Journal of Memory and Language 63 (2010) 367–386

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

# Recognition memory reveals just how CONTRASTIVE contrastive accenting really is

Scott H. Fraundorf *, Duane G. Watson, Aaron S. Benjamin

*Department of Psychology, University of Illinois at Urbana-Champaign, Champaign, IL 61820, United States*

ARTICLE INFO

*Article history:*
Received 26 November 2009
revision received 18 June 2010
Available online 23 July 2010

*Keywords:*
Discourse
Language comprehension
Pitch accenting
Focus
Recognition memory

ABSTRACT

The effects of pitch accenting on memory were investigated in three experiments. Participants listened to short recorded discourses that contained contrast sets with two items (e.g. *British scientists* and *French scientists*); a continuation specified one item from the set. Pitch accenting on the critical word in the continuation was manipulated between non-contrastive ($H^*$ in the ToBI system) and contrastive ($L + H^*$). On subsequent recognition memory tests, the $L + H^*$ accent increased hits to correct statements and correct rejections of the contrast item (Experiments 1–3), but did not impair memory for other parts of the discourse (Experiment 2). $L + H^*$ also did not facilitate correct rejections of lures not in the contrast set (Experiment 3), indicating that contrastive accents do not simply strengthen the representation of the target item. These results suggest comprehenders use pitch accenting to encode and update information about multiple elements in a contrast set.



## Introduction

The present study investigates the mechanisms by which pitch accents affect memory in language comprehension. Theories of intonation have proposed that prominent words are marked with *pitch accents*, realized acoustically as changes in fundamental frequency ($F_0$), increased duration, and greater intensity (see Ladd (1996), for a review). Pitch accents have been argued to reflect discourse structure; for instance, referents that are new to a discourse, that are unpredictable, or that are in focus are more likely to be produced with a pitch accent. In turn, pitch accents influence both online language comprehension and offline memory and meta-linguistic judgments (for review, see Cutler, Dahan, and van Donselaar (1997) and Wagner and Watson (in press)).

However, the specific mechanisms by which pitch accenting affects processing remain to be determined. In the present paper, we compare identification, granularity, and contrast-based accounts of the effects of pitch accents. We present evidence that differences between pitch accents may affect discourse comprehension through their effects on a listener's representation of contrast items, as revealed by later memory for the discourse.

## How do pitch accents affect processing?

One possibility is that pitch accenting serves mainly to facilitate initial identification of the discourse status of referents. For instance, Bock and Mazzella (1983) found that listeners are faster to report comprehending a discourse when new information is accented and given information is de-accented. Using the visual world paradigm (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995), Dahan, Tanenhaus, and Chambers (2002) found that listeners rapidly interpreted pitch accented words as referring to new referents, and de-accented words as referring to given referents. Results of this kind have been interpreted by Bock and Mazzella in terms of the given-new contract proposed by Haviland and Clark (1974). In this model, discourse comprehension requires that listeners first identify which information is new and which is given. Denoting

* Corresponding author.
*E-mail address:* sfraund2@illinois.edu (S.H. Fraundorf).

0749-596X/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jml.2010.06.004

discourse status with pitch accents (or the lack thereof) may allow listeners to complete this process more quickly. In a strong version of this view, pitch accents may speed the construction of a discourse representation, but once the discourse status of all referents has been fixed, pitch accenting does not alter the nature of the resulting representation.

Indeed, in some cases, pitch accenting has been observed to affect only early stages of language comprehension. Almor and Eimas (2008) presented participants with short, recorded discourses such as (1a) and (1b) below. (Capital letters indicate contrastive accents throughout.)

(1a) It was the ROBIN that ate the apple. The bird seemed very satisfied.
(1b) What the robin ate was the APPLE. The bird seemed very satisfied.

In (1a), the *it*-cleft placed *the robin* in linguistic focus (i.e., as the answer to the implicit question "Who ate the apple?"), and *robin* received a contrastive pitch accent. In (1b), the cleft places *the apple* in focus, and *apple* rather than *robin* is pitch accented. Almor and Eimas found that placing *the robin* in focus sped an immediate lexical decision—cued by a tone—to the anaphor *the bird*. However, clefting and pitch accenting had no effect on performance in a later cued recall task in which participants were asked, *Who ate the fruit*? Almor and Eimas interpreted these results as suggesting that pitch accents and other manipulations of focus speed initial processing but have no consequence for long-term memory for typical discourses.

But other theories propose that pitch accents can also modulate later memory for a discourse. For example, Sanford, Sanford, Molle, and Emmott (2006) proposed a granularity account in which pitch accenting and other manipulations of linguistic focus increase the specificity of semantic representations. Supporting this, Sanford et al. found that focus modulates performance in a change detection task (Simons & Levin, 1997). In the task used by Sanford et al., participants were twice presented with short, recorded discourses and asked to detect whether any of the words changed between the first and second presentations. In a broad focus condition such as (2a), the target word was spoken with a non-contrastive accent (an $H^*$ accent in the ToBI system; reviewed below) and the discourse established an implicit question (*What happened?*) that is answered by the entire second sentence. In a narrow focus condition such as (2b), the target word was spoken with a contrastive accent ($L + H^*$ in ToBI) and the implicit question specifically questioned which money had stolen.

(2a) They wanted to find out what had happened. The money from the wallet had gone missing. Thefts in the area were becoming all too common.
(2b) They wanted to find out which money had been stolen. The money from the WALLET had gone missing. Thefts in the area were becoming all too common.

Sanford et al. found that the ability of participants to detect a one-word alteration such as changing *wallet* to *purse* was superior in the narrow focus condition than in the broad focus condition. This difference indicates that manipulations of focus can affect not only initial processing time, as predicted by an identification account, but also subsequent representations of a discourse. However, because both the pitch accents and prior discourse context differed between conditions in their experiment, it is not clear which of the two drove this effect.

Sanford et al. argued that the difference in performance between (2a) and (2b) is consistent with an account in which focus modulates the specificity of semantic representations. For instance, the broad focus in (2a) might lead listeners to encode *wallet* at a superordinate level (*accessory*), which would be insufficient for detecting the change to *purse*, whereas the narrow focus in (2b) leads to a more specific encoding of *wallet*. This theory predicts that the focus manipulation should matter less in distinguishing two words that are members of separate semantic categories, such as *wallet* and *bank*, because encoding at even a superordinate level should be sufficient for this distinction. And, as predicted, the magnitude of the focus effect was larger when the target word was changed to a word in the same semantic category (*wallet* to *purse*) than one in a different one (*wallet* to *bank*).

The granularity account is also consistent with the effects of other manipulations of linguistic focus. Performance on a change detection task is also enhanced by pseudoclefting (Sturt, Sanford, Stewart, & Dawydiak, 2004) and by italicization of written words (Sanford et al., 2006), which, like pitch accenting, have been argued to be related to linguistic focus (McAteer, 1992). Using *it*-clefts, Birch and Garnsey (1995) tested the effects of focus on priming of word naming. They visually presented subjects with sentences such as (3a) or (3b). In (3a), the critical word *present* is not in focus, but in (3b) the cleft construction brings *present* into focus.

(3a) The card on the present was addressed to the little girl's mother.
(3b) It was the present that had surprised and thrilled the girl's mother.

Immediately after the prime sentence, participants completed a naming task. The target in the naming task was phonologically or semantically related to the prior critical word but was not a discourse alternative. For example, after *present*, participants might have to name the semantically related word *gift*. These related items were slower to be named when the critical word had been in focus. For example, naming *gift* after *present* took longer when *present* appeared in focus than when it did not. These data are consistent with a granularity account: when a word like *present* is focused, readers may represent it more specifically and that leads to less activation of related words like *gift*. Crucially, the granularity account predicts that any benefit to discourse representation occurs because the target word itself is represented more specifically.

Effects of pitch accenting can also be explained by what we will call a contrast representation account: differences in pitch accents might also modulate the representation of a *contrast* item, not just the accented word itself. For

instance, the L + H* pitch accent on *wallet* in the Sanford et al. (2006) study might have led participants to consider *purse* as a possible contrast item for *wallet*, which could facilitate their later awareness of this change. Evidence for this position comes from Braun and Tagliapietra (in press), who compared the effects of pitch accenting types on two types of words: words that were semantically related to a prime but not necessarily contrastive (similar to Birch & Garnsey's items), and words that specifically contrasted with a prime. Participants listened to utterances like (4) and then had to make a lexical decision for a visually presented word.

(4) Our neighbors assembled an antenna.

Braun and Tagliapietra manipulated the type of pitch accent placed on the last word of the spoken utterance. A contrastive accent on the word facilitated lexical decision times to the visual target, but only if the visual target contrasted with the word that received the contrastive accent. For instance, the contrastive accent facilitated responses to *dish*, which contrasts with *antenna*, but not to *television*, which is semantically related to *antenna* but not contrastive. This facilitation is not predicted by the granularity account, which proposes that contrastive accents lead to a more specific representation of the target that should have been less apt to spread activation to the contrastive word. Instead, this result suggests that pitch accents may influence representations of a contrast item instead of, or in addition to, the accented item itself. Importantly, this result also suggests that the effects of contrastive accents may differ for items inside and outside of a contrast set.

These three accounts of the mechanisms of pitch accent effects have not yet been directly pitted against one another. Comparisons are further complicated by the fact that, although clefting and contrastive accents have both been argued to imply contrast, in many experiments the contrast set has not been explicitly specified. For instance, in (1a), *the robin* is implicitly contrasted with other agents that could have eaten the apple, but the specific set of such agents is not specified. The contrast representation account predicts that comprehenders should encode information about this contrast set and thus treat referents inside the contrast set differently from those outside it. However, testing this prediction is difficult without an independent measure of what participants understand to be inside or outside of the contrast set.

In the present study, we investigate the mechanisms underlying the processing of pitch accents by assessing recognition memory for facts about discourses in which the contrast set was explicitly specified. This technique allows us to test predictions of the accounts regarding the effects of pitch accents on hits to correct discourse information, correct rejections of false information that was part of the contrast set, and correct rejections of false information *not* part of the contrast set. We compare two of the most studied pitch accent types in American English: the accents labeled H* and L + H* in the ToBI framework of prosodic transcription for American English (Beckman & Elam, 1997; Pitrelli, Beckman, & Hirschberg, 1994; Silverman et al., 1992).

## Pitch accent types

In the ToBI system, the H* accent represents a high pitch target with fundamental frequency ($F_0$) high in the speaker's range. Pierrehumbert and Hirschberg (1990) argued that this accent type is associated with information that is discourse-new but not contrastive. Some experimental work has supported this account. For example, Chen, den Os, and de Ruiter (2007) compared listeners' interpretation of several types of pitch accents in British English using the Transcription of Dutch Intonation system (ToDI; Gussenhoven, 2005). Using the visual world paradigm, Chen et al. found that one of two ToDI pitch accent types corresponding to American English H* cued attention to new referents, although results were mixed for the other accent type. However, other experiments (e.g. Watson, Gunlogson, & Tanenhaus, 2008) have found that H* may be interpreted by listeners more broadly to refer to any information that had been less salient, whether it is discourse-new or not.

ToBI distinguishes the H* accent from the L + H* accent, which consists of an initial low pitch followed by a sharp rise to a high target on the accented syllable. Pierrehumbert and Hirschberg (1990) argued that the L + H* accent is associated with information that is contrastive. For instance, in (5b) below, *projector* is simply new information and would likely receive an H* accent in Pierrehumbert and Hirschberg's account. In (6b), however, *projector* contrasts with an already mentioned referent, *monitor*, and would receive an L + H* accent.

(5a) What did Eric fix?
(5b) Eric fixed the projector.
(6a) Did Eric fix the monitor?
(6b) Eric fixed the PROJECTOR.

Eye-tracking experiments have supported the association of the L + H* accent with contrast: an L + H* accent, but not an H* accent, cued attention to a referent that contrasts with a previously mentioned one. For instance, consider a set of utterances like (7), (8) and (9) below (from Watson et al., 2008):

(7) Click on the camel and the dog.
(8) Move the dog to the right of the square.
(9) Now, put the camel above the square.

