The implicit prosody of corrective contrast primes appropriately intonated probes (for some readers)

Speer, S.R.; Foltz, A.; Frazier, L.; Gibson, E.

Explicit and Implicit Prosody in Sentence Processing

Published: 31/07/2015

Other version

Dyfyniad o'r fersiwn a gyhoeddwyd / Citation for published version (APA):
Speer, S. R., Foltz, A., Frazier, L. (Ed.), & Gibson, E. (Ed.) (2015). The implicit prosody of corrective contrast primes appropriately intonated probes (for some readers). In *Explicit and Implicit Prosody in Sentence Processing* (2015 ed., pp. 263-285). Springer.

The implicit prosody of corrective contrast primes appropriately intonated probes - for some readers

Shari R. Speer, The Ohio State University

Anouschka Foltz, Bielefeld University

Abstract

Studies of silent sentence reading have shown indirect evidence for the impact of a default projected "implicit prosody" on sentence processing, such as longer processing times for sentences with final syntactic interpretations inconsistent with their assumed default prosody (e.g. Fernández, 2003). While explanations of such effects associate them with the presence of an auditory image, or "inner speech", there is no direct evidence tying such an image directly to the measured effects--that is, there is little information about how implicit prosody "sounds" in the mind of the reader. In two visual-to-auditory cross-modal priming experiments, we look for evidence of a link between the implicit prosodic contour generated during silent reading and the explicit prosodic contour pronounced on a subsequently-presented auditory probe word. Pairs of text sentences that contained corrective contrasts (e.g. *Jacquelyn didn't pass the test. Belinda passed the test.*) were immediately followed by probes pronounced with pitch accent patterns consistent (BELINDA) or inconsistent (BELINDA) with the corrective contrast in the read text. Results in an initial experiment showed no priming from implicit to overt prosody, but instead a processing advantage for contrasts in verb position over those in subject position, as well as an advantage for no-accent probes over contrastively accented ones. The experiment was repeated with participants separated into two groups on the basis of the overt prosody they produced when reading a passage aloud. Results showed that individual differences in the read-aloud production correlate with the response pattern in the cross-modal task. The results provide some initial evidence about the nature of the auditory image produced as inner speech during silent reading.

Introduction

The Implicit Prosody Hypothesis (IPH) (Fodor, 1998, 2002) proposes that a 'default' prosodic contour is projected onto sentences during silent reading, and predicts a relationship between the projected contour and the final interpretation of the text. The IPH was initially proposed to explain the effect of constituent length on attachment decisions made after reading ambiguous relative clause (RC) sentences, such as *Someone shot the servant of the actress who was on the balcony (drinking tea)*. Such off-line judgments show that readers show a stronger preference for low-attachment of a short ambiguous RC than of a longer one (e.g. Fernández & Bradley, 1999). According to the IPH a longer sentence-final RC induces a prosodic phrase break before the RC, setting it off from the two preceding noun phrases. Numerous experiments examining reader's syntactic judgments for sentences in many languages (e.g. Fernández & Bradley, 1999 (Spanish); Hirose, 1999 (Japanese); Wijnen, 2004 (Dutch); Vasishth, Agnihotri, Fernández, & Bhatt, 2004 (Hindi), among others) have supported the IPH. Online studies of sentence reading have also shown indirect evidence for implicit prosody effects, such as shorter reading times for sentences with final syntactic interpretations that were consistent with their assumed default prosodic phrasing, as compared to those inconsistent with that phrasing (e.g. Fernández, 2003).

Although explanations of implicit prosody effects associate them with the presence of an auditory image, or "inner speech" in the mind of the reader, there is no direct evidence tying such an image to the measured off-line judgment and reading time effects. Such evidence would be valuable, as it would increase our understanding of the source of implicit prosodic contours within the language processing system, and might lend support to claims about the nature and location of the specific prosodic features posited (such as accents and phrasing), on which implicit prosody-based arguments rest. There is some evidence that an auditory image generated during silent reading can affect subsequent processing of related auditory material. Abramson (2007) asked participants to listen to a pair of male and female interlocutors who each uttered one statement and one question. Participants then read silently a set of sentences that began with either 'He said' or 'She said' and were punctuated to indicate either a question or a statement. They were told the text items had been spoken by the interlocutors previously heard. After a 5 minute delay, they gave auditory lexical decisions to words that had been final in the silently read sentences, pronounced with either rising or falling final intonation by a male or a female speaker. Lexical decision times were facilitated when sex, intonation or both matched the text presentation of the words, as compared to non-matching items. However, in this design it is difficult to determine if the auditory lexical decisions were primed by the initial prosody of the interlocutor's intonation, the implicit prosody generated during the reading of the text sentences, or some combination of these factors. In addition, effects due to the evoked memory for the voice of a particular speaker that is not the reader may be different from those generated during normal reading.

