeners would expect a *given* entity to be referred to with a pronoun, rather than with a definite noun phrase. Upon hearing

the word "the," listeners might therefore be expected to fixate pictures which had not been mentioned previously, an expectation subsequently modulated by accent. To evaluate this hypothesis, we reanalyzed fixations in each discourse-context condition, beginning from the onset of the determiner. The proportion of fixations at the offset of the determiner were virtually identical to those before the determiner was encountered, making it unlikely that the bias to look at new pictures was due to the determiner. A more likely possibility is that the bias was related to visual attention. Participants preferred to look at unmentioned pictures in order to gather more information about the visual display. Regardless of the source of the bias, however, the tendency to fixate on a previously unmentioned picture after performing the first instruction did not mask or interact with the effects of accent.

To summarize, Experiment 1 demonstrated that accent placement on a definite noun phrase has immediate effects on reference resolution. A noun phrase is more likely to be interpreted anaphorically if it is deaccented than accented and is more likely to be interpreted as referring to a *new* entity if it is accented than deaccented. While this hypothesis had previously been proposed, the current results provide the first direct evidence that accent constrains the initial referential domain for definite noun phrases. Note that these effects are probabilistic: Accent introduces a bias rather than completely determining the referential domain of the definite noun phrase. The fact that the effects are probabilistic should not be surprising given that the effects we are monitoring are carried by information in the vowel before the end of the referring noun and while the phonetic information is consistent with multiple lexical alternatives in the screen-defined referential world.

To this point, we have operationalized the *given/new* distinction in terms of whether an entity has been referred to previously. However, it is important to recognize that the information status of discourse entities is sensitive to a range of factors beyond previous mention. For example, sentences containing expressions referring to recently mentioned antecedents or to topicalized entities take less time to read than sentences referring to distant antecedents or to nontopicalized antecedents (e.g., Anderson, Garrod, & Sanford, 1983; Clark & Sengul, 1979; Hudson-D'Zmura & Tanenhaus, 1998). In addition, the type of linguistic expression used to refer to an entity can also influence its information status. Sanford,

Moar, and Garrod (1988) showed that characters introduced in a text by proper name, as opposed to a definite description, are very likely to be treated as topic characters. Finally, the grammatical function played by entities can modulate their information status. For example, grammatical subjects may be perceived as more salient in the discourse than nonsubjects (Gordon & Chan, 1995; Gordon, Grosz, & Gilliom, 1993). Importantly, while the importance of these factors has been established in resolving anaphoric pronouns, their role is less clear in the resolution of other referring expressions, such as definite noun phrases (Garrod et al., 1994; Sanford et al., 1988).

Experiment 2 further investigated the interpretation of accented noun phrases in order to determine whether they are specifically interpreted as introducing *new*, unmentioned information, as the results of Experiment 1 suggest, or alternatively as referring to previously mentioned, but less prominent entities.

Experiment 2

In Experiment 1, we operationalized the *given/new* distinction by introducing an entity in the first instruction in the role of Theme (i.e., the object to be moved). Thus, the *given* entity had been mentioned in the preceding context, but was presumably also the most important or *focal* entity in the event described, whereas the *new* entity had not been mentioned in the preceding context. Entities can have other information status, such as having been mentioned before but not as the discourse focus.[4] Experiment 2 investigated how an accented definite noun phrase is interpreted with respect to the information status of potential referents (focused or nonfocused).

As in Experiment 1, participants were presented with a display composed of four pictured objects and four geometric shapes and spoken instructions to perform two consecutive actions.

---

[4] We will reserve the term "focus" or "discourse focus" to refer to the most salient or activated entity in the listener's discourse representation at any given moment (see Grosz & Sidner, 1986; Sanford & Garrod, 1981). This entity is obviously *given* and likely to be deaccented. This use of "focus" contrasts with the way it is used in the intonational work (see Birch & Clifton, 1995; Ladd, 1996; Selkirk, 1984), where it refers to the most informative part of the utterance and is likely to be accented.

Table 5
Experiment 2: Illustration for the three conditions

Conditions	First instruction	Second instruction
Theme	Put the candle below the triangle	Now put the CANDLE above the square
Goal	Put the necklace below the candle	Now put the CANDLE above the square
Theme-moveableGoal	Put the candle below the necklace	Now put the CANDLE above the square

Note. Capitals indicate which part of the second instruction the speaker was instructed to emphasize. The target word in the second instruction is underlined.

We manipulated information status by varying the thematic role used to introduce in the first instruction the entity that would be referred to in the second instruction.[5] This entity was either realized as the Theme, which for our utterances defined it as the current focus of the discourse (e.g., the entity candle in "Put the candle below the triangle"), or as the Goal (e.g., the entity candle in "Put the necklace below the candle"). In the second instruction, the same entity was realized as the Theme, using a definite noun phrase with an accented noun (e.g., "Now put the CANDLE above the square"). As in Experiment 1, the referent picture (e.g., candle) was presented along with a competitor picture (e.g., candy). The competitor picture was never mentioned in the first instruction; thus it represented a completely *new* entity. In contrast, the target picture was always mentioned in the first instruction in the current experiment, but was realized in different thematic roles. The temporary lexical ambiguity between the referent and its competitor allowed us to compare how accent was interpreted with respect to the information status of the target and its competitor.