After (7), *dog* and *camel* formed a paired set, providing the opportunity for *camel* to be contrasted with *dog* in (9). Using the visual world paradigm, Watson et al. found that when listeners heard an H* accent, there was an increase in fixations to both the contrast item *camel* and a similar sounding unmentioned object like *candle*. But when listeners heard an L + H* accent, there was an increase in fixations only to the contrast item *camel*, supporting the theory that L + H* is related to a contrastive interpretation. Similarly, an L + H* accent on a color adjective directed listeners' attention to an item that contrasts in color with the last-mentioned item, such as a red ball followed by a blue ball (Ito & Speer, 2008; Weber, Braun, & Crocker, 2006; but see Sedivy, Tanenhaus, Chambers, and Carlson (1999), for a

similar task in which no effect of L + H* vs. H* was observed). These effects did not obtain when the target referent did not contrast with the previous one; in fact, an L + H* accent placed on a non-contrastive referring expression misdirected attention to a referent that is contrastive rather than the correct target referent (Ito & Speer, 2008).

It is presently debated whether the difference between H* and L + H* is qualitative (e.g. Chafe, 1974; Cutler & Isard, 1980; Pierrehumbert & Hirschberg, 1990; Selkirk, 2002) or quantitative (e.g. Ladd & Schepman, 2003). In the present study, we investigate the effects of differences in pitch accents without making assumptions about the nature of the differences between those accents.

## Present study

Given that H* and L + H* accents have different effects on online processing, they can be used to test the mechanisms by which pitch accents affect language comprehension. In the present study, we examined the effects of these accents on listeners' memory for short, recorded discourses. Each discourse began with a short context passage that established two contrast sets, each of which contained two items. For example, the context passage (10) establishes two sets: one between *British* and *French*, and the other between *Malaysia* and *Indonesia*. Each context passage was then followed by a continuation passage such as (11) that mentioned one word from each contrast set.

(10) Both the British and the French biologists had been searching Malaysia and Indonesia for the endangered monkeys.
(11) Finally, the (British/French) spotted one of the monkeys in (Malaysia/Indonesia) and planted a radio tag on it.

In our experiments, we tested how the pitch accent type (H* vs. L + H*) on the critical words in the continuation affected subsequent recognition memory for those details (e.g., whether the British or the French spotted the monkey). In Experiment 1, we tested whether pitch accents have consequences for later memory for a discourse. Effects of pitch accent type on long-term memory for a discourse is predicted by both the granularity account, in which a contrastive accent leads to encoding of the *accented* information at a deeper level, and the contrast representation account, in which a contrastive accent leads to additional encoding about the *contrast* item. However, in an identification account, pitch accenting is largely limited to facilitating the initial recognition of discourse status and might not have consequences for long-term memory. In Experiment 2, we then tested whether the mnemonic effects of pitch accent type are due to enhanced memory for portions of the discourse with the contrastive L + H* accent, as predicted by the granularity and contrast representation accounts, or impaired memory for portions without it. Finally, in Experiment 3, we compared the granularity and contrast representation accounts by independently assessing rejections of the contrast item and of an unmentioned but same-category item. The contrast representation account predicts that a contrastive accent should selectively facilitate rejections of the contrast item, whereas the granularity account predicts that any benefit in rejection should apply equally to both types of lures.

## Experiment 1

Experiment 1 used a forced-choice recognition memory to investigated whether differences in pitch accent type affect memory in language comprehension. Both the granularity and contrast representation accounts predict that pitch accents should alter encoding of a discourse and facilitate memory for it. On the other hand, the identification account describes pitch accenting as mainly affecting listeners' initial identification of the discourse status of referents and not long-term memory.

To date, evidence on the effects of pitch accenting on long-term memory has been mixed. Sanford et al. (2006) found that the L + H* accent facilitated performance on a change detection task, relative to the H* accent. However, in the Sanford et al. experiment, the presence of an L + H* accent was confounded with the prior discourse context; it is possible that the differences they observed were due solely to the discourse context. Evidence for this possibility comes from Cutler and Fodor (1979), who manipulated focus using prior discourse context and found that words in focus were better remembered, even while intonational contours were hold constant. Thus, it is possible that the effects observed by Sanford et al. are the result of the manipulation of discourse context and not that of prosody. Effects of pitch accenting on test might also depend on retention interval or test type: Almor and Eimas (2008) found no effect of contrastive pitch accenting on a cued recall test that came after a longer retention interval. Thus, Experiment 1 tested effects of pitch accent types on memory while holding discourse content constant and with a lag between study and test.

### Method

#### *Participants*

In this and all subsequent experiments, the participants were native speakers of English at the University of Illinois at Urbana-Champaign with no reported vision or hearing difficulties, who participated for course credit or cash compensation. Fourteen individuals participated in Experiment 1.

#### *Materials*

Participants listened to 48 prerecorded discourses recorded by a female research assistant with an Inland Northern American English accent (Labov, Ash, & Boberg, 2006), appropriate for the region. Each discourse began with a short context passage such as (10), reproduced below as (12), that established two contrast sets, each containing two items. The context passage was followed by a continuation passage such as (13) that mentioned one member of each contrast set. The items chosen from each contrast

set in the continuation passage were independently randomized across subjects. Contrast sets differed in their grammatical and thematic roles across stories.

(12) Both the British and the French biologists had been searching Malaysia and Indonesia for the endangered monkeys.
(13) Finally, the (British/French) spotted one of the monkeys in (Malaysia/Indonesia) and planted a radio tag on it.

A complete list of stories used is available in Appendix A.

Pitch accenting was manipulated in 24 of the stories, the *critical stories*. Each critical story was presented in one of two conditions: either the first critical word in the continuation received an L + H* (contrastive) accent and the second critical word received an H* accent, or vice versa. Critical stories were randomly assigned to conditions for each participant, with the constraint that an equal number of stories were presented in each condition, and presented in a random order.

Acoustic analyses were conducted to verify that the H* and L + H* conditions in the critical stories differed. Because pitch accents are argued to be realized on the syllable carrying primary stress (e.g. Ladd, 1996), analyses were conducted both on the stressed syllable of the critical word alone and the entire critical word. Table 1 presents means and standard errors for intensity, duration, maximum pitch, difference between maximum and minimum pitch, and mean pitch for both stressed syllables and on entire words. The H* and L + H* conditions differed reliably on all measures both within the stressed syllable and across the entire word. Fig. 1 displays a stylized representation of the pitch contours on the stressed syllable for the H* and L + H* conditions.

To ensure that the stimuli differed only in the pitch accents on the target words, stimuli were created by splicing the critical word into a carrier sentence that was identical across conditions. This procedure raises the concern that the resulting stimuli may have sounded unnatural to the participants. However, in a post-experiment survey, none of the participants in any of the three experiments reported noticing that the audio had been spliced or sounded unusual.

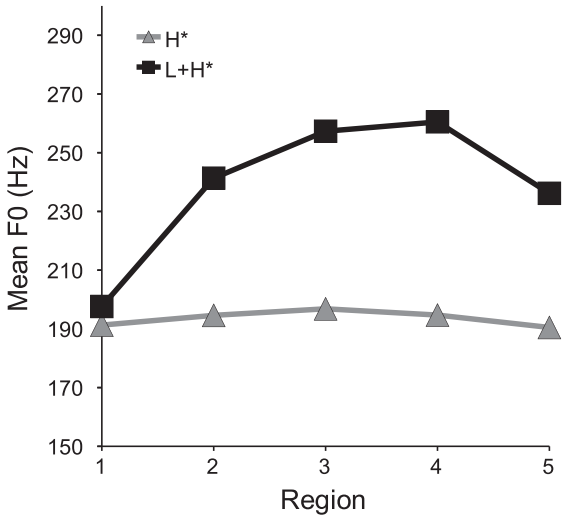


**Fig. 1.** Mean $F_0$ contour of the stressed syllable for words in the H* and in the L + H* conditions in Experiment 1 materials. *Note:* The stressed syllable of each critical word was divided into five equal regions, and the mean $F_0$ was calculated across all items.

The other 24 stories were *filler stories* that followed the same format as the critical stories, but were heard with H* accents on both critical words.

### Procedure

Each participant was informed that they would be listening to stories and that their memory for the stories would later be tested. The specific format of the memory test was not described to participants. Participants performed the task on a computer running MATLAB 7.1 and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997).

Participants were first allowed to adjust the volume of the computer speakers to a comfortable level. Participants then began with a study phase in which they listened to all 48 stories, presented in randomized order. There was a 5 s delay between each story. Additionally, after participants had listened to 24 stories, a message on the computer informed participants that they were halfway through and allowed them to take a break before continuing with the other 24 stories.

**Table 1**
Mean acoustic measures by accent type for Experiment 1 materials.

| Measure | H* | | L + H* | | $F(1, 95)$ | $p$ |
|---|---|---|---|---|---|---|
| | *M* | *SE* | *M* | *SE* | | |
| *Stressed syllable* | | | | | | |
| Mean intensity (dB) | 68.97 | 0.29 | 72.45 | 0.24 | 100.44 | <.0001 |
| Duration (ms) | 231.52 | 7.30 | 388.44 | 15.94 | 191.09 | <.0001 |
| Maximum pitch (Hz) | 221.80 | 3.15 | 305.69 | 3.33 | 462.42 | <.0001 |
| Pitch difference (Hz) | 53.69 | 4.02 | 131.09 | 6.32 | 108.53 | <.0001 |
| Mean pitch (Hz) | 194.93 | 2.75 | 249.45 | 4.37 | 155.83 | <.0001 |
| *Entire word* | | | | | | |
| Mean intensity (dB) | 68.40 | 0.28 | 71.08 | 0.19 | 77.52 | <.0001 |
| Duration (ms) | 399.50 | 11.92 | 658.13 | 15.90 | 633.81 | <.0001 |
| Maximum pitch (Hz) | 230.24 | 3.40 | 301.61 | 3.08 | 399.45 | <.0001 |
| Pitch difference (Hz) | 81.27 | 4.26 | 180.66 | 5.29 | 274.36 | <.0001 |
| Mean pitch (Hz) | 191.78 | 2.71 | 212.58 | 3.26 | 33.36 | <.0001 |

Once participants had listened to all 48 stories, they proceeded to the test phase. In the test phase, memory for each story was tested in the same order in which the stories were presented in the study phase. Each story was displayed with both the context and continuation passages displayed visually on the screen, with the two critical words in the continuation replaced by underscores, as in (14). The stories were not re-presented aurally during the test phase.

(14) Both the British and the French biologists had been searching Malaysia and Indonesia for the endangered monkeys. Finally, the _____ spotted one of the monkeys in _____ and planted a radio tag on it.

Within each story, memory was tested one contrast set at a time. The two words in the contrast set were displayed on the screen and participants indicated the correct word by pressing a key on the keyboard. There was a 500 ms delay between the two tests within a story and a 1000 ms delay between stories.

## Results

Recognition memory accuracy was analyzed as a function of accent type and position within a story (first or second critical word). Mean accuracy in each condition is displayed in Fig. 2.

For categorical measures such as forced choice responses, ANOVA models are inappropriate (Jaeger, 2008). Consequently, data were analyzed using a multi-level logit model to model the log odds of a correct response on each trial. The model included fixed effects for pitch accent, position, and their interaction, and with crossed random intercepts for subjects and items. Fixed effects were coded using mean-centered contrast coding. The model was fit in the R software package (R Development Core Team, 2008) with Laplace estimation using the *lmer()* function of the *lme4* package (Bates, Maechler, & Dai, 2008).

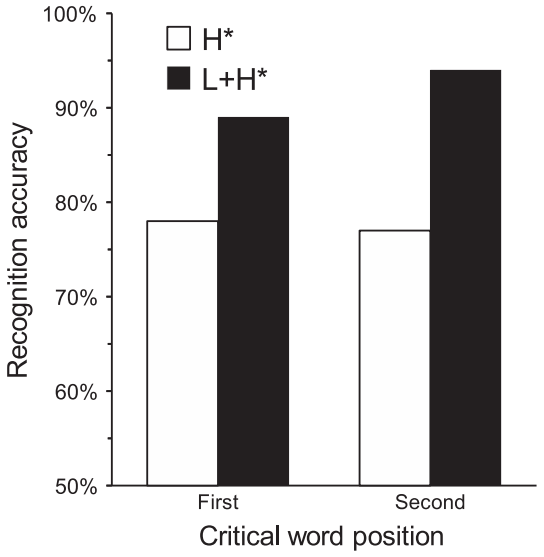


**Fig. 2.** Mean recognition accuracy in Experiment 1 as a function of pitch accent type and critical word position.

Table 2 displays parameter estimates for the model. The odds of correct recognition were approximately 3.38 times (95% CI: [2.09, 5.50]) greater for words heard with the L + H* accent ($M$ = 91% correct) than for words heard with the H* accent ($M$ = 77% correct). The main effect of position was not significant, nor was the interaction of position and accent type.