Most studies on implicit prosody, however, provide only indirect information about such “inner speech”. For example, several investigators have experimentally compared the overt prosodic phrasing readers produce when reading a sentence aloud to their implicit prosody (as determined by their preferred syntactic parse of the sentence during silent reading). These studies have shown mixed results. For example, Jun and Kim (2004; see also Hwang & Schafer, 2009, and Jun & Koike, 2003, for similar experiments with Japanese) conducted production experiments with Korean relative clauses, recording readers who either skimmed a text and then read it aloud, or read aloud without skimming. Participants then completed an off-line syntactic judgment task for the same sentences. Results showed that judged syntactic preferences and Korean ToBI-annotated overt prosodic phrasing patterns were consistent with one another - about a two-thirds preference for high-attachment readings and a corresponding phrasal break location about two-thirds of the time, with skimmers showing a stronger high-attachment preference. Here, the correlation between offline preference and overt produced prosodic phrasing supports the predictions of the IPH. However, similar work on English has not consistently supported IPH predictions. The most frequently produced prosodic phrasing pattern for ambiguous RC sentences in English read-aloud studies is one where the RC is set off in its own phrase (NP1 NP2)(RC), a pattern consistent with a high attachment preference, when English is generally considered to 'prefer' low-attachment (Jun, 2010). Bergmann and Ito (2007, 2009) asked participants to read ambiguous RC sentences aloud and manipulated the length of N1 and of the RC (e.g. *The (defense) lawyer of the mayor who smokes (like a chimney) impressed the guest*). After reading aloud, participants answered a comprehension question indicating attachment, and overwhelmingly gave low attachment interpretations. In contrast, prosodic boundaries were produced much more frequently at N2 than N1 - the pattern expected for high attachment interpretations. In addition, longer N1s generated more prosodic boundaries at N1, but longer RCs did not systematically generate more prosodic boundaries at N2. This pattern of results is inconsistent with the IPH explanation of the low attachment preference for English. Follow-up studies (Bergmann, Ito, & Maday, 2008; Foltz, Maday, & Ito, 2011) had participants either (a) read aloud an ambiguous RC sentence and then give a syntactic judgment, or (b) silently read the sentence, then give a syntactic judgment, and then read the sentence aloud. Only for (a) were prosodic patterns correlated with readers’ interpretations. In another experiment, a separate group of participants listened to the read sentences and judged the speakers' intended syntax. Only the prosodic patterns of the (a) sentences modulated participants’ interpretations. Although these findings are inconsistent with the notion of a predictable default prosodic contour, it is not possible to know whether parsing preferences in reading were due to "inner speech" contours, because the silent and overt prosodies were necessarily produced on different trials.

Jun (2010) has argued that it might not be possible to assess implicit prosody by comparing readers' overt read-aloud prosody to the judgments they make while or after silent reading. The prosody of read speech is easily recognizable as such, and differs substantially from spontaneous speech and "laboratory speech"(speech created for use in experiments), with shorter constituent phrases (so more pitch accents and breaks) (Howell & Kadi-Hanifi, 1991). In both laboratory speech and spontaneous speech, the speaker begins with an intended message-level meaning that contributes to the generation of a corresponding prosodic structure. In contrast, reading aloud involves recovering a meaning provided by the writer, often a bit at a time as the words are recognized and produced. Thus the overt prosody/message mapping may be shallow and based on minimal constraints as compared to that for implicit prosody, which readers may generate at any point during text comprehension. In addition, goals in reading aloud may differ from goals for silent reading, with the former more focused on articulatory processes, such as preventing word-level pronunciation errors, and the latter more on message comprehension. Jun's (2010) production results showed that readers probably produced prosodic breaks at NP2 in long RC sentences before they had comprehended the lexical content of the RC, thus producing an infelicitous break pattern for the syntactic structure. This suggests that length constraints rather than meaning controlled the prosody produced. Some of these readers also produced hesitation lengthening later in the same utterances, suggesting the operation of a prosodic repair process based on self-monitoring as the overt prosody unfolds. These issues could easily result in stark differences between overt and implicit prosodies for read text, making it difficult to learn about the prosodic phrasal structure and accent patterns available during silent reading on the basis of productions gathered during reading aloud.