Table 5 presents the conditions tested in Experiment 2. In the Theme condition, the referent object in the first instruction (e.g., candle) was introduced as the Theme and thus played the role of discourse focus. This condition was analogous to one of the conditions tested in Experiment 1, specifically the anaphoric-accented condition. We predicted that the same pattern would emerge, namely, that there would be more initial fixations to the competitor picture than to the target picture. In the Goal condition, the referent entity was mentioned in the first instruction but as the Goal, and thus not as the discourse focus. This condi-

tion allowed us to evaluate three hypotheses about how discourse context affects the interpretation of an accented noun phrase. First, accent might be interpreted as introducing a completely *new*, unmentioned entity. Thus, interpretation of the accented expressions would be biased toward the set of unmentioned entities. This would result in the same pattern of results as in the Theme condition described above, i.e., more fixations to the competitor picture than to the target picture early in the target word. Second, accent might be interpreted as conveying a focus shift, resulting in a bias for nonfocused entities, regardless of whether the entity had been mentioned previously. This hypothesis predicts the same proportion of fixations to the target and competitor pictures early in the target word, because neither entity was the previous focus. Finally, accent might indicate a shift of focus within the set of referents established by the previously mentioned entities. According to this hypothesis, the referential interpretation of an accented expression would favor a nonfocused entity that was previously introduced in the discourse. This hypothesis predicts that the accented target word would be interpreted as referring to the object playing the role of Goal in the first instruction and, consequently, that few fixations will be made to the competitor.

The first instruction in the Goal and Theme conditions also differed in whether the mentioned but nonfocused entity (i.e., the entity introduced in the role of Goal) was one of the four pictured objects (e.g., the candle in the Goal condition) or one of the geometric shapes (e.g., the triangle in the Theme condition). As in Experiment 1, the geometric shapes differed from the pictured objects in their visual properties (colored vs. black and white) and presentation (present in each trial vs. occurring only once in the whole experiment). Most importantly, they differed in whether they were moveable. The geometric shapes were fixed landmarks, whereas the pictures were moveable and thus were pragmatically possible Themes for a "put" instruction. Recent research by Chambers

---

[5] Although most experimental studies to date have addressed the role of structural factors in focusing (e.g., Gordon & Chan, 1995; Gordon et al., 1993), thematic considerations may also play a role (see, e.g., Stevenson & Urbanowicz, 1995).

and his colleagues using a similar setting as here has provided evidence suggesting that action-relevant pragmatic constraints can immediately affect reference resolution (Chambers, 2001; Chambers et al., 2002). In this work, participants performed actions involving real objects (e.g., ''Put the cube inside the can''). Analysis of participants' eye movements revealed that size-incompatible containers (e.g., a can too small to accommodate the cube) were excluded from the initial set of potential referents upon hearing a command, which carries the presupposition that the evoked action is possible, but not upon hearing a yes/no question (e.g., ''Can you put the cube inside the can?''). This demonstrates that the referential domain for a definite noun phrase is circumscribed by pragmatic constraints that are relevant to goal-oriented actions; potential referents that are not pragmatically plausible participants in the action are excluded from the initial referential domain. We explored whether similar effects would occur with accent by including a third condition, labeled Theme-moveableGoal, exemplified by the instruction sequence ''Put the candle below the necklace. Now put the CANDLE above the square.'' As in the Theme condition, the focused entity in the first instruction was the referent of the accented definite noun phrase in the critical instruction. However, the previously mentioned but nonfocused entity (e.g., the necklace) was moveable and thus was in the pragmatically relevant referential domain for the noun phrase in the second instruction. If an accented noun phrase is initially interpreted as referring to a previously mentioned but nonfocused object, the necklace is likely to be considered a referential candidate, as opposed to an unmentioned entity. Note that in this condition, the mentioned nonfocused entity (e.g., necklace) does not overlap at onset with the target word in the second instruction (e.g., candle). Thus an effect of accent would be reflected in (a) delayed fixations to the eventual referent and (b) a preference to look at the previously mentioned distractor compared to the unmentioned distractor.

To summarize, the referential resolution of the accented expression in the second instruction was compared across three conditions, which differed according to the characteristics of the first instruction (see Table 5). All three conditions mentioned the referent picture in the first instruction. However, the thematic role in which the referent was introduced varied (Theme condition vs. Goal condition). These two conditions allowed us to address how an accented expression is interpreted with regard to the set of already mentioned entities and the discourse focus. The third condition (Theme-moveableGoal condition) is analogous to the Theme condition, with the exception that the Goal referent is moveable. Of interest was whether the mention of a such moveable entity would affect early referential interpretation, compared to the Theme condition, where no such referential alternative was provided.

Method

Participants

Eighteen native speakers of American English were recruited at the University of Rochester and were paid a small amount for their participation. None of them had taken part in Experiment 1.

Materials

Twenty-one pairs of picturable nouns that overlapped at onset (e.g., candle–candy) were selected. One was assigned the role of target and the other of competitor. This role reflected which item of the pair was target or competitor in the second instruction of each trial. In contrast to Experiment 1, the competitor was not mentioned in either instruction. The target picture was the to-be-moved picture in the first instruction or not, depending on the condition (cf. Table 5). The mean lexical frequency, taken from Francis and Kučera (1982), was 32.6 per million for the targets and 39.8 per million for the competitors. Each of the 21 item pairs was associated with two picturable, distractor nouns. The set of materials is presented in Appendix.

In addition to these 21 experimental trials, 36 filler trials were constructed. Because all the experimental trials were anaphoric (i.e., the second instruction referred to an object mentioned in the first instruction) and involved pictures with phonetically similar names, it was important to neutralize participants' expectations about what pictured object was likely to be target in the second instruction. We thus included 20 filler trials where two of the four pictures had similar names (e.g., peanut, peacock). For 10 of these fillers, the first instruction mentioned one of the phonetically related items (e.g., peanut) and the target of the second instruction was its counterpart (e.g., peacock). Within these 10 fillers, the first instruction could refer to the phonetically similar item as the Theme and to a geometric shape as the Goal (e.g., ''Put the peanut below the triangle'') (3 fillers), or

to a phonetically similar Theme and a distractor-picture Goal (e.g., ''Put the peanut below the bottle'') (3 fillers), or to a phonetically similar Goal and a distractor-picture Theme (e.g., ''Put the bottle below the peanut'') (4 fillers). For the other 10 fillers with two phonetically similar pictures, the first instruction mentioned one of the phonetically related items but the target of the second instruction did not refer to its counterpart but rather to one of the unmentioned distractor pictures. The structure of the first instruction was varied in same way as for the other 10 fillers. We also constructed 12 other filler trials containing four phonetically unrelated pictures. All referred in the second instruction to an object that was not mentioned in the first instruction. The structure of the first instruction was also varied across these 12 fillers. Finally, 4 additional filler trials with four phonologically unrelated pictures were constructed. These fillers referred in the second instruction to an object that was mentioned in the first instruction and the structure of the first instruction was also varied.