## Discussion

Experiment 1 found that words receiving an L + H* accent were remembered better than words receiving an H* accent. Because memory was not tested for any story until all the stories had been presented, participants' memory was being tested approximately 30 min after hearing the story. Yet, a memory benefit from the L + H* accent persisted.

This difference is predicted by both the granularity and contrast representation hypotheses, which propose that pitch accents result in improved encoding of semantic or discourse information, but not one in which pitch accents serve only to speed the initial identification of discourse status.

Experiment 1 thus extends prior findings about the effects of pitch accents on immediate memory for a discourse (e.g. Sanford et al., 2006) to later memory for that discourse. Moreover, because pitch accent type was the only factor manipulated, the results cannot be attributed to some of the confounding factors, such as prior discourse context, present in prior work. This benefit of pitch accenting on memory differs from the results of Almor and Eimas (2008), who found no effect of pitch accenting on memory in normal discourse. Although the reason for this discrepancy is not immediately clear, one reason may be that the discourses used by Almor and Eimas were substantially shorter and simpler than the ones in the present experiment.

An alternate interpretation of the mnemonic effects in Experiment 1, however, is not that the L + H* accent facilitated memory relative to its baseline level, but that memory for a word was *impaired* if a *different* word received the L + H*. Recall that when one critical word received an H* accent, the other always received an L + H* accent. Possibly, this L + H* accent elsewhere in the story might have diverted attention away from the H* word and resulted in poorer memory for that contrast set.

**Table 2**
Fixed effect estimates (top) and variance estimates (bottom) for multi-level logit model of correct recognition in Experiment 1 ($N$ = 672, log-likelihood: −273.1).

| Fixed effect | Coefficient | *SE* | Wald *z* | *p* |
|---|---|---|---|---|
| Intercept | 1.92 | 0.19 | 10.14 | <.001 |
| L + H* accent | 1.21 | 0.25 | 4.96 | <.001 |
| Second position | 0.32 | 0.25 | 1.32 | .19 |
| L + H* accent × second position | 0.77 | 0.50 | 1.55 | .12 |
| Random effect | $s^2$ | | | |
| Subject | 0.25 | | | |
| Item | 0.06 | | | |

Some data from Experiment 1 provide a test of this account. Recall that in filler continuations, both test words received an H* accent, whereas in critical continuations one test word received an H* accent and one received an L + H* accent. If the pitch accent effect is due to the L + H* word impairing memory for other information, memory for the H* words should be worse when a different word has received an L + H* (critical trials) compared to when no word received an H* (filler trials). To test this possibility, an additional multi-level logit model compared memory for H* words on filler trials, which contained no L + H* accent at all, with H* words on critical trials, where the other referent had an L + H* accent. This model found a reliable effect of trial type; the odds of correctly remembering a word with an H* were 1.52 times greater (95% CI: [1.09, 2.10]) when no words received L + H* ($M$ = 84% accurate) than when another word received an L + H* accent ($M$ = 77% accurate), Wald $z = 2.50$, $p < .05$.

However, in Experiment 1, the text of the stories was not rotated through the critical and filler conditions. That is, the stories which did not contain an L + H* accent were always different stories than the stories that did, so the presence of an L + H* accent was confounded with the story content. This confound makes it difficult to tease apart enhancement or impairment accounts using the Experiment 1 data.

Experiment 2 more effectively evaluated the effect of an L + H* accent on both contrast sets by independently manipulating the accent on each critical word in a $2 \times 2$ factorial design.

## Experiment 2

In Experiment 2, participants listened to stories of the same format as in Experiment 1. However, the pitch accent type on each critical word in the continuation was independently manipulated, so that some contained two words with L + H* accents, some stories had one, and some stories had none. The four possible versions of the accenting in the continuation are presented in (15a)–(15d).

(15a) Finally, the British spotted one of the monkeys in Malaysia and planted a radio tag on it.
(15b) Finally, the BRITISH spotted one of the monkeys in Malaysia and planted a radio tag on it.
(15c) Finally, the British spotted one of the monkeys in MALAYSIA and planted a radio tag on it.
(15d) Finally, the BRITISH spotted one of the monkeys in MALAYSIA and planted a radio tag on it.

Because pitch accenting is independently manipulated on each critical word, Experiment 2 provides a test of whether the mnemonic effects of the L + H* accent are due to facilitation, impairment, or both. If the L + H* accent works by impairing memory for other parts of the discourse, a particular target word should be less apt to be remembered if the other critical word in the continuation has an L + H* accent rather than an H* accent. For instance, hearing an L + H* accent on *British* might impair memory for *Malaysia* in (15b) relative to (15a). On the other hand, if pitch accents work only by enhancing encoding of the information on which they are placed, as suggested by the granularity and contrast representation accounts, then the pitch accent type on one word may not affect memory for the word in the other contrast set. Experiment 2 tests these predictions.

### Method

#### *Participants*

Fifteen individuals participated in Experiment 2.

#### *Materials*

The same passages were used in Experiment 2 as in Experiment 1. However, the type of pitch accent (H* or L + H*) on each word was now independently varied. Thus, each item could appear in one of four conditions: an L + H* on the first, the second, both, or neither of the critical words. Items were randomly assigned to conditions for each participant, with the constraint that 12 items were presented in each condition. As in Experiment 1, the referent from each contrast set that was specified in the continuation was also randomized for each participant. Stories were presented in a random order.

All 48 stories were used as critical stories in Experiment 2; there were no filler stories. Because the filler stories in Experiment 1 had not been recorded with L + H* accents, all of the stories had to be re-recorded for Experiment 2. The re-recordings were made by a different female research assistant who also had the appropriate Inland Northern American accent.

Again, to verify the manipulation, acoustic analyses were conducted on both the stressed syllable of the critical word alone and the entire critical word. Table 3 presents means and standard errors for intensity, duration, maximum pitch, difference between maximum and minimum pitch, and mean pitch for the Experiment 2 recordings. The H* and L + H* conditions again differed on all measures, both when the analyses were conducted on only the stressed syllable and when the analyses were conducted on the entire word. Fig. 3 displays a stylized representation of the pitch contours on the stressed syllables in Experiment 2.

#### *Procedure*

Aside from the change in stimuli, the procedure in Experiment 2 was the same as in Experiment 1.

### Results

Recognition memory accuracy for each critical word was analyzed as a function of accent type on the critical word itself, the accent type on the *other* critical word in the story, and the critical word's position within a story. Mean percent accuracy in each condition is displayed in Fig. 4. The first panel displays performance for the first critical word and the second panel displays performance for the second critical word.

**Table 3**
Mean acoustic measures by accent type for Experiment 2 materials.

| Measure | H* | | L + H* | | $F(1, 157)$ | $p$ |
|---|---|---|---|---|---|---|
| | *M* | *SE* | *M* | *SE* | | |
| *Stressed syllable* | | | | | | |
| Mean intensity (dB) | 71.64 | 0.24 | 74.72 | 0.17 | 242.84 | <.0001 |
| Duration (ms) | 266.55 | 8.42 | 392.32 | 15.06 | 164.38 | <.0001 |
| Maximum pitch (Hz) | 243.26 | 1.63 | 312.22 | 2.32 | 1007.10 | <.0001 |
| Pitch difference (Hz) | 39.76 | 1.92 | 104.15 | 3.41 | 325.86 | <.0001 |
| Mean pitch (Hz) | 221.48 | 1.29 | 268.38 | 2.23 | 510.02 | <.0001 |
| *Entire word* | | | | | | |
| Mean intensity (dB) | 71.02 | 0.23 | 73.33 | 0.13 | 152.45 | <.0001 |
| Duration (ms) | 503.15 | 13.05 | 746.37 | 15.31 | 428.65 | <.0001 |
| Maximum pitch (Hz) | 251.50 | 1.71 | 305.14 | 1.99 | 696.69 | <.0001 |
| Pitch difference (Hz) | 54.41 | 1.98 | 123.22 | 3.13 | 500.90 | <.0001 |
| Mean pitch (Hz) | 220.24 | 1.34 | 242.53 | 1.77 | 121.17 | <.0001 |

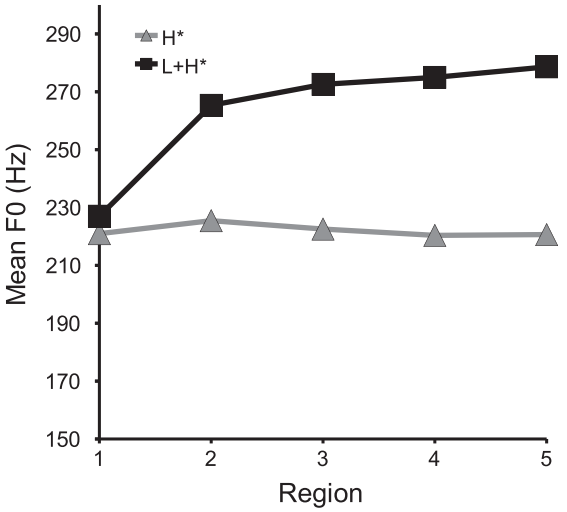


**Fig. 3.** Mean $F_0$ contour of the stressed syllable for words in the H* and in the L + H* conditions in Experiment 2 materials.

A multi-level logit model was fit to the data. Log odds of a correct response on each trial were modeled as a function of the fixed effects for the accent on the target word, accent on the other critical word, word position, and all interactions of these factors, and of crossed random intercepts for subjects and items. Fixed effects were again coded using mean-centered contrast coding. The model was fit using the same software as the Experiment 1 model. Table 4 displays parameter estimates for this model.

The effect of the accent type on the critical word being tested was reliable. The odds of a correct recognition memory response were 1.82 times greater (95% CI: [1.34, 2.47]) for critical words heard with the L + H* accent ($M$ = 90% correct) than words heard with the H* accent ($M$ = 83% correct). No other effects or interactions were reliable. The effect of accent on the other critical word did not approach significance. Moreover, the difference between the means was in the opposite direction predicted by an impairment effect; performance was numerically greater when the other word had an L + H* accent rather than an H* accent.

## Discussion

Experiment 2 replicated the benefit of the L + H* accent on memory that was observed in Experiment 1, providing additional evidence that differences in pitch accenting have consequences for long-term memory for a discourse.

Experiment 2 also directly tested whether changes in pitch accent type facilitate memory for the accented information, impair memory for other parts of the discourse, or both. The L + H* accent facilitated memory for a contrast set on which it was placed, but there was no evidence that an L + H* accent on one contrast set impaired memory for the other contrast set. This finding suggests that the effects of pitch accents are due to facilitation, not impairment.

Recall that in Experiment 1, memory for H*-accented words was worse if another word received an L + H* accent. In Experiment 1, however, the presence of a L + H* accent was confounded with the discourse content. Because Experiment 2 did not replicate this effect when this confound was absent, we attribute the difference between filler and target items observed in Experiment 1 to differences in discourse content.

Both Experiments 1 and 2 found that differences in pitch accenting enhance long-term memory for a discourse. This effect is not predicted by a theory in which pitch accents only speed the initial identification of discourse status. Instead, it seems likely that the L + H* accent actually facilitates the encoding of the discourse, as proposed by both the granularity and contrast representation theories. However, these theories differ in their explanations of *how* accenting facilitates memory for a discourse.