Individual differences may also contribute to inconsistent findings when comparing implicit to overt prosody, because readers may differ from one another and from themselves on repeated trials in the prosody they assign to a particular utterance. The fact that speakers produce a variety of prosodic structures for any given lexico-syntactic sequence is well-established for the overt prosody of spontaneous and quasi-spontaneous speech. For example, 13 participants in a game task were restricted to the use of particular syntactic frames, but could insert lexical items to convey their intended meaning. ToBI annotations of pitch accents and phrasal tones showed that when pronouncing a PP attachment ambiguity, they used 62 different prosodies for 78 productions of the high-attached version, and 87 different prosodies for 101 productions of the low-attached version (Schafer, Speer, Warren, & White, 2000; Speer, Warren, & Schafer, 2011). Similarly, when 10 uninstructed speakers in a separate study gave instructions to an interlocutor in a tree-decoration task, they produced 580 adjective-noun sequences in contrastive contexts (e.g. *blue ball* preceded by *green ball*). Annotation showed 223 prosodic patterns across these utterances (Ito & Speer, 2008). There is some evidence that readers may differ in the implicit prosody they assign to text sentences as well. Swets, Desmet, Hambrick, & Ferreira (2007) showed differences in attachment preferences for ambiguous relative clause sentences depending on the working memory capacity of the reader. When NP1 NP2 RC sentences were presented in their entirety, participants with less memory capacity showed a stronger high-attachment preference than participants with greater capacity. The authors suggested that low capacity participants had generated an implicit prosodic break between NP2 and the RC, while high capacity participants were able to process the sentence as an integrated unit with a low-attachment final parse.

The role of implicit prosody in silent reading can also be seen in effects due to lexical stress, the pattern of strong and weak syllables associated with individual words in languages like English, Dutch and German. Findings here are more consistent than those for the syntax-implicit prosody correspondence summarized above. The number of stressed syllables in a word has been demonstrated to affect reading time, such that a word with two stressed syllables (e.g. *RAdiAtion*) takes longer to read than one with the same number of syllables, but only one stressed syllable (e.g. *inTENsity*); this effect was interpreted to indicate that the time to prepare the implicit pronunciation of a word depends on the number of stressed syllables it contains, as the overall number of syllables in a word did not itself correlate with reading time (Ashby & Clifton, 2005). Additional evidence for the influence of the phonological representation of words on processing during silent reading comes from studies of stress-alternating verb-noun homographs. Breen & Clifton (2011, 2013) found longer reading times when sentence context necessitated a revision of syntactic category that included a revision of lexical stress pattern than when it did not (e.g. revision of the noun *ABstract* to the verb *abSTRACT* was more costly than revision of the noun *rePORT* to the verb *rePORT*). In a separate experiment, they used limerick contexts to manipulate metrical expectations for the lexical stress pattern of a phrase-final homograph. For example, consistent context for the verb *present* was (SMALL CAPS indicate metrically strong syllables): *There ONCE was a CLEVer young GENT // who HAD a nice TALK to preSENT*, while an inconsistent context was: *There ONCE was a PENniless PEASant // who WENT to his MASter to preSENT*. Evidence from eyetracking showed that when the metrically predicted location for lexical stress was inconsistent with the syntactic category of the phrase-final word, reading times were longer than when metrical context was consistent. This difference was not shown for phrase-final noun homographs in metrically biasing contexts -- that is, a processing penalty was found for revising from a SW to a WS lexical stress pattern, but not for revising from a WS to a SW pattern. These studies clearly implicate a conflict between implicit prosody and the final interpretation of a silently read sentence as the source of a processing deficit, and do so more convincingly than the previously discussed evidence from relative clause processing. However, there are two possible sources of the conflicting implicit prosody. On the one hand, sentence level metrical structure (implemented in spoken sentences by the alignment of pitch accents with particular lexically-stressed syllables) was manipulated to create consistent and inconsistent contexts for the critical words. This would suggest that the effects were due to a sentence-level implicit prosodic representation. On the other hand, the critical words' lexical stress patterns were also either consistent or inconsistent with the syntactic role of the critical words. Thus it is possible that the source of the conflict between implicit prosody and the final interpretation of the sentences was at the lexical level, due to the necessity of re-accessing the properly stressed phonological form from the lexicon. But longer processing times could also have been due to the necessity of re-assigning the location of sentence-level implicit pitch accents, or to some combination of these processes.