To summarize, Experiment 2 consisted of 57 trials (21 experimental trials, 36 fillers). For 25 of these trials, the second instruction referred to a previously mentioned object (i.e., anaphoric trials), while for the 32 remaining trials, it did not (i.e., nonanaphoric trials). Of the 41 trials with two phonologically similar pictures, 21 were anaphoric and 20 were nonanaphoric. Out of these 20 nonanaphoric trials, half referred to the phonologically similar counterpart in the second instruction, while the other half referred to a distractor picture. Five of the fillers (two anaphoric, and three nonanaphoric) were presented at the beginning of the session to familiarize participants with the task and the procedure. The 228 pictures [(21 experimental + 36 filler) × 4 pictures] were selected from the same sources as those in Experiment 1 and all were black and white line drawings.

The spoken instructions were recorded by the same speaker as in Experiment 1, sampling at 22,050 Hz. As in Experiment 1, the recording script was designed to prevent the speaker from producing nonfelicitous accent patterns (e.g., accenting the target word in the second instruction when the object it referred to was the focus of the preceding instruction) by pairing the various versions of the first instruction of one experimental item with the second instruction of another experimental item (e.g., ''Put the candle below the necklace. . .now put the CARD above the triangle''). The noun phrase to accentuate in the second instruction was indicated with capitals. While the second instruction of all experimental trials carried an accent on the name of the referent picture, accent was varied among the filler trials. Sixteen of the filler trials had no accent on the name of the referent picture in second instruction, whereas the other 20 filler trials did. These 20 accented fillers were all nonanaphoric to counterbalance the 21 anaphoric experimental trials; consequently, if participants were able to anticipate the presence of an accent before hearing the accented word in the second instruction, the probability of an anaphoric or nonanaphoric referent was roughly equal.

Each utterance was edited and its prosody transcribed using the ToBI labeling system. The majority of the first instructions were characterized by the prosodic pattern described in (8) (49 of 63 cases), (9) (6 cases), or (10) (7 cases), which differ in the pitch contour on the Theme. (For one instruction, this pitch contour was L* H − L%.) The types of pitch contour were roughly balanced across the three conditions (Theme, Goal, and Theme-moveableGoal). Thus, the Theme in the first instruction was made prominent by the presence of a pitch accent and an intonational-phrase boundary. In 60 of the 63 instructions, the Goal received a low pitch accent followed by a rise and a high boundary tone (in the three remaining cases, the pitch accent was H*).

(8) Put the candle   below the square
     H*   H* H − H%   L* H − H%

(9) Put the cow   below the square
     H*   H* L − H%   L* H − H%

(10) Put the bell   above the square
     H*   L* H − H%   L* H − H%

Table 6
Experiment 2: Mean F0 peak on stressed vowels (in Hz) and duration (in ms) of various parts of the second instruction

F0 peak				Duration	
"Now"	Theme	Preposition	Goal	Theme noun	Pause following Theme (if any)
142	182	121	122	670	185 ($N = 5$)

The second instructions were characterized by the prosodic pattern exemplified in (11) and (12) in, respectively, 10 and 11 cases of 21. Additional acoustic analyses (shown in Table 6) confirm that the Theme was made prominent by the presence of a high pitch accent (sometimes preceded by sharp rise from a valley) resulting in a F0 peak in the high range for the speaker, followed by an intonational-phrase boundary and an occasional silent pause.

(11) Now put the BELL   below the triangle
     H*   L+H* L–H%   H* L–L%
(12) Now put the CANDLE   above the circle
     H*   H* L–H%   H* L–L%

Procedure

Three lists were constructed by varying in which of the three conditions the experimental items were presented (i.e., Theme, Goal, Theme-moveableGoal). Within each list, seven experimental items were assigned to each condition. Participants were randomly assigned to each list. For each list, three random orders were created and participants were randomly assigned to each order. The procedure for data collection and coding was identical to that used in Experiment 1.

Results and discussion

For two participants, a few trials were missing because of track loss (six trials in total). In addition, three trials were excluded from the analyses because some participants did not fixate the target picture in the second instruction while clicking on it and moving it with the computer mouse. Missing trials represented 2.4% of the data.

Consider first the results from the Theme condition. Recall that this condition is analogous to the anaphoric-accented condition from Experiment 1. Fig. 2 presents the mean proportion of fixations to displayed objects, starting from the onset of the accented noun. Fixations to the two distractors were averaged because neither of them had been mentioned in the first instruction and therefore their status was equivalent. The results were similar to those for Experiment 1. Participants showed an initial bias in favor of fixating the competitor and the distractors, which were all unmentioned pictures. The proportion of fixations to both the target and the competitor began to
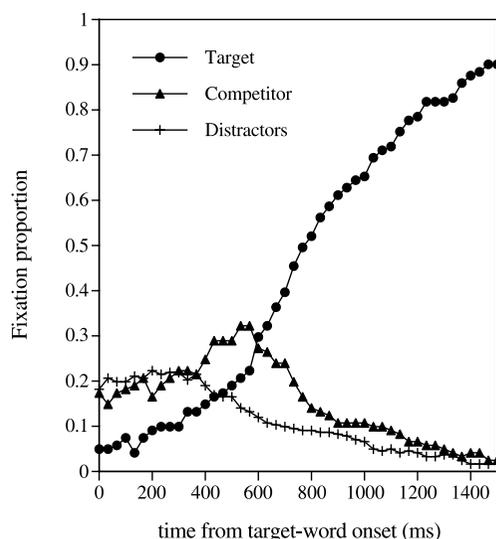

Fig. 2. Experiment 2. Fixation proportions over time to the target, the competitor, and the averaged distractors from the onset of the target word in the Theme condition.