The granularity account proposes that the L + H* accent results in semantic encoding as a subordinate rather than a superordinate category—for instance, encoding that the scientists who found the monkey were British, specifically, rather than encoding that they were European. Crucially, this account predicts that the L + H* should facilitate rejection of *any* other in-category items, such as *Portuguese*, whether they were part of the discourse or not. The contrast representation theory, alternately, proposes that the L + H* accent leads to enhanced encoding of what did *not* happen; i.e., memory that it was not the French who found the monkey. Consequently, under the contrast

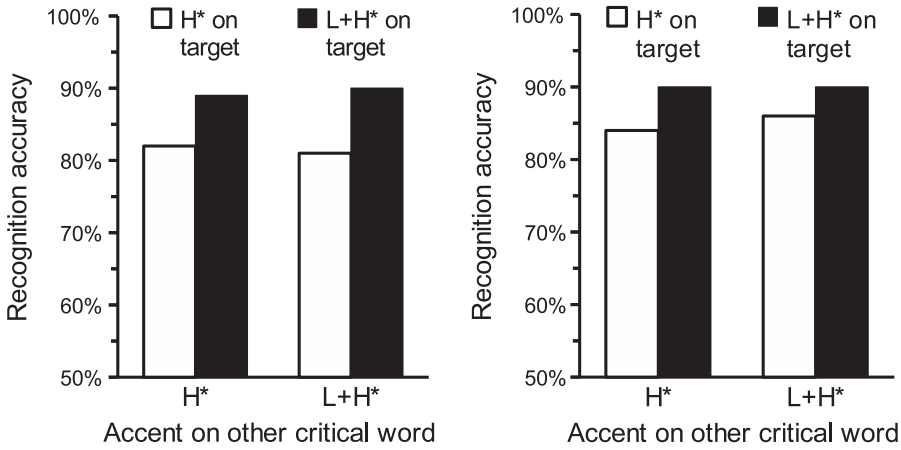


**Fig. 4.** Mean recognition accuracy in Experiment 2 as a function of pitch accent type on tested critical word and pitch accent type on other critical word. The left panel displays critical words in the first position and the right panel displays critical words in the second position.

**Table 4**
Fixed effect estimates (top) and variance estimates (bottom) for multi-level logit model of correct recognition in Experiment 2 ($N = 1600$, log-likelihood: −607.8).

| Fixed effect | Coefficient | *SE* | Wald *z* | *p* |
|---|---|---|---|---|
| Intercept | 2.08 | 0.18 | 11.56 | .001 |
| L + H* accent on this word | 0.60 | 0.16 | 3.89 | .001 |
| L + H* accent on other word | 0.03 | 0.16 | 0.19 | .85 |
| Second position | 0.15 | 0.15 | 0.94 | .34 |
| L + H* on this word × L + H* on other word | −0.04 | 0.31 | −0.12 | .91 |
| L + H* on this word × second position | −0.13 | 0.31 | −0.41 | .68 |
| L + H* on other word × second position | 0.01 | 0.31 | 0.04 | .97 |
| L + H* on this word × L + H* on other × second position | −0.35 | 0.61 | −0.57 | .57 |
| Random effect | $s^2$ | | | |
| Subject | 0.41 | | | |
| Item | 0.21 | | | |

representation account, the benefit from the L + H* should only apply to rejections of particular contrast items from the discourse, not to any potential alternative.

The forced choice task used in Experiments 1 and 2 does not provide information that would distinguish between these two possibilities. In a forced choice task, identification of the correct item could reflect memory for the correct item or rejection of any number of lures.

To tease apart the granularity and contrast representation accounts, we introduced a new test format in Experiment 3.

## Experiment 3

Experiment 3 tested participants' memory for discourse using a true–false verification task. Stories were presented in the same format as Experiments 1 and 2, consisting of a context passage (such as 10, reproduced below as 16) followed by a continuation (11, reproduced below as 17).

(16) Both the British and the French biologists had been searching Malaysia and Indonesia for the endangered monkeys.
(17) Finally, the British spotted one of the monkeys in Malaysia and planted a radio tag on it.

During the test phase, participants saw statements about the original stories and were told to indicate whether they were true or false. Each statement was tested in one of the three probe conditions: a correct statement (such as 18), a statement about the contrast item (such as 19), or a statement about a previously unmentioned, but within-category, lure (such as 20).

(18) The British scientists spotted the endangered monkey and tagged it.
(19) The French scientists spotted the endangered monkey and tagged it.
(20) The Portuguese scientists spotted the endangered monkey and tagged it.

Unlike the two alternative forced choice task used in Experiments 1 and 2, this task allows for a separate assessment for rejections of the contrast item, and rejections of an unmentioned but within-category item. If the effect of contrastive accenting is to create or enhance a memory representation of what did *not* happen (e.g. that the French scientists did not find the monkey), as predicted by the contrast representation theory, then the L + H* accent should facilitate rejections only of the contrast item (19). The benefit of remembering *the French scientists did not find the monkey* should not extend to rejections of an unmentioned lure like (20), which was not part of the original contrast set.

In comparison, the granularity theory proposes that L + H* accents primarily enhance representations of the *correct* item. Enhanced knowledge of the correct answer can facilitate rejections of lures in general (the phenomenon of recollection rejection; e.g. Brainerd, Reyna, and Estrada (2006) in the domain of discourse comprehension),

but if this rejection is based on increased strength of the correct answer, it should apply to rejections of both lure types. That is, increasing the strength of the correct *British* statement in (18–20) may facilitate rejection of the lures, but it should not selectively facilitate rejections of the *French* lure over rejections of the *Portuguese* lure. Thus, the contrast representation account predicts that the L + H* accent should facilitate rejections of contrast lures but not unmentioned lures, whereas the granularity account predicts that any facilitation in rejection should apply to both lure types.

Experiment 3 also increased the lag between study and test to 24 h, providing a further test of the degree to which effects on pitch accenting on discourse processing persist.

## Method

### *Participants*

Thirty individuals participated in Experiment 3.

### *Materials*

The same recordings as in Experiment 2 were used; however, only 36 of the 50 stories were included in Experiment 3. Stories that were eliminated from Experiment 2 were eliminated because there was not a third semantically plausible lure that could be used in the unmentioned lure condition (e.g. in one story, the original contrast was between *boys* and *girls*).

In Experiment 3, subjects were randomly assigned to view one of six presentation lists. There were six critical conditions, resulting from a $2 \times 3$ factorial design: pitch accent (H* or L + H*) $\times$ probe type (correct item, contrast item, or unmentioned item). The assignment of each fact to conditions was counterbalanced across lists using a Latin Square design. Every list contained all 36 stories. The order of the stories was randomized in creating each list but was fixed for every use of that list.

Because only a subset of items was used in Experiment 3, the acoustic measures were re-conducted to verify that the H* and L + H* conditions differed in this subset. Table 5 presents means and standard errors of these measures. Within the Experiment 3 items, the H* and L + H* conditions differed reliably on all measures both on the stressed syllables and across the entire word.

For each story, two sets of test items were then constructed. Each set of items tested one of the two critical facts from the story. Within each set were three statements about one of the critical facts about the story. The statements varied only in whether they named the correct referent, the contrast item, or an unmentioned lure. For example, items (18–20), reproduced below as (21–23) below, test one of the facts from the monkey story, while items (24–26) test the other fact. Crucially, the items in each set did not reveal the answer to items in the other set. For example, viewing (21) would not reveal whether (24) was a true or false statement. This allowed both facts from each story to be tested separately.

(21) The British scientists spotted the endangered monkey and tagged it.
(22) The French scientists spotted the endangered monkey and tagged it.
(23) The Portuguese scientists spotted the endangered monkey and tagged it.
(24) The endangered monkey was finally spotted in Malaysia.
(25) The endangered monkey was finally spotted in Indonesia.
(26) The endangered monkey was finally spotted in the Philippines.

The complete list of test items is available in Appendix B.

As in prior experiments, the words chosen in the continuation from each contrast set were counterbalanced between participants; thus either (21) or (22) could be the correct statement depending on the counterbalancing. However, to control the contents of the contrast set between conditions, the lexical item used as the unmentioned lure was never used as a member of the contrast set; that is, (23) was always a false statement.

### *Procedure*

Participants visited the lab for two sessions. The first session consisted of the study phase, in which the stories were presented aurally as in the first two experiments. Par-

**Table 5**
Mean acoustic measures by accent type for Experiment 3 materials.

| Measure | H* | | L+H* | | $F(1, 157)$ | $p$ |
|---|---|---|---|---|---|---|
| | *M* | *SE* | *M* | *SE* | | |
| *Stressed syllable* | | | | | | |
| Mean intensity (dB) | 71.76 | 0.25 | 74.81 | 0.17 | 213.34 | <.0001 |
| Duration (ms) | 269.48 | 8.74 | 396.89 | 16.25 | 140.35 | <.0001 |
| Maximum pitch (Hz) | 243.40 | 1.78 | 311.28 | 2.43 | 931.90 | <.0001 |
| Pitch difference (Hz) | 39.95 | 2.04 | 103.35 | 3.61 | 281.09 | <.0001 |
| Mean pitch (Hz) | 221.59 | 1.39 | 267.34 | 2.34 | 454.62 | <.0001 |
| *Entire word* | | | | | | |
| Mean intensity (dB) | 71.13 | 0.25 | 73.41 | 0.14 | 132.91 | <.0001 |
| Duration (ms) | 497.16 | 12.93 | 744.34 | 16.23 | 397.68 | <.0001 |
| Maximum pitch (Hz) | 252.00 | 1.88 | 305.62 | 2.07 | 636.25 | <.0001 |
| Pitch difference (Hz) | 54.05 | 2.11 | 121.54 | 3.31 | 433.72 | <.0001 |
| Mean pitch (Hz) | 220.96 | 1.47 | 243.90 | 1.81 | 112.41 | <.0001 |

ticipants returned to the lab twenty-four hours after the first session for the test phase. This delay was introduced to test the effects of pitch accenting over a longer retention interval and to ensure that correct rejection of the unmentioned lure was not at ceiling.

At the start of the second session, participants were instructed that they would see a series of statements about the stories from the previous session, and that some of the statements would be true and others false. Participants were instructed that they should respond *true* only if the statement was exactly true, or *false* if any part of the statement was false.

Test items were presented one at a time and participants indicated whether each was *true* or *false* by pressing a key on the keyboard. The two facts about each story were tested separately. Each fact was tested in only one probe condition; that is, participants would see only one of (21), (22) and (23) and one of (24), (25), and (26). Facts were presented in random order. The two facts about each story were not necessarily tested one after another.

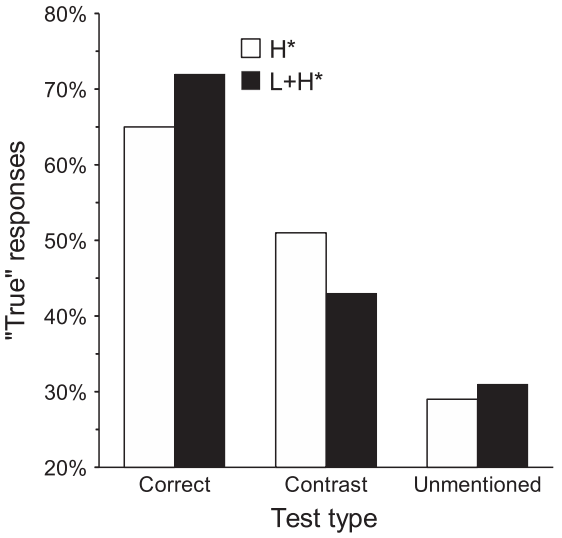


**Fig. 5.** Mean rate *of true* responses in Experiment 3 as a function of probe type and pitch accent type. Responding *true* is a hit to a correct probe and a false alarm to a contrast or unmentioned probe.

## Results

In Experiments 1 and 2, participants made a forced choice between the correct item and contrast item; performance in those experiments could be analyzed in terms of accuracy of these choices. However, Experiment 3 used a task in which participants responded *true* or *false* independently to correct probes, contrast probes, and unmentioned probes. Accuracy is an inappropriate dependent measure in this case because differences in mean accuracy across probe conditions might reflect general preferences to respond *true* or *false* rather than differences in memory fidelity. This confound can be addressed by detection-theoretic analyses, which permit decomposition of behavior into response bias and sensitivity (Green & Swets, 1966; Macmillan & Creelman, 2005; see also Wright, Horry, and Skagerberg (2008), for an application of detection theory to multi-level models using the log odds ratio).