In an effort to find more direct evidence of the nature of the *sentence-level* auditory image present during silent reading, in the work presented here we conduct cross-modal priming experiments. In particular, we attempt to use the implicit prosody of a read text to prime an appropriately-intonated probe word. Readers were induced to generate an implicit pitch accent pattern with paired text sentences that contained a corrective contrast,e.g. *Jacqueline didn't pass the test. Belinda passed the test.* (cf. Bock & Mazzella, 1983, who used stimuli like *ARNOLD didn't fix the radio. DORIS fixed the radio.*). Corrective contrasts were in either subject or verb position. The participant's task was to read the sentence pair and then respond to the auditory probe by indicating whether it had been the initial word in the second sentence. The prosody of the probe words was manipulated to contain either a high-rising contrastive pitch accent (L+H*), or no pitch accent. When the corrective contrast was in subject position, readers should be induced to generate an implicit L+H* accent on the initial word in the second sentence, and show shorter response times for a spoken probe with a L+H* accent. But when the corrective contrast was in verb position, readers should be induced to generate a prosodic contour with an unaccented initial word in the second sentence, and show shorter response times to an unaccented probe. Bock & Mazzella (1983) showed faster sentence comprehension times when new information was accented and repeated information was not, as compared to the reversed pattern of accentuation. In Experiment 1, we present results from this cross-modal priming task. In Experiment 2, we group readers based on the prosody they used when reading aloud to see whether overt reading style can predict the pattern of response times in the cross-modal priming task. Note that we do not assume that read-aloud speech versions of sentences should have the same prosody as implicit prosody versions. Instead, we look at people's reading styles as an indicator of how their inner speech might sound.

We have employed a cross-modal priming paradigm in previous studies (Bergmann & Speer, 2007a; 2007b) to prime an appropriately intonated probe word by implicit prosody generated during silent reading. While the results from these studies were quite tentative, they have allowed us to pilot the methodology and develop a paradigm that may advance our understanding of the intonational composition of implicit prosody. In particular, we are using short sentences in contrasting pairs that will have been read silently and completely understood when the auditory probe is presented. In addition, the sentence location of the target word, whose implicit prosody is meant to prime an appropriately intonated auditory probe, is predetermined and consistent across trials. Our previous studies differed from this approach in that they combined self-paced reading with the presentation of auditory probes at varying, unpredictable sentence locations. We believe that our new approach has several advantages: It allows participants to silently read the complete sentence pairs in a natural reading situation and without the interruptions of the reading flow that are associated with button-presses in a self-paced reading task. This allows participants to assign an implicit prosodic pattern that is similar to what we would expect in natural reading tasks. In addition, a self-paced reading paradigm may have induced sentence-medial prosodic boundaries at button-press locations. A predetermined and consistent target word location may reduce the variability in response times because the participants' attention is drawn to the prime word, and the auditory probe is expected at a consistent time. In contrast, in previous studies response times may have been affected by varying levels of expectations about when an auditory probe might occur.

In addition to the above methodological difference, this study also focuses on a different aspect of sentence level prosody: Whereas our previous studies using cross-modal priming focused on edge tones, here we shift our attention to pitch accents. This has the advantage of reducing possible individual differences in implicit prosody while still investigating a sentence-level prosodic phenomenon. The contrastive accent vs. no accent manipulation in the short, simple sentence pairs that we used is a prosodic feature that is assigned at the discourse or sentence level and cannot be simply associated with a particular lexical item as lexical stress is. The pitch accent manipulation is also in some sense tonally simpler and more predictable than phrasal boundaries in English. Prosodic phrasal boundaries that mark syntactic constituency in English potentially have variable amounts of final lengthening and many tonal shapes (In the ToBI system, intermediate phrase accents may be H, !H or L, and boundary tones at intonation phrase breaks may also be either H or L, with the combinations creating many possible end contours). In contrast, repeated unaccented words have stressed syllables; words carrying corrective contrastive marking should have a high-rising contrastive pitch accent (L+H*). Some evidence for the frequency of L+H* used to mark contrast (although not corrective contrast) can be found in an analysis of spontaneous speech from a task in which speakers gave directions that included contrasting adjective-noun pairs such as '…hang a red ball. Now hang a green ball…'. Contrast-bearing adjectives (*green* in the example) bore a L+H* pitch accent on 53% of trials, while this accent appeared on only 3% of adjectives in comparable but non-contrastive contexts such as '…hang a red ball. Now hang a green drum…' (Ito & Speer, 2006). Thus, in the current study the number of possible appropriate prosodies that might allow for the probe word to be primed by implicit prosody is reduced.

Experiment 1

In this experiment, we present data from a cross-modal priming task. Participants silently read pairs of sentences that contained a corrective contrast either in subject or verb position. They then responded *yes* or *no* to an auditory probe, saying whether it