increase about 200 ms after the onset of the target word, whereas fixations to the distractors began to decrease. Analyses on mean fixation proportions over various time windows confirmed this pattern. During the 0–300-ms window, fixation proportions to the competitor and the distractors did not differ significantly (t_1 and $t_2 < 1$). However, between 300 and 1000 ms, the proportion of fixations to the competitor was greater than to the distractors ($t_1(17) = 3.90$, $p < .005$; $t_2(20) = 2.31$, $p < .05$). Comparisons between fixation proportions to the target and competitor in Fig. 2 confirmed the initial tendency to fixate unmentioned pictures that overlapped with the target word at onset (i.e., the competitor) over mentioned overlapping pictures (i.e., the target). The advantage of the competitor over the target remained up to 600 ms, although the difference did not reach significance between 300 and 600 ms, probably due to the sharp increase in target fixations (over 0–300 ms, $t_1(17) = 2.65$, $p < .05$, $t_2(20) = 2.68$, $p < .05$; over 300–600 ms, $t_1(17) = 1.61$, $p = .12$, $t_2(20) = 1.37$, $p > .10$).

Fig. 3 presents the proportion of fixations to displayed objects in the Goal condition. As is readily apparent, the pattern of fixations to the competitor and to the target are strikingly different from those in the Theme condition, even though the second instruction was identical in each case. In the Goal condition, fixation
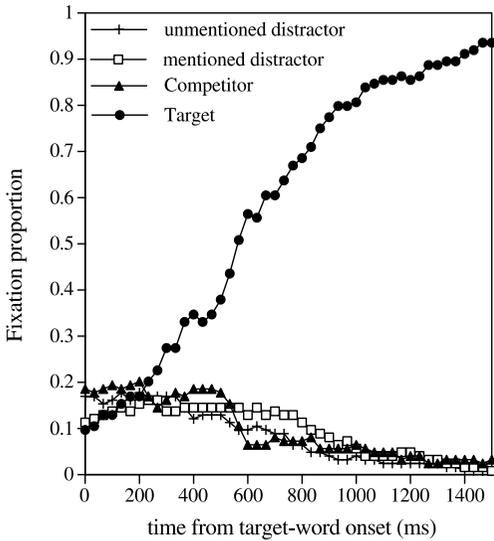
Fig. 3. Experiment 2. Fixation proportions over time to the target, the competitor, the mentioned distractor, and the unmentioned distractor from the onset of the target word in the Goal condition.

proportions to the target began to increase around 200 ms after word onset, and fixation proportions to all other pictures decreased correspondingly. Fixations to the competitor and the unmentioned distractor remained similar over time, despite the fact that the competitor's name overlapped with the initial sounds of the target word. Fixation proportions to the unmentioned and mentioned distractors were similar. The initial bias in fixating unmentioned pictures (i.e., the competitor and the unmentioned distractor) over mentioned ones (i.e., the target and the mentioned distractor) was present but small. Statistical analyses confirmed that fixation proportions to the competitor did not significantly differ from those to the unmentioned distractor, nor from those to the mentioned distractor, for any of the time windows. This indicates that the accented noun phrase was immediately interpreted as referring to the entity that had been realized as the Goal in the preceding instruction, i.e., the target picture itself. This suggests that an accented noun phrase is not preferentially interpreted as introducing a completely new entity, that is, an entity that has not been previously mentioned. If this were the case, the competitor would have been initially favored over the target as the referent of the accented noun phrase, as this was the case in the Theme condition. Rather, the results suggest that accent is interpreted as indicating a shift of attention away from the focused entity, under the assumption that the entity introduced as the Theme in the first instruction was the local focus. However, this does not fully account for the fixation pattern in the Goal condition. Indeed, neither the target nor the competitor were focused, and both overlapped with the target word at onset. If accent were interpreted as introducing an entity other than the focused entity, both the target and the competitor would have been equally considered as the referent of the accented noun early on. Instead, we observed an immediate and sharp increase in target fixations and no difference in fixations to the competitor compared to the unrelated distractors. This suggests that the preferred referent for the accented noun phrase was the previously mentioned nonfocused entity, namely, the target.

The Theme-moveableGoal condition provides a further test of the hypothesis that an accented noun phrase is preferentially interpreted as referring to a mentioned but nonfocused entity. Recall that in this condition, unlike in the Theme and Goal conditions, a moveable nonfocused distractor entity was mentioned in the first instruction. If accent is preferentially interpreted as signaling reference to a previously mentioned nonfocused entity, then the pattern of fixations should reveal a tendency for interpreting the accented noun phrase as referring to this entity, compared to the unmentioned nonfocused distractor. Fig. 4 presents the proportion of fixations to the target, the competitor, the mentioned distractor (which was the Goal in the first instruction, e.g., <u>necklace</u> in "Put the candle below the necklace"), and the unmentioned distractor, beginning from the onset of the target word of the second instruction. As in the other conditions, an early bias in fixating unmentioned over mentioned pictures was present, except for the mentioned distractor, which was fixated as often as the unmentioned distractor. At 300 ms and onward, the proportion of fixations to the competitor was greater than that to the unmentioned distractor, but similar to that of the mentioned distractor. Furthermore, the proportion of fixations to the mentioned distractor was greater than that to the unmentioned distractor. Comparisons over the 300–1000 ms time window confirmed this pattern On average, the competitor was fixated more than the unmentioned distractor ($t_1(17) = 2.42$, $p < .05$, $t_2(20) = 1.97$, $p = .06$) and the mentioned distractor was fixated more than the unmentioned distractor ($t_1(17) = 3.65$, $p < .005$, $t_2(20) = 2.10$, $p < .05$). Mean fixation proportions to the competitor and the mentioned
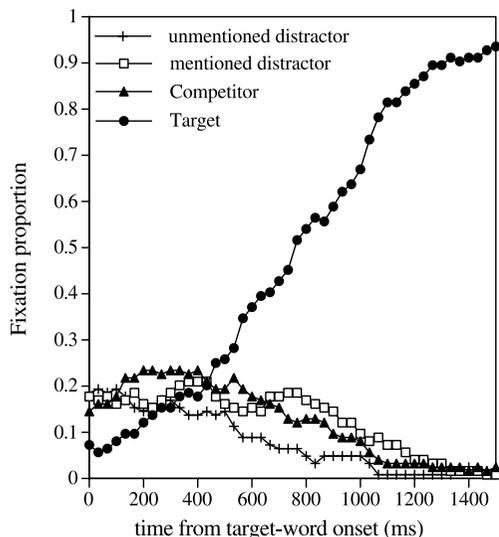
Fig. 4. Experiment 2. Fixation proportions over time to the target, the competitor, the mentioned distractor, and the unmentioned distractor from the onset of the target word in the Theme-moveableGoal condition.

distractor did not differ significantly over any time window. This confirms that accent directed the listener's attention to mentioned but nonfocused entities, even in the absence of direct lexical support for this candidate.