To conduct this analysis, data from Experiment 3 were parameterized to reflect whether the participant responded *true* or *false* to each item. This parameterization permits an analysis of subjects' *true* or *false* classifications in which sensitivity to an item's truthfulness can be statistically and logically dissociated from baseline biases to respond *true*. Mean percentage of *true* responses in each condition is displayed in Fig. 5; these *true* responses are hits when responding to correct probes but are false alarms when responding to contrast and unmentioned probes.

Performance was analyzed using a multi-level logit model with *true* responses as the dependent variable, probe type and accent type and their interaction as fixed effects, and subjects and items as crossed random effects. The probe type variable was coded using effects coding to separately compare responses to the correct probes (to which *true* is a hit) with each of the two types of lures (to which *true* is a false alarm). Accent type was coded using mean-centered contrast coding, as before. Table 6 presents parameter estimates for this model.

The first parameter in this model represents participants' overall bias to respond *true* or *false*. This parameter was not reliable, indicating that participants did not have an overall preference to respond *true* or *false*.

Subsequent parameters tested what factors affected participants' likelihood of responding *true*. The contrast between correct and unmentioned probes was reliable; the odds ratio between *true* and *false* response was 5.08 times (95% CI: [3.92, 6.60]) greater for correct probes compared to unmentioned probes, indicating that participants correctly said *true* more to correct probes than to unmentioned probes. The contrast between correct probes and contrast probes in *true* responses was not reliable; the odds ratio between *true* and *false* responses was only 1.15 times (95% CI: [0.90, 1.48]) greater for correct probe than contrast probes.

**Table 6**
Fixed effect estimates (top) and variance estimates (bottom) for multi-level logit model of *true* responses in Experiment 3 ($N = 2160$, log-likelihood: −1366).

| Fixed effect | Coefficient | *SE* | Wald *z* | *p* |
|---|---|---|---|---|
| Intercept (response bias) | −0.06 | 0.09 | −0.64 | .52 |
| Correct probe vs. contrast probe (sensitivity) | 0.14 | 0.13 | 1.12 | .26 |
| Correct probe vs. unmentioned (sensitivity) | 1.63 | 0.13 | 12.24 | <.001 |
| L + H* accent (effect on response bias) | 0.02 | 0.09 | 0.17 | .87 |
| L+H* accent × correct vs. contrast probe (effect on sensitivity) | 0.78 | 0.26 | 3.07 | <.01 |
| L + H* accent × correct vs. unmentioned probe (effect on sensitivity) | −0.13 | 0.27 | −0.49 | .62 |
| Random effect | $s^2$ | | | |
| Subject | 0.07 | | | |
| Item | 0.12 | | | |

Again, the important tests were the effects of pitch accent on participants' responses within these probe types. If the effect of a contrastive accent was to induce an overall bias to respond *true* or *false*, regardless of the actual truthfulness of the probe, there would be a main effect of pitch accent independent of probe type. Importantly, however, the main effect of pitch accent on *true* or *false* responses was *not* reliable, indicating that the contrastive accent did not induce an overall bias to respond *true* or *false*.

Effects of pitch accent on the ability to actually discriminate *true* from *false* probes are re-presented in the model as interactions between pitch accent type and probe type on the rate of *true* responses. Pitch accent type reliably interacted with the contrast comparing correct and contrast items, reflecting the fact that the L + H* accent enhanced discrimination between the correct items and contrast item lures. The L + H* accent increased the odds ratio between correct and contrast probes by 2.18 times (95% CI: [1.33, 3.61]).

Crucially, however, the L + H* pitch accent did *not* reliably improve discrimination between correct and unmentioned items. In fact, the effect was numerically in the opposite direction; the L + H* accent *decreased* the odds ratio between correct and unmentioned probes by 0.88 times (95% CI: [0.52, 1.48]) relative to the H* accent, but this difference was not reliable. This absence of a benefit is predicted by the contrast-encoding theory but not by the granularity theory. The failure of the L + H* accent to benefit rejections of the unmentioned probe cannot be attributed to a ceiling effect in the unmentioned probe condition; participants still made false alarms 30% of the time in the unmentioned condition.

## Discussion

Experiment 3 compared the contrast-encoding and granularity accounts of pitch accenting by comparing the effects of pitch accenting on three types of probes: correct items, contrast items, and unmentioned items.

Overall, participants were more accurate in rejecting contrast probes than unmentioned probes. This baseline difference between lure types might reflect the effects of familiarity or prior discourse status; the contrast items had been seen during the study phase whereas the unmentioned items were novel in the context of the experiment and new to the discourse.

The key test of theories of pitch accents, however, is the differing effect of the L + H* accents within each of the two lure types. The L + H* accent facilitated discrimination between correct items and contrast item lures. However, the L + H* accent did *not* facilitate discrimination between correct items and unmentioned lures that belonged to the same semantic category but that were not part of the contrast set in the discourse. The failure of the L + H* accent to affect discrimination between correct and unmentioned lures cannot be attributed to a ceiling effect; participants failed to reject the unmentioned lures nearly a third of the time. The effects of L + H* are also not due to response bias; the change in accent type did not alter the overall tendency to respond *true* or *false*.

Experiment 3 also added an additional challenge for participants in that it introduced a new test format that required participants to affirm or reject statements. In general, participants often perform poorly on such tasks; they frequently affirm erroneous statements such as (27), even when debriefing questionnaires confirm they indeed knew that it was Noah, not Moses, who took animals on the ark. (This phenomenon is often termed the *Moses illusion*; for review, see Park and Reder (2004) and Sanford and Sturt (2002).)

(27) Moses took two animals of each kind of the ark.

We attempted to reduce participants' tendency to endorse such statements by emphasizing in the instructions that statements should be rejected if any part of the statement seemed false. Nevertheless, the general difficulty of rejecting such erroneous statements might explain why participants sometimes incorrectly affirmed the false statements—especially those lures that involved the contrast items, which had been previously mentioned in the discourse. Importantly, however, task difficulty *cannot* explain the selective benefit of the L + H* accent in facilitating rejections of the contrast statements. If the contrast statements were simply too confusing to reject, no manipulation should have facilitated performance in that condition, yet this was precisely the condition where the L + H* accent was beneficial. Moreover, the detection-theoretic analysis found that the benefit from the L + H* accent in rejecting the contrast statements was independent of any overall bias to respond *true* to these statements.

In summary, Experiment 3 found that contrastive accents facilitated rejections of contrast lures but not of unmentioned lures. These results are most consistent with the contrast encoding hypothesis: changes in pitch accenting may lead to encoding of a contrast, which facilitates rejection of the contrast set, but does not facilitate rejections of unmentioned items that were never part of the contrast set. The present results are inconsistent with a granularity account in which the accented information itself is encoded more precisely. To the extent the L + H* accent improves rejection of the lures by facilitating recollection of the correct target, rejections of *any* within-category lure should have been improved by the contrastive accent, contrary to the results of Experiment 3.

## General discussion

In three experiments, we assessed participants' memory for short, recorded discourses as a function of the type of pitch accent applied to critical words within the discourses. In all three experiments, memory for words spoken with an L + H* accent was more accurate than for words spoken with an H* accent; in Experiment 3, these effects were found to persist at least as long as one day. These downstream effects of pitch accents are not predicted by a theory of pitch accenting in which pitch accents only speed the initial identification of the discourse status of referents.

Several accounts of this memory effect were tested. Experiment 2 tested whether the advantage of the L + H* was due to enhanced memory for information receiving the L + H*, as predicted by the granularity and contrast representation accounts, or due to impaired memory for information that did not receive the L + H*. In Experiment 2, pitch accents were independently manipulated on two different details in the story. The presence of an L + H* accent on one word facilitated memory for that detail, but did not impair memory for the other detail in the story. This suggests that the advantage for contrastive accenting lies in improved memory for contrastively accented information and not in impaired memory for other parts of the discourse.

Experiment 3 pitted the granularity account, in which manipulations of focus enhance the encoding of the *target* information, against a contrast representation account, in which contrastive accents lead to greater encoding of the contrast set. Because the stimuli explicitly specified the contrast set in the discourse, rejections of lures regarding the contrast set could be compared to rejections of lures regarding unmentioned items. If the L + H* accent benefits memory because superior recollection of the target information enhances rejection of lures, as predicted by the granularity account, the effect should apply equally to rejections of both types of lures. However, Experiment 3 found that the memory benefit extended only to rejections of items in a previously mentioned contrast set and not to other within-category items.

These results are most consistent with the contrast representation theory. An L + H* accent may lead listeners to encode additional information about what did *not* happen, improving later rejections of contrast set items but not rejection of information outside the contrast set. This theory is consistent with linguistic theories (e.g. Pierrehumbert & Hirschberg, 1990) that have described the L + H* accent as contrastive. It is also consistent with the findings of Braun and Tagliapietra (in press), who found that contrastive accents primed features of contrast items, but not features of related but non-contrastive items. The present experiments extend this work to show that differences in pitch accenting have long-term consequences for encoding and representation of a discourse. Pitch accents are not merely a tool for setting up an initial discourse representation but have consequences for the representation of a discourse at least a day later.

## Pitch accents and perception

An alternate account of the effects of pitch accenting on memory is that they result only from the acoustic or perceptual characteristics of accented words, rather than discursive or pragmatic information associated with the accents. In the present study, the words produced with the L + H* accent had greater duration, intensity, and pitch excursion than the H* words, so the words themselves were more salient.

However, several pieces of evidence suggest that the effects of pitch accents are not limited to their perceptual characteristics. If the L + H* accent served only to make a word more perceptually salient, it should facilitate memory performance for that word regardless of the probe type. However, Experiment 3 found that the benefit in memory from the L + H* accent applied more selectively: the L + H* accent facilitated rejections of contrast items but not of unmentioned words. This result is predicted by the contrast representation account but not by a perceptual account.

Second, facilitation of memory from pitch accenting is consistent with findings that other manipulations of linguistic focus improve memory even if they do not necessarily increase perceptual availability. Manipulating focus through pseudocleft constructions or prior discourse context improves change detection in printed sentences (Sturt et al., 2004), without necessarily changing the acoustic or perceptual characteristics of the words themselves.

Finally, when perceptual salience is manipulated, it does not always enhance semantic processing. Kamas, Reder, and Ayers (1996) tested participants' ability to detect semantic anomalies such as the aforementioned Moses illusion: Noah, not Moses, took animals on the ark, making (28) a false statement. Kamas et al. found that increasing the salience of the anomalous word by presenting it in all caps, as in (28), did not increase participants' sensitivity to the presence of an anomaly. Rather, capitalization simply increased participants' bias to respond that the sentence contained an anomaly, regardless of whether or not it actually did.

(28) MOSES took two animals of each kind of the ark.

Nevertheless, pitch accenting clearly does affect the perceptual characteristics of words. Accented words tend to be longer and more intense than de-accented words (Ladd, 1996), and words with contrastive accents may have further differences in pitch, boundary tones, and post-word pausing (e.g. Selkirk, 2002). These differences are probably not coincidental. If pitch accents signal new or contrastive information, making those words particularly perceptually salient may facilitate communication.

## Identification and beyond

In all three experiments, pitch accents were observed to influence representations of a discourse well after its initial presentation. In Experiment 3, pitch accents affected memory for the discourse even a day after the study phase. These results are evidence against a strict identification account of pitch accenting in which pitch accents affect only the initial identification of the discourse status of referents and have no effect on the final representation of the discourse.

Effects of pitch accenting on memory have been mixed in the literature to date. All three experiments reported here found benefits on contrastive pitch accenting on long-term memory, and Sanford et al. (2006) found similar benefits across a shorter retention interval. However, Almor and Eimas (2008), using similar short, recorded discourses, found no effect of contrastive pitch accenting on long-term memory. The reason for the differences in results across experiments is not immediately clear. One

reason may be that the discourses used by Almor and Eimas were simpler and shorter than those in other experiments, which may have permitted less opportunity for contrastive accents to benefit memory.

It is unlikely that all of the effects of pitch accents on later memory can be attributed to initial identification processes. To some extent, quickly identifying the discourse status of referents might facilitate memory simply because it permits more time to encode details about those referents. However, such a mechanism cannot explain the qualitative differences observed in Experiment 3, in which changes in pitch accenting were observed to influence rejection of contrast lures but not unmentioned lures. If the contrastive accents simply resulted in a quantitative shift in the amount of time for encoding, they should improve performance across all conditions. Instead, these results suggest that contrastive accents modulate discourse representation even beyond the initial interpretation of discourse status; e.g., by leading to the encoding of information about a contrast set.