In addition to the independent comparisons for each condition, we carried out comparisons across conditions. Fig. 5 presents the proportions of fixations to the target (Fig. 5A) and to the competitor (Fig. 5B), from the onset of the accented noun of the second instruction, in all three conditions. While starting at approximately equivalent values for all three conditions, the fixation proportion to the competitor increased in the Theme condition as the target word unfolded over time, while it decreased steadily in the Goal condition, with the Theme-moveableGoal condition showing intermediate fixation proportions. One-way ANOVAs conducted on mean fixation proportions to the competitor over various time windows confirmed an effect of conditions between 300 and 1000 ms after target-word onset (from 0 to 300 ms, $F_1(2, 34) < 1$, $F_2(2, 40) < 1$; from 300 to 1000 ms, $F_1(2, 34) = 6.63$, $p < .005$, $MSE = .0081$, $F_2(2, 40) = 4.24$, $p < .05$, $MSE = .0148$). Pairwise comparisons showed greater fixation proportions to the competitor in the Theme condition than in the Goal condition over the 300–1000-ms window ($t_1(17) = 3.32$, $p < .005$; $t_2(20) = 3.50$, $p < .005$). Pairwise comparisons
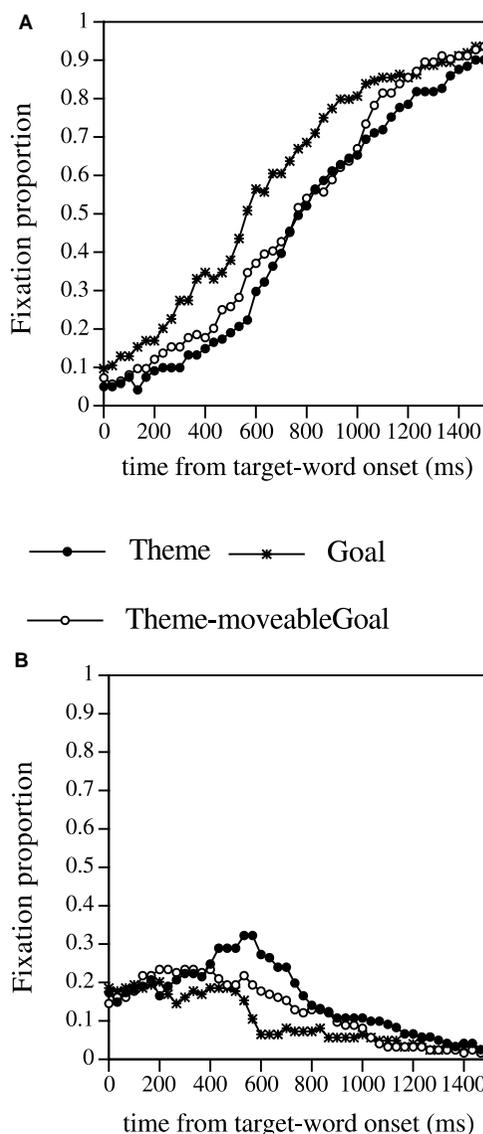

Fig. 5. Experiment 2. Fixation proportions over time to the target (A) and to the competitor (B) from the onset of the target word in the Theme, Goal, and Theme-moveableGoal conditions.

between the Theme and Theme-moveableGoal conditions showed more fixations in the Theme condition, but this trend did not reach significance over the 300–1000-ms time window ($t_1(17) = 1.52$, $p = .15$; $t_2(20) = 1.15$, $p = .27$).

Fixation proportions to the target (Fig. 5A) increased fastest in the Goal condition and increased the most slowly in the Theme condition. However, the Theme-moveableGoal condition

tended to pattern with the Theme condition. One-way ANOVAs on mean target fixation proportions over various time windows revealed an effect of conditions. The effect was already present in the 0–300-ms window, due to a slightly faster rise in the Goal condition than in the other two conditions. The effect remained significant from 300 to 1000 ms after noun onset (from 0 to 300 ms, $F_1(2, 34) = 4.67$, $p < .05$, $MSE = .0081$, $F_2(2, 40) = 4.02$, $p < .05$, $MSE = .0105$; from 300 to 1000 ms, $F_1(2, 34) = 13.16$, $p < .0001$, $MSE = .0142$, $F_2(2, 40) = 10.18$, $p < .0001$, $MSE = .0194$). Pairwise comparisons between conditions indicated that the mean fixation proportion was significantly higher in the Goal condition than in the Theme condition (over 300–1000 ms, $t_1(17) = 4.27$, $p < .001$, $t_2(20) = 3.75$, $p < .001$). No reliable difference was found between the Theme and the Theme-moveableGoal conditions.

The pattern of results for the combined conditions reinforces the conclusions from the individual conditions. Most strikingly, fixations to targets were faster in the Goal condition than in the Theme and the Theme-moveableGoal conditions. Since the Goal condition was the only condition in which the referent was the mentioned but not focused entity, this result provides additional support for the conclusion that an accented noun phrase is preferentially interpreted as referring to a nonfocused but mentioned entity. Furthermore, the trend toward more fixations to competitors in the Theme condition compared to the Theme-moveableGoal condition provides support for the hypothesis that pragmatic constraints circumscribe initial referential domains.

General discussion

The current studies make several important contributions to our understanding of how accent affects reference resolution. First, they provide striking evidence that listeners can use accentual information in tandem with information from the discourse model and phonetic input remarkably quickly during reference resolution. Within many theoretical frameworks, pitch accents and phrasal tones are used to construct the intonational meaning of an utterance and provide crucial information about the relationships among utterances in a discourse (e.g., Pierrehumbert & Hirschberg, 1990). The current results suggest that pitch accents, which form the basis for this in-

formation, can be rapidly used during real-time processing. Recall that the effects of accent observed in the current experiments reflected processing within the brief time interval during which an unfolding spoken word was temporarily ambiguous between two task-relevant referents. Our results indicate that information from the prosodic stream can be rapidly integrated with segmental information, as demonstrated by the fact that accent combined with segmental information to influence the relative proportion of fixations to temporarily ambiguous lexical competitors. Moreover, information about accent can be processed in parallel with segmental information, as demonstrated by effects of accent on fixations to mentioned and nonmentioned distractors in Experiment 2. In the Theme-moveableGoal condition, accent on the noun led to more fixations toward the previously mentioned distractor (introduced as the Goal in the first instruction) than toward the unmentioned distractor, even though there was no supporting bottom-up segmental evidence for this referential candidate. This suggests that prosodically conveyed discourse factors may have a stronger influence on lexical access than has previously been assumed (see Kjelgaard & Speer, 1999, for a similar proposal). It will be important to examine exactly how prosodic and segmental constraints are integrated during lexical access in future research.

Second, the current studies provide information about how listeners used accent information to circumscribe referential domains. While the hypothesis that a deaccented noun phrase biases interpretation in favor of a *given* entity had previously been proposed, Experiment 1 provided the first direct evidence that listeners preferentially interpret an unfolding deaccented noun phrase as coreferential with a mentioned and focused noun entity compared to a previously unmentioned discourse entity. Experiment 2 further demonstrated that under some discourse conditions, an accented noun phrase is preferentially interpreted as referring to a previously mentioned but nonfocused discourse entity rather than to an unmentioned discourse entity, as hypothesized by Terken and Nooteboom (1987). Thus, the discourse status of an entity (i.e., focused or nonfocused), rather than just its previous mention, constrains the referential interpretation of an accented noun phrase. It is important to note that we are not claiming that an accented noun phrase is preferentially interpreted as referring to a previously mentioned but nonfocused discourse en-

tity under *all* discourse conditions. Indeed, pragmatic constraints can modulate whether a mentioned nonfocused entity is considered as a potential referent for an upcoming noun phrase. The differences in target and competitor fixations between the Theme and Theme-moveableGoal conditions, in which the Goal of the first instruction corresponded to a fixed landmark or to a moveable object, respectively, suggest the influence of such pragmatic constraints. Furthermore, the interpretation of accented expressions as referring to a previously mentioned nonfocused entity may be modulated by intonational information that helps establish links between utterances in a discourse. In the current studies, the utterance containing the accented noun phrase followed an utterance that ended with a high boundary tone. As discussed earlier, Pierrehumbert and Hirschberg (1990) proposed that an utterance ending in a high boundary tone is interpreted in the context of the following utterance. If this conjecture is correct, the high boundary tone might have encouraged listeners to interpret the definite noun phrase in the next utterance as referring to one of the previously mentioned entities. Moreover, the second instruction began with "now," a so-called cue word that can signal that an utterance is to be interpreted as part of the same discourse segment as the previous utterance (Hirschberg & Litman, 1987). These intonational and lexical cues may have biased the interpretation of the noun phrase as referring to a previously mentioned entity. If the noun phrase was deaccented, it was preferentially interpreted as referring to the discourse focus; if the noun phrase was accented, this strongly disfavored the focus and favored a mentioned but nonfocused entity. Some further support for this bias comes from examining the pattern of fixation proportions to the target picture and the asymmetry it revealed in Experiment 1.

In the anaphoric condition, target fixations rose faster in the deaccented than in the accented condition because the second instruction referred to the discourse focus. However, in the nonanaphoric condition, target fixations did not differ significantly between the accented and deaccented conditions. We argue that target fixations did not rise faster in the accented than in the deaccented condition, as predicted by the hypothesis that an accented noun phrase is interpreted as referring to a new unmentioned entity, because intonational and lexical cues led people to interpret an ac-

cented noun phrase as referring to a mentioned but nonfocused entity. Note that the difference in competitor fixations observed in the nonanaphoric conditions reflected the preference for (wrongly) interpreting the deaccented noun phrase as referring to the discourse focus, here the competitor. In future research, it will be important to manipulate factors such as boundary tones and intonation patterns on cue words such as "now" to determine whether they trigger a possible interpretation bias and how they interact with accent during reference resolution. It will also be important to examine how the details of pitch accents on linguistic antecedents interact with the accentual status of the anaphor. For instance, the Goal of the first instruction in Experiment 2 almost always received a low pitch accent, while the Theme received a high pitch accent. This may have caused the listeners to interpret the following accented noun phrase as referring to that entity by preestablishing a contrastive set between the two entities.

The current results are incompatible with simple hypotheses such as the hypothesis that accented noun phrases always refer to new discourse entities or the hypothesis that deaccented noun phrases refer to the most recently mentioned entity. Rather, the results suggest that the use of accent in reference resolution, along with use of other prosodic factors in real-time comprehension, involves a complex, but systematic, interaction between information provided by the utterance and information provided by the goal structure that helps modulate the listener's attentional and intentional states during discourse, as argued by Grosz and Sidner (1986) among many others. Thus, a research strategy that combines discourse manipulations with natural tasks that have a clearly articulated goal structure is particularly well suited to examining how prosody is used in real-time language comprehension.