Of course, it is likely that pitch accents also affect the initial interpretation of discourse status. Using eye-tracking, Dahan et al. (2002) find that participants rapidly interpret accented noun phrases as referring to discourse-new referents and de-accented noun phrases as referring to given referents. The present study does not imply that pitch accents have *no* effect on initial interpretation of a discourse. Rather, it suggests that pitch accents can produce qualitative changes in the representation of a discourse in addition to their role in speeding its initial interpretation.

## Facilitation and impairment

Across three experiments, target words receiving an L + H* pitch accent were remembered better than words receiving an H* pitch accent. Experiment 2 independently manipulated the accent on each referent to test whether these effects were due to the L + H* accent facilitating memory for information on which it was placed or impairing memory for other parts of the discourse. The results of this experiment found relatively local effects of the L + H* accent: the L + H* accent facilitated memory for the contrast set on which it was placed and did not impair memory for the other contrast set.

Other experiments on memory for a discourse have used manipulations that did impair memory relative to a baseline level. For example, when one referent was placed in a cleft construction, referents outside the cleft were remembered worse than when no cleft was present (Sanford, Price, & Sanford, 2009). Differences in the manipulations across experiments suggest possible boundary conditions for this memory impairment effect. In the experiments on clefting, the non-clefted referent was outside the scope of linguistic focus, and memory for it decreased. In the present experiments, the targets were always accented and in focus; the manipulation was whether the accent *type* was a presentational (H*) or contrastive (L + H*) accent. In this case, placing a contrastive accent on one piece of information did not impair memory for other information that received a presentational accent.

One possible interpretation of these differing results is that focus on just one referent may impair memory for other referents. However, when focus is distributed over multiple pieces of information, manipulations of the *type* of pitch accenting on each item—such as a presentational vs. contrastive accent—may not affect memory for the other information.

## Focus and contrast

Pitch accents have been argued to reflect linguistic focus (e.g. Schwarzschild, 1999; Selkirk, 1986). However, linguistic theories differ as to the precise role of focus. For example, Rooth (1992) proposed an alternative semantics account of focus in English, in which focus evokes a set of relevant alternatives for the focused referent. In other accounts, focus may be interpreted as existential presupposition (Rooth, 1999). Or, focus may reflect any of several constraints and not have any specific interpretation; rather, it is only lack of focus that is specifically interpreted to signal givenness (Schwarzschild, 1999). Findings from Experiment 3 are consistent with an alternative semantics account in which focus is associated with relevant alternatives: the L + H* pitch accent facilitated correct rejections of the alternative item in the contrast set, but not rejections of lures outside of the contrast set.

The alternative semantics theory of focus also integrates the current findings with prior findings on the effects of pitch accenting and other manipulations of focus. Recall that manipulations of pitch accenting have sometimes been observed to facilitate semantic processing even when the prior discourse did not provide an explicit contrast set; for example, the L + H* accent increased the odds of detecting the change of *wallet* to *purse* in the experiment by Sanford et al. (2006). However, the alternative semantics theory of focus proposes that focused constituents evoke alternatives. That is, hearing *the money from the WALLET* may have led participants to consider *the money from the purse* as an alternative, and thus increased the likelihood that participants noticed the change. This explanation is supported by the finding that a contrastive accent on a word like *antenna* can prime possible contrast items like *dish*, even when those items have not been specifically mentioned in the discourse (Braun & Tagliapietra, in press).

The present experiments differed from past work in that the discourse established an explicit contrast set (e.g. between *British* and *French* scientists only). In this case, an L + H* accent on one of the items facilitated correct rejection only of the item from the contrast set and not other unmentioned lures. This suggests that the discourse context may modulate the size of the set of alternatives evoked by focus. When an item has been mentioned as part of an explicit contrast set, the alternatives may be limited to those items in the contrast set; when the discourse does not provide a contrast set, a wider set of alternatives may be considered.

One question that remains unresolved is the specific nature of the contrast evoked by the L + H* accent. For example, does considering a contrast item like *French* lead listeners to consider only general properties of *French*, or to elaborate on the role of the contrast item in the particular

discourse? To the extent that contrastive accents did not shift encoding so greatly as to completely avert errors in memory, any shifts in encoding may be relatively subtle, but the specific nature of these changes remains to be determined.

## Conclusion

In three experiments, the L + H* pitch accent was found to facilitate memory for a discourse compared to the H* accent. This difference extended to memory for the discourse up to a day later, extending prior findings about the effects of pitch accenting on online language comprehension. Moreover, Experiment 3 provides evidence that this difference is driven by facilitated rejection only of contrast items, and not of other false items, consistent with theories that relate some pitch accents to contrast.

## Acknowledgments

We thank Kathryn Bock and members of the Communication and Language Lab and Human Memory and Cognition Lab for their comments and suggestions, Jessica George and Shefali Khanna for recording stimulus materials, Euil Lim, Dipika Mallya, Kevin Park, Kirstin Shafer, and Jonathan Tullis for running experimental sessions, and Alexandria Packard for assistance with the acoustic measurements.

This work was supported by National Science Foundation Graduate Research Fellowship 2007053221 to Scott H. Fraundorf and National Institutes of Health Grants R01DC008774 to Duane G. Watson and R01AG026263-03 to Aaron S. Benjamin.

Portions of this work were submitted as part of the first author's Master's thesis.

## Appendix A

*Stimuli for all experiments*

1. Experiment 1: critical

*Context*: After the old mansion was finally sold to a new owner, a plumber was sent in to inspect the bathroom and kitchen to make sure that they were up to standards. He was horrified to discover that there were pests as well as leaks.

*Continuation*: He was able to get rid of the (pests/leaks) in the (bathroom/kitchen), but he had to send for another team to fix the rest. It was one of the worst messes he'd ever seen.

2. Experiment 1: critical, Experiment 2, Experiment 3

*Context*: The local parks commission had a busy meeting on Wednesday to decide how to spend its money for the year. People disagreed on whether the commission should focus its resources on expanding the parks or tidying the existing parkland. The parents wanted a playground and the teenagers wanted a skateboard park.

*Continuation*: After a long debate, a compromise was worked out to (tidy/expand) the parks and build a (playground/skateboard park), but it didn't seem like anyone was very happy with the decision.

3. Experiment 1: critical, Experiment 2, Experiment 3

*Context*: When the strange respiratory disease began spreading, the scientists knew they had to determine whether it was caused by a bacteria or by a virus, and if it infected the lungs or the throat. The research labs worked 'round the clock.

*Continuation*: After discovering the cause was a (bacteria/virus) found in the (lungs/throat), the scientists passed the information on to medical specialists who began to work on a cure.

4. Experiment 1: critical

*Context*: As a rising star in the world of chess, the young prodigy had received challenges from both a well-known grandmaster and from the creators of a chess-playing computer. The prodigy wanted to accept the challenges, but knew he would have to play harder than he ever had before. He considered both a more aggressive style and a more defensive one.

*Continuation*: After losing the first game against the (computer/grandmaster) using the (aggressive/defensive) style, the prodigy changed tactics and rallied back to win the second and third games.

5. Experiment 1: critical, Experiment 2, Experiment 3

*Context*: The director and producer of a forthcoming film were having a big dispute over the budget and screenplay.

*Continuation*: When an agreement couldn't be reached on the (budget/screenplay), the (director/producer) quit.

6. Experiment 1: critical, Experiment 2, Experiment 3

*Context*: Both the doctors and the nurses were pleased with the renovation of the hospital, which brightened the rooms and improved the air conditioning during the hot summer months.

*Continuation*: The (doctors/nurses) were particularly pleased to have the (brighter/cooler) rooms.

7. Experiment 1: critical, Experiment 3

*Context*: Thanks to an anonymous donor, the playground at Jefferson Elementary recently acquired a new jungle gym and a slide for the playground.

*Continuation*: The (jungle gym/slide) seemed especially popular among the kids, as all the (boys/girls) would gather there every recess.

8. Experiment 1: critical, Experiment 2, Experiment 3

*Context*: Both the British and the French biologists had been searching Malaysia and Indonesia for the endangered monkeys.

*Continuation*: Finally, the (British/French) spotted one of the monkeys in (Malaysia/Indonesia) and planted a radio tag on it.

9. Experiment 1: critical

*Context*: The new product line had tested poorly, and the CEO of the company was uncertain whether to revise it or abandon it. She met with the research team and the marketing team to get their input.

*Continuation*: Eventually, she decided to heed the advice of the (research/marketing) team and (revise/abandon) the product line.

10. Experiment 1: critical

*Context*: The newspaper didn't have the resources to cover both the fire and the robbery, so the editor assigned the paper's best reporter and photographer to cover one of the two stories.

*Continuation*: This turned out to be a good decision, because the (reporter's/photographer's) work on the (fire/robbery) story was later nominated for an award.

11. Experiment 1: critical, Experiment 2, Experiment 3

*Context*: Strangely enough, on the same day, both the Boston-based and the Houston-based publisher released new books about the Civil War and the Great Depression.

*Continuation*: Critics judged the book about the (Civil War/Great Depression) from the (Boston/Houston) publisher to be the best of the bunch.

12. Experiment 1: critical, Experiment 2, Experiment 3

*Context*: Mike's doctor told him that he should get more exercise, so Mike considered walking or biking to work. He also thought about swimming or hiking.

*Continuation*: The only way he could fit all those activities into his schedule, though, was to (walk/bike) to work and (swim/hike) on the weekends. After a month, Mike's doctor was quite pleased with his progress.

13. Experiment 1: critical, Experiment 2, Experiment 3

*Context*: To win the hand of the Baroness's daughter, the English and the Scottish knights competed in a tournament of jousting and archery.

*Continuation*: Both knights gave it their best, but the (English/Scottish) knight emerged victorious during the (jousting/archery) competition and married the daughter.

14. Experiment 1: critical, Experiment 2, Experiment 3

*Context*: Stephanie had planned on going to the grocery store and post office yesterday afternoon. However, when she left her house, she saw the snowstorm had left many of the streets covered with snow and ice.

*Continuation*: She managed to complete one of her errands, but since the (grocery store/post office) was surrounded by (snow/ice), she decided to come back another day rather than risk driving.

15. Experiment 1: critical, Experiment 2, Experiment 3

*Context*: The local bowling lanes had gotten to be quite popular. The lanes were inexpensive and open to everyone of all skill levels, as long as they obeyed two simple rules: no outside food or drinks were allowed, and everyone had to wear bowling shoes. Thanks to these friendly policies, bowling leagues had been formed by both the firefighters and the architects.

*Continuation*: But when some of the (architects/firefighters) were found with (drinks/shoes) that were not allowed, management had no choice but to temporarily ban them from the lanes.

16. Experiment 1: critical, Experiment 2, Experiment 3

*Context*: Originally, the space probe was designed to fly past Mars and Jupiter and send photographs and videos back to NASA from both planets.

*Continuation*: However, due to a glitch in the system, the (photos/videos) from (Mars/Jupiter) were lost.

17. Experiment 1: critical, Experiment 2, Experiment 3

*Context*: Before each game, the quarterback always polished his helmet and shoes to give himself good luck. He thought that if he didn't, he might throw an interception or fumble the ball.

*Continuation*: So, when he lost the ball to (a fumble/an interception) during the second quarter, he blamed it on the fact that he had forgotten to polish his (helmet/shoes) before the game.

18. Experiment 1: critical

*Context*: When Janice was vacationing in Japan, she visited both a beautiful temple and an ancient castle. She had promised her friends that she would bring back lots of photographs and souvenirs.

*Continuation*: So, she was disappointed when she was not able to get any (photos/souvenirs) at the (temple/castle). The rest of the trip, however, was fantastic.

19. Experiment 1: critical, Experiment 2, Experiment 3

*Context*: Elizabeth was going to bake a cake for her best friend's birthday, but by the time she made it to the store, she couldn't remember if the recipe called for vanilla extract or lemon extract, and if she needed two eggs or three eggs. She bought everything she might have needed.