Appendix

Target	Competitor	Distractor 1	Distractor 2
Experimental trials for Experiment 1			
bell (23)	bed (139)	sock	knife
cage (11)	cake (16)	glasses	headphones
candle (23)	candy (18)	necklace	pear
card (61)	cart (9)	tiger	whistle
cat (42)	cap (22)	leaf	wheel
chain (60)	chair (89)	buffalo	trumpet
clown (6)	cloud (64)	shark	grill
comb (6)	coat (52)	onion	bird
cow (46)	couch (2)	pumpkin	jar
flask (5)	flag (18)	deer	iron
hammock (5)	hammer (6)	fly	avocado
horn (33)	horse (233)	sun	plug
lamb (14)	lamp (24)	orange	gun
penguin (0)	pencil (38)	grapes	binoculars
picture (277)	pitcher (29)	tie	kite
plate (44)	plane (2)	scissors	well
rooster (6)	ruler (13)	basket	ladybug
sandal (5)	sandwich (13)	globe	telephone
scale (62)	skate (1)	moon	strawberry
sheep (24)	shield (8)	kettle	duck
snail (3)	snake (70)	zipper	rocking chair
trunk (13)	truck (80)	butterfly	watch
turtle (9)	turkey (4)	ring	box
window (172)	windmill (1)	pipe	pen
Experimental trials for Experiment 2			
bell (23)	bed (139)	sock	knife
cage (11)	cake (16)	glasses	headphones
candle (23)	candy (18)	necklace	pear
card (61)	cart (9)	tiger	whistle
chain (60)	chair (89)	buffalo	trumpet
cloud (64)	clown (6)	shark	grill
comb (6)	coat (52)	onion	bird
cow (46)	couch (2)	pumpkin	jar
flask (5)	flag (18)	deer	iron
horn (33)	horse (233)	sun	plug
lamb (14)	lamp (24)	orange	gun
penguin (0)	pencil (38)	grapes	binoculars
plate (44)	plane (2)	scissors	well
rooster (6)	ruler (13)	basket	ladybug
sandal (5)	sandwich (13)	globe	telephone
scale (62)	skate (1)	moon	strawberry
sheep (24)	shield (8)	kettle	duck
snail (3)	snake (70)	zipper	rocking chair
trunk (13)	truck (80)	butterfly	watch
turtle (9)	turkey (4)	ring	box
window (172)	windmill (1)	pipe	pen

Note. Lexical frequencies, as reported in Francis and Kučera (1982), are indicated between parentheses. A frequency of 0 indicates that no frequency was available.

References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: evidence for continuous mapping models. *Journal of Memory and Language, 38*, 419–439.

Anderson, A., Garrod, S., & Sanford, A. J. (1983). The accessibility of pronominal antecedents as a function of episode shifts in narrative text. *Quarterly Journal of Experimental Psychology, 35a*, 427–440.

Arnold, J. E., Eisenband, J. G., Brown-Schmidt, S., & Trueswell, J. C. (2000). The rapid use of gender information: evidence of the time course of pronoun resolution from eye-tracking. *Cognition, 76*, B13–B26.

Bard, E. C., & Aylett, M. P. (1999). The dissociation of de-accenting, givenness, and syntactic role in spontaneous speech. In *Proceedings of the International Conference on Phonetic Sciences, San Francisco* (pp. 1753–1756).

Bard, E. G., Cooper, L., Kowtko, J., & Brew, C. (1991). Psycholinguistic studies on incremental recognition of speech: A revised and extended introduction to the messy and the sticky. DYANA Deliverable R1.3B, University of Edinburgh.

Beckman, M. E., & Hirschberg, J. (1994). *The ToBI annotation conventions*. Columbus, OH: Ohio State University.

Birch, S., & Clifton, C., Jr. (1995). Focus, accent, and argument structure: effects on language comprehension. *Language and Speech, 38*, 365–391.

Bock, J. K., & Mazzella, J. R. (1983). Intonational marking of given and new information: some consequences for comprehension. *Memory & Cognition, 11*, 64–76.

Bolinger, D. (1972). Accent is predictable (if you're a mind-reader). *Language, 58*, 505–533.

Carlson, K., Clifton, C., Jr.,, & Frazier, L. (2001). Prosodic boundaries in adjunct attachment. *Journal of Memory and Language, 45*, 58–81.

Chambers, C. G. (2001). The dynamic construction of referential domains. Unpublished Dissertation, University of Rochester.

Chambers, C. G., Tanenhaus, M. K., Eberhard, K. M., Filip, H., & Carlson, G. N. (2002). Circumscribing referential domains during real-time language comprehension. *Journal of Memory and Language, 47*, 30–49.

Clark, H. H., & Haviland, S. E. (1977). Comprehension and the given- new contract. In R. O. Freedle (Ed.), *Discourse production and comprehension* (pp. 1–40). Norwood: Ablex.

Clark, H. H., & Sengul, C. J. (1979). In search of referents for nouns and pronouns. *Memory & Cognition, 7*, 35–41.

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language. A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology, 6*, 84–107.

Cutler, A. (1976). Phoneme-monitoring reaction time as a function of preceding intonation contour. *Perception & Psychophysics, 20*, 55–60.

Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: a literature review. *Language and Speech, 40*, 141–201.

Cycowicz, Y. M., Friedman, D., Rothstein, M., & Snodgrass, J. G. (1997). Picture naming by young children: norms for name agreement, familiarity, and visual complexity. *Journal of Experimental Child Psychology, 65*, 171–237.

Dahan, D., Magnuson, J. S., & Tanenhaus, M. K. (2001a). Time course of frequency effects in spoken-word recognition: evidence from eye movements. *Cognitive Psychology, 42*, 317–367.

Dahan, D., Magnuson, J. S., Tanenhaus, M. K., & Hogan, E. M. (2001b). Subcategorical mismatches and the time course of lexical access: evidence for lexical competition. *Language and Cognitive Processes, 16*, 507–534.

Dahan, D., Swingley, D., Tanenhaus, M. K., & Magnuson, J. S. (2000). Linguistic gender and spoken-word recognition in French. *Journal of Memory and Language, 42*, 465–480.