*Continuation*: When she got home, she checked the recipe and saw that it used (vanilla/lemon) extract and (two/three) eggs.

20. Experiment 1: critical, Experiment 2, Experiment 3

*Context*: After the McKee Company struck business deals with firms in China and South Korea, the company hired Suzanne as translator. When the company received the initial contract and invoice from one of its partners...

*Continuation*: ...her first job was to translate the (contract/invoice) from (Chinese/Korean).

21. Experiment 1: critical, Experiment 2

*Context*: The newlyweds had considered both Hawai'i and Italy for their honeymoon.

*Continuation*: But, they ultimately chose (Hawai'i/Italy) because the (bride/groom) had never been there before.

22. Experiment 1: critical, Experiment 2, Experiment 3

*Context*: The German and the Japanese engineers had been competing to build the first working version of the new computer chip. Neither had been able to figure out how to make the chip small enough or keep it cool enough.

*Continuation*: The crucial breakthrough came out when the (Germans/Japanese) figured out how to make the chip (smaller/cooler), and the chip was soon rushed to production.

23. Experiment 1: critical, Experiment 2, Experiment 3

*Context*: A new Mexican and a new Greek restaurant had recently opened in the city. Both were waiting to hear whether or not the notoriously harsh food critic would give his approval to their special and entrees. The critical originally planned to dine at both restaurants during the week.

*Continuation*: But because he caught the flu, he only had a chance to visit the (Mexican/Greek) restaurant, where he gave the (specials/entrees) a favorable review.

24. Experiment 1: critical

*Context*: The students thought that between the calculus homework and the English paper, they had too much homework. They decided to ask their teachers to either postpone the homework or make it shorter.

*Continuation*: The teachers agreed to (postpone/shorten) the (English/calculus) assignment.

25. Experiment 1: filler

*Context*: Rachel's home business was quite successful, but she often had a hard time finding enough time during the day to get everything done. Yesterday, she had some letters and some packages to mail during her trips to UPS and FedEx, and she was supposed to meet a client for lunch.

*Continuation*: She mailed the (letters/packages) at (UPS/FedEx) and then hurried to the restaurant to meet her client.

26. Experiment 1: filler

*Context*: Expectations were high for the Danish ski team as the Olympics began. The famed veteran, who was still recovering from an injury, would be competing against the up-and-coming rookie.

*Continuation*: The race was thrilling and everyone was surprised when the (rookie/veteran) won the (gold/silver).

27. Experiment 1: filler

*Context*: Mary was having an awful time trying to renew her driver's license. She would been on the phone all day trying to figure out if she was supposed to pay by mail or in person. She talked to several people at the DMV, and no one could tell her whether to pay by check or by credit card.

*Continuation*: Finally, she was able to talk to a customer service representative who explained that she could pay (by mail/in person) using a (check/credit card).

28. Experiment 1: filler, Experiment 2, Experiment 3

*Context*: Steve and his wife had been wanting to visit the Grand Canyon and the Everglades. Steve's employer told him he could take a week of vacation either in the spring or in the fall.

*Continuation*: Steve and his family considered their options before eventually deciding to visit the (Everglades/Grand Canyon) in the (spring/fall).

29. Experiment 1: filler, Experiment 2, Experiment 3

*Context*: Annette completed all her holiday shopping her brother and her father while on vacation in Germany. She stopped in a gift shop in Munich and bought a T-shirt and a book of German item.

*Continuation*: She decided to give her (brother/father) the (book/T-shirt). He was very happy to get it and said it was his favorite of the gifts he received that year.

30. Experiment 1: filler, Experiment 2, Experiment 3

*Context*: Two best friends from college, Matt and Eric, ended up as rival salesmen at Coke and Pepsi.

*Continuation*: They had a friendly competition going, but the winner was almost always (Matt/Eric), who worked for (Coke/Pepsi).

31. Experiment 1: filler, Experiment 2, Experiment 3

*Context*: The cop surveyed the room where the body was found. It wasn't obvious whether it was a murder or a suicide, so he sent the fingerprints and blood samples he found on the gun to forensics.

*Continuation*: Analysis of the (fingerprints/blood) revealed the death to be a (murder/suicide).

32. Experiment 1: filler, Experiment 2, Experiment 3

*Context*: Andrea never enjoyed flying, but with both a professional conference and a wedding to attend this month, she knew she'd have to grin and bear it. She just hoped her flights didn't get delayed or canceled.

*Continuation*: Unfortunately, her fears were confirmed when, on her way to the (conference/wedding), her flight was (delayed/canceled).

33. Experiment 1: filler, Experiment 2, Experiment 3

*Context*: When the rock band first formed, the bandleader both sang and played guitar. Later, after Lauren and Chris joined the band...

*Continuation*: ...(Lauren/Chris) took over as (vocalist/guitarist).

34. Experiment 1: filler, Experiment 2

*Context*: Before the presidential debate, the moderator prepared a number of questions for the two candidates. She selected a number of questions about both Social Security and the environment.

*Continuation*: During the debate, she decided to ask the (Democrat/Republican) about (Social Security/the environment).

35. Experiment 1: filler, Experiment 2

*Context*: When the book club was first formed, its members decided they should read both fiction and non-fiction works. They also wanted to read a mix of classics and new books.

*Continuation*: So, the leaders of the club decided that the first book they would read would be a (classic/new) work of (fiction/non-fiction), but that they would read something different for the book after that.

36. Experiment 1: filler, Experiment 2, Experiment 3

*Context*: Although Jennifer owned both a cat and a rabbit, the two pets got along great with each other. There wasn't a problem until her cousin visited and brought along her ferret and bird.

*Continuation*: Jennifer's (cat/rabbit) hated the cousin's (ferret/bird), and chaos broke out in the house.

37. Experiment 1: filler, Experiment 2, Experiment 3

*Context*: When Mary and Jessica saw each other at the high school reunion, they were surprised and amused by each other's career choices. Both of them had swore they'd never end up as accountants or secretaries.

*Continuation*: But, sure enough, (Mary/Jessica) was now (an accountant/a secretary).

38. Experiment 1: filler, Experiment 2

*Context*: Michelle had finally set aside enough money to buy a new computer, but she wasn't sure whether she wanted a PC or a Mac, and if she should get a desktop or a laptop computer. She talked to her friend Jim, who was a computer expert.

*Continuation*: Jim recommended a (PC/Mac) (desktop/laptop) as best suiting her needs.

39. Experiment 1: filler, Experiment 2, Experiment 3

*Context*: Tina was disappointed when she realized that both the concert and the play were both scheduled for Thursday evening, and that she could only go to one of them. She figured she would just go to the one that was closest, but then her boyfriend suggested the cheaper alternative to save money.

*Continuation*: By Thursday morning, she still hadn't decided, but she and her boyfriend eventually decided on the (concert/play) since it was (closer/cheaper).

40. Experiment 1: filler, Experiment 2, Experiment 3

*Context*: The meteorologist had predicted bad weather for the weekend, saying that there would either be sleet or hail. Julia thought about canceling her trip to the beach when she heard the forecast.

*Continuation*: But she was glad that she didn't, because the only bad weather ended up being the (sleet/hail) on (Saturday/Sunday).

41. Experiment 1: filler, Experiment 2, Experiment 3

*Context*: Ben had made it to the last round of the game show. Now, he had to choose to open either the blue door or the red door to claim his prize. Behind one of the doors

was a new car and behind the other was a goat. But he only got one chance to pick.

*Continuation*: Nervously, Ben opened the (blue/red) door and discovered the (goat/car).

42. Experiment 1: filler, Experiment 2, Experiment 3

*Context*: The renowned fashion designer divided his time between Venice and Paris while working on his new lines of shirts and dresses.

*Continuation*: When the designs had been finished, he decided to unveil the (shirts/dresses) at an upcoming show in (Venice/Paris).

43. Experiment 1: filler, Experiment 2

*Context*: The grade school class took a trip to the natural history museum, where the children were enthralled by the dinosaur skeleton. They had just been learning about what different animals ate, so the children wanted to know whether the dinosaur ate plants or animals, and whether it used sight or smell to look for food.

*Continuation*: Happy to see the children so excited, the museum guide explained that this species ate (plants/animals) that it found by (sight/smell).

44. Experiment 1: filler, Experiment 2, Experiment 3

*Context*: A sporting goods manufacturer was looking for some new athletes for its lines of jackets and hats. Representatives from the company met with a golfer and a baseball pitcher...

*Continuation*: ...before the company decided to sign the (golfer/pitcher) to endorse the (jackets/hats).

45. Experiment 1: filler, Experiment 2, Experiment 3

*Context*: Bridget was planning a day in the city, since she'd been thinking about the museum or the art gallery. She knew there would be a lot of traffic, so she planned to take either the bus or the train instead of driving.

*Continuation*: After checking a city map, she decided the best plan was to take the (bus/train) to the (museum/gallery).

46. Experiment 1: filler, Experiment 2, Experiment 3

*Context*: The fraternity and sorority were competing to raise the most money to fight cancer and AIDS, so they held a number of fundraisers.

*Continuation*: The most successful fundraiser was the haunted house, which helped the (fraternity/sorority) raise money for (cancer/AIDS).

47. Experiment 1: filler, Experiment 2, Experiment 3

*Context*: Dorothy's car had been hit in an accident. Fortunately, she was unharmed thanks to her seat belt and airbag, and it didn't seem like the damage to the car was too bad either. Just to make sure, though she took the car to a mechanic to make sure that the engine and transmission were still working fine. The mechanic told her the collision had been worse than she'd thought.

*Continuation*: The (engine/transmission) was broken and she'd have been seriously injured if she hadn't been using the (seatbelt/airbag).

48. Experiment 1: filler, Experiment 2, Experiment 3

*Context*: The city was thrilled when both the poet and the novelist who lived there won Pulitzer Prizes in the same year. The mayor planned to hold a parade and banquet in their honor.

*Continuation*: However, shunning publicity, the (poet/novelist) declined to attend the (parade/banquet).

49. Experiment 1: filler, Experiment 2, Experiment 3

*Context*: A new children's movie had just been released about a farm where the cows and horses can all talk. The farm is threatened by bankruptcy and crop disease.

*Continuation*: However, the efforts of a heroic (cow/horse) end up saving the farm from (bankruptcy/disease).

50. Experiment 1: filler, Experiment 2, Experiment 3

*Context*: Brad went fishing at the lake this past weekend, hoping to catch some catfish or bass. He hadn't been fishing in a long time, though, and was out of practice. Most of the fish that he hooked got away.

*Continuation*: But, he did manage to reel in a few (catfish/bass) during the (morning/afternoon).

## Appendix B

*Test probes for Experiment 3*

2A. The local parks commission decided to build a (playground/skatepark/dog park).

2B. The local parks commission decided to focus its resources on (tidying/expanding/landscaping) the parkland.

3A. The research labs discovered that the cause of the respiratory disease was a (bacteria/virus/fungus).

3B. The research labs discovered that the respiratory disease infected the (lungs/throat/mouth).

5A. The forthcoming film ran into trouble when the (director/producer/star) quit.

5B. Someone quit the forthcoming film because of a dispute over the (budget/screenplay/marketing).

6A. The (doctors/nurses/patients) at the hospital were particularly pleased by the renovations.

6B. People at the hospital were particularly pleased by the (brighter/cooler/larger) rooms.

7A. The new (jungle gym/slide/swingset) at Jefferson Elementary was particularly popular.

7B. The (boys/girls/teachers) at Jefferson Elementary always gathered at the same place during every recess period.

8A. The (British/French/Portuguese) scientists spotted the endangered monkey and tagged it.

8B. The endangered monkey was finally spotted in (Malaysia/Indonesia/the Philippines).

11A. Critics were especially pleased by the new book about the (Civil War/Great Depression/American Revolution).

11B. The history book that the critics liked best was from the publisher in (Boston/Houston/Philadelphia).

12A. To get more exercise, Mike decided to (walk/bike/jog) to work.

12B. To get more exercise, Mike decided to (swim/hike/climb) on the weekends.

13A. The (English/Scottish/Welsh) knight married the baroness's daughter.

13B. The competition to marry the baroness's daughter was resolved by the (jousting/archery/fencing) contest.

14A. Because of the storm yesterday, Stephanie couldn't make it to the (grocery store/post office/gas station).