Eberhard, K. M., Spivey-Knowlton, M. J., Sedivy, J. C., & Tanenhaus, M. K. (1995). Eye movements as a window into real time spoken language comprehension in natural contexts. *Journal of Psycholinguistic Research, 24*, 409–436.

Fischer, B. (1992). Saccadic reaction time: implications for reading, dyslexia and visual cognition. In K. Rayner (Ed.), *Eye movements and visual cognition: Scene perception and reading* (pp. 31–45). New York: Springer-Verlag.

Fowler, C. A., & Housum, J. (1987). Talkers' signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language, 26*, 489–504.

Francis, W. N., & Kučera, H. (1982). *Frequency analysis of English usage. Lexicon and grammar*. Boston: Houghton Mifflin Company.

Garrod, S., Freudenthal, D., & Boyle, E. (1994). The role of different types of anaphor in the on-line resolution of sentences in a discourse. *Journal of Memory and Language, 33*, 39–68.

Garnham, A. (2001). *Mental models and the interpretation of anaphora*. Philadelphia: Psychology Press/Taylor & Francis.

Gordon, P. C., & Chan, D. (1995). Pronouns, passives, and discourse coherence. *Journal of Memory and Language, 34*, 216–231.

Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science, 17*, 311–347.

Grosz, B. J., & Sidner, C. L. (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics, 12*, 175–204.

Halliday, M. A. K. (1967). Notes on transitivity and theme in English: II. *Journal of Linguistics, 3*, 199–244.

Hirschberg, J., & Litman, D. (1987). Now lets talk about now: identifying cue phrases intonationally. *Proceedings of the Association for Computational Linguistics, 25*, 163–171.

Hudson-D'Zmura, S., & Tanenhaus, M. K. (1998). Assigning antecedents to ambiguous pronouns: the role of the center of attention as the default assignment. In M. A. Walker, A. K. Joshi, & E. F. Prince (Eds.), *Centering theory in discourse* (pp. 199–226). Oxford: Clarendon Press.

Kjelgaard, M. M., & Speer, S. R. (1999). Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. *Journal of Memory and Language, 40*, 153–194.

Ladd, D. R. (1980). *The structure of intonational meaning*. Bloomington: Indiana University Press.

Ladd, D. R. (1996). *Intonational phonology*. Cambridge: Cambridge University Press.

Marslen-Wilson, W. (1987). Functional parallelism in spoken word-recognition. *Cognition, 25*, 71–102.

Nooteboom, S. G., & Kruyt, J. G. (1987). Accent, focus distribution, and the perceived distribution of given and new information: an experiment. *Journal of the Acoustical Society of America, 82*, 1512–1524.

Pierrehumbert, J., & Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In P. R. Cohen, J. Morgan, & M. E. Pollack (Eds.), *Intentions in communication* (pp. 271–311). Cambridge, MA: MIT Press.

Prince, E. F. (1981). Toward a taxonomy of given-new information. In P. Cole (Ed.), *Radical pragmatics* (pp. 223–255). New York: Academic Press.

Sanford, A. J., Moar, K., & Garrod, S. (1988). Proper names as controllers of discourse focus. *Language and Speech, 31*, 43–56.

Sanford, A. J., & Garrod, S. (1981). *Understanding written language: Explorations in comprehension beyond the sentence*. Chichester: Wiley.

Saslow, M. G. (1967). Latency for saccadic eye movement. *Journal of the Optical Society of America, 57*, 1030–1033.

Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition, 71*, 109–147.

Selkirk, E. O. (1984). *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.

Silverman, K., Beckman, M. E., Pitrelli, J., Ostendorf, M., Wightman, C., Price, P., Pierrehumbert, J., & Hirschberg, J. (1992). ToBI: a standard for labeling English prosody. Paper presented at the Second International Conference on Spoken Language Processing, Banff, Canada.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 174–215.

Stevenson, R. J., & Urbanowicz, A. J. (1995). Structural focusing, thematic focusing, and the comprehension of pronouns. In J. Moore, & J. Lehman (Eds.), *Proceedings of the seventeenth annual meeting of the cognitive science society* (pp. 328–333). Mahwah, NJ: Erlbaum.

Tanenhaus, M. K., Magnuson, J. S., Dahan, D., & Chambers, C. G. (2000). Eye movements and lexical access in spoken language comprehension: evaluating a linking hypothesis between fixations and linguistic processing. *Journal of Psycholinguistic Research, 29*, 557–580.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*, 1632–1634.

Tanenhaus, M. K., Spivey-Knowlton, M., Eberhard, K. M., & Sedivy, J. C. (1996). Using eye-movements to study spoken language comprehension: evidence for visually mediated incremental interpretation. In T. Inui, & J. McClelland (Eds.), *Attention performance XVI: Integration in perception and communication* (pp. 457–478). Cambridge, MA: MIT Press.

Terken, J., & Hirschberg, J. (1994). Deaccentuation of words representing 'given' information: effects of persistence of grammatical function and surface position. *Language and Speech, 37*, 125–145.

Terken, J., & Nooteboom, S. G. (1987). Opposite effects of accentuation and deaccentuation on verification latencies for given and new information. *Language and Cognitive Processes, 2*, 145–163.

Turk, A. E., & Sawusch, J. R. (1997). The domain of accentual lengthening in American English. *Journal of Phonetics, 25*, 25–41.

van Donselaar, W., & Lentz, J. (1994). The function of sentence accents and given/new information in speech processing: different strategies for normal-hearing and hearing-impaired listeners? *Language and Speech, 37*, 375–391.

Viviani, P. (1990). Eye movements in visual search: cognitive, perceptual, and motor control aspects. In E. Kowler (Ed.), *Eye movements and their role in visual and cognitive processes* (pp. 353–393). Amsterdam: Elsevier.

Vonk, W., Hustinx, L. G. M. M., & Simons, W. H. G. (1992). The use of referential expressions in structuring discourse. *Language and Cognitive Processes, 7*, 301–333.

Zwitserlood, P. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition, 32*, 25–26.