14B. Stephanie couldn't complete all of her errands yesterday because of the (snow/ice/debris) on the road.

15A. The (architects'/firefighters'/electricians') bowling league was temporarily banned from the local bowling lanes due to a rule violation.

15B. The management of the bowling lanes had to ban one of the leagues for bringing (drinks/shoes/food) not allowed.

16A. NASA lost some of the (photos/videos/measurements) from the space probe due to a glitch.

16B. Due to a glitch in the space probe, NASA lost some of the recordings from (Mars/Jupiter/Saturn).

17A. The quarterback blamed a (fumble/interception/penalty) on the fact that he hadn't polished everything for good luck.

17B. The quarterback blamed his bad performance during the second quarter on the fact that he hadn't polished his (helmet/shoes/ring) before the game.

19A. Elizabeth baked a cake for her best friend's birthday that used (vanilla/lemon/almond) extract.

19B. Elizabeth baked a cake for her best friend's birthday that used (two eggs/three eggs/one egg).

20A. Suzanne's first job as a translator at the McKee Company was to translate (a contract/an invoice/a prospectus).

20B. The McKee Company hired Suzanne to translate documents from (Chinese/Korean/Vietnamese).

22A. The (Germans/Japanese/Dutch) were responsible for the crucial breakthrough in the development of the new computer chip.

22B. The breakthrough in developing the new computer chip was making the chip (smaller/cooler/faster).

23A. Because the critic caught the flu, he only had a chance to visit the (Mexican/Greek/Indian) restaurant.

23B. The food critic gave a favorable review to the (specials/entrees/desserts) at one of the new restaurants.

28A. Steve and his wife decided to go to (the Everglades/the Grand Canyon/Yellowstone) on vacation.

28B. Steve and his wife decided to visit a national park for vacation during the (spring/fall/summer).

29A. Annette gave her (brother/father/son) a gift from the gift shop in Munich that he was very happy to get.

29B. Annette's gift of a (shirt/book/mug) from Munich was a favorite for its recipient.

30A. The winner of the sales contest between college friends was almost always (Matt/Eric/Dan).

30B. The winner of the sales contest between college friends was almost always the (Coke/Pepsi/Dr. Pepper) salesman.

31A. The forensics team discovered that the death was the result of a (murder/suicide/accident).

31B. After receiving the gun, the forensics team determined the cause of death based on the (fingerprints/blood/ballistics).

32A. Andrea had to suffer through a (delayed/canceled/redirected) flight while traveling this month.

32B. Andrea encountered flight problems on her way to a (conference/wedding/job interview).

33A. The leader of the rock band changed roles when someone else took over as (vocalist/guitarist/keyboardist).

33B. After joining the rock band, (Lauren/Chris/Dave) took over one of the bandleader's roles.

36A. Jennifer's (cat/rabbit/hamster) had a problem with one of her cousin's pets.

36B. Chaos broke out at Jennifer's house because of her cousin's (ferret/bird/dog).

37A. At the high school reunion, the friends were amused by (Mary/Jessica/Catherine)'s career choice.

37B. At the high school reunion, the friends were surprised that one of them was now (an account/a secretary/a banker).

39A. Tina and her boyfriend made their decision about what to do on Thursday by choosing the (closer/cheaper/shorter) event.

39B. Tina and her boyfriend didn't decide to attend the (concert/play/movie) until Thursday morning.

40A. The only bad weather during Julia's trip to the beach was on (Friday/Saturday/Sunday).

40B. The only bad weather during Julia's trip to the beach was the (sleet/hail/rain).

41A. On the game show, Ben chose the prize behind the (blue/red/green) door.

41B. The door that Ben opened on the game show had a (goat/car/boat) behind it as the prize.

42A. The fashion designer decided where he would unveil his new line of (shirts/dresses/shoes).

42B. The fashion designer planned to unveil some of his new work at an upcoming show in (Venice/Paris/London).

44A. The sporting goods manufacturer decided to sign the (golfer/pitcher/swimmer) to endorse one of its products.

44B. After the meetings, the sporting goods manufacturer signed an athlete to endorse its (jackets/hats/watches).

45A. There was a lot of traffic in the city, so Bridget took the (bus/train/subway) instead of driving.

45B. Bridget visited the (museum/gallery/planetarium) during her day in the city.

46A. The haunted house fundraiser was organized by the (fraternity/sorority/dorm).

46B. The haunted house fundraiser raised money to fight (cancer/AIDS/MS).

47A. The car accident was worse than Dorothy had thought because of the damage to her car's (engine/transmission/carburetor).

47B. Dorothy would have been hurt in the car accident if it hadn't been for her (seatbelt/airbag/fenders).

48A. The (parade/banquet/press conference) was not attended by one of the city's Pulitzer Prize winners.

48B. The (poet/novelist/composer) who won the Pulitzer Prize shunned publicity and declined to attend one of the celebratory events.

49A. In the new children's movie, a talking (cow/horse/chicken) saves the farm.

49B. In the new children's movie, a talking animal saves the farm from (bankruptcy/a disease/an earthquake).

50A. Brad caught a few (catfish/bass/trout) at the lake this past weekend.

50B. When Brad went fishing at the lake this past weekend, he caught most of the fish during the (morning/midday/afternoon).

## References


Almor, A., & Eimas, P. D. (2008). Focus and noun phrase anaphors in spoken language comprehension. *Language and Cognitive Processes, 23*, 201–225.

Bates, D., Maechler, M., & Dai, B. (2008). *lme4: Linear mixed-effects models using s4 classes [computer software manual]*. <http://lme4.r-forge.r-project.org>. (R package version 0.999375-28). Retrieved.

Beckman, M. E., & Elam, G. A. (1997). *Guidelines for ToBI labelling (version 3.0)*. <http://www.ling.ohio-state.edu/~tobi/ame_tobi/labelling_guide_v3.pdf>.

Birch, S. L., & Garnsey, S. M. (1995). The effect of focus on memory for words in sentences. *Journal of Memory and Language, 34*, 232–267. doi:10.1006/jmla.1995.1011.

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

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433–436. doi:10.1163/156856897X00357.

Brainerd, C. J., Reyna, V. F., & Estrada, S. (2006). Recollection rejection of false narrative statements. *Memory, 14*, 672–691. doi:10.1080/09658210600648449.

Braun, B., & Tagliapietra, L. (in press). The role of contrastive intonation contours in the retrieval of contextual alternatives. *Language and Cognitive Processes*. doi:10.1080/01690960903036836.

Chafe, D. (1974). Language and consciousness. *Language, 50*, 111–133.

Chen, A., den Os, E., & de Ruiter, J. P. (2007). Pitch accent type matters for online processing of information status: Evidence from natural and synthetic speech. *The Linguistic Review, 24*, 317–344. doi:10.1515/TLR.2007.012.

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

Cutler, A., & Fodor, J. A. (1979). Semantic focus and sentence comprehension. *Cognition, 7*, 49–59.

Cutler, A., & Isard, S. D. (1980). The production of prosody. In B. Butterworth (Ed.), *Language production. Speech and talk* (Vol. 1, pp. 245–269). London: Academic Press.

Dahan, D., Tanenhaus, M. K., & Chambers, C. G. (2002). Accent and reference resolution in spoken-language comprehension. *Journal of Memory and Language, 47*, 292–314.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Gussenhoven, C. (2005). Transcription of Dutch Intonation. In S.-A. Jun (Ed.), *Prosodic typology and transcription: A unified approach* (pp. 118–171). Oxford: Oxford University Press.

Haviland, S. E., & Clark, H. H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior, 13*, 512–521. doi:10.1016/S0022-5371(74)80003-4.

Ito, K., & Speer, S. R. (2008). Anticipatory effects of intonation: Eye movements during instructed visual search. *Journal of Memory and Language, 58*, 541–573. doi:10.1016/j.jml.2007.06.013.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformations or not) and towards logit mixed models. *Journal of Memory and Language, 59*, 434–446. doi:10.1016/j.jml.2007.11.007.

Kamas, E. N., Reder, L. M., & Ayers, M. S. (1996). Partial matching in the Moses illusion: Response bias not sensitivity. *Memory & Cognition, 24*, 687–699.

Labov, W., Ash, S., & Boberg, C. (2006). *The atlas of North American English: Phonetics, phonology, and sound change*. New York: Walter de Gruyter.

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

Ladd, D. R., & Schepman, A. (2003). "Sagging transitions" between high pitch accents in English: Experimental evidence. *Journal of Phonetics, 31*, 81–112. doi:10.1016/S0095-4470(02)00073-6.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory* (2nd ed.). New York: Erlbaum.

McAteer, E. (1992). Typeface emphasis and information focus in written language. *Applied Cognitive Psychology, 6*, 345–359. doi:10.1002/acp.2350060406.

Park, H., & Reder, L. M. (2004). Moses illusion: Implication for human cognition. In R. F. Pohl (Ed.), *Cognitive illusions*. Hove: Psychology Press.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437–442. doi:10.1163/156856897X00357.

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

Pitrelli, J. F., Beckman, M. E., & Hirschberg, J. (1994). Evaluation of prosodic transcription labeling in the ToBI framework. In *Proceedings of the third international conference on spoken language processing* (pp. 123–126).

R Development Core Team (2008). *R: A language and environment for statistical computing [Computer software manual], Vienna, Austria*. <http://www.r-project.org> (ISBN:3-900051-07-0). Retrieved.

Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics, 1*, 75–116.

Rooth, M. (1999). Association with focus or association with presupposition? In P. Bosch & R. van der Sandt (Eds.), *Focus: Linguistic, cognitive, and computational perspectives* (pp. 232–244). Cambridge, United Kingdom: Cambridge University Press.

Sanford, A. J. S., Price, J., & Sanford, A. J. (2009). Enhancement and suppression effects resulting from information structuring in sentences. *Memory & Cognition, 37*, 880–888. doi:10.3758/MC.37.6.880.

Sanford, A. J. S., Sanford, A. J., Molle, J., & Emmott, C. (2006). Shallow processing and attention capture in written and spoken discourse. *Discourse Processes, 42*, 109–130. doi:10.1207/s15326950dp4202_2.

Sanford, A. J., & Sturt, P. (2002). Depth of processing in language comprehension: Not noticing the evidence. *Trends in Cognitive Sciences, 6*, 382–386. doi:10.1016/S0960-9822(06)00291-0.

Schwarzschild, R. (1999). GIVENness, AVOIDF, and other constraints on the placement of accent. *Natural Language Semantics, 7*, 141–177.

Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition, 71*, 109–147. doi:10.1016/S0010-0277(99)00025-6.

Selkirk, E. (1986). On derived domains in sentence phonology. *Phonology Yearbook, 3*, 371–405.

Selkirk, E. (2002). Contrastive *FOCUS* vs. presentational *focus*: Prosodic evidence from right node raising in English. In *Speech prosody*, Aix-en-Provence (pp. 643–646).

Silverman, K., Beckman, M., Pitrelli, J., Ostendorf, M., Wightman, C., Price, P., et al. (1992). ToBI: A standard for labeling English prosody. In *Proceedings of the third international conference on spoken language processing* (pp. 867–870).

Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends in Cognitive Science, 1*, 261–267.

Sturt, P., Sanford, A. J., Stewart, A., & Dawydiak, E. (2004). Linguistic focus and good-enough representations: An application of the change-detection paradigm. *Psychonomic Bulletin & Review, 11*, 882–888.

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. doi:10.1126/science.7777863.

Wagner, M., & Watson, D. G. (in press). Experimental and theoretical advances in prosody: A review. *Language and Cognitive Processes*. doi:10.1080/01690961003589492.

Watson, D., Gunlogson, C., & Tanenhaus, M. (2008). Interpreting pitch accents in on-line comprehension: $H^*$ vs. $L+H^*$. *Cognitive Science, 32*, 1232–1244.

Weber, A., Braun, B., & Crocker, M. W. (2006). Finding referents in time: Eye-tracking evidence for the role of contrastive accents. *Language and Speech, 49*, 367–392. doi:10.1177/00238309060490030301.

Wright, D. B., Horry, R., & Skagerberg, E. M. (2008). Functions for traditional and multilevel approaches to signal detection theory. *Behavior Research Methods, 41*, 257–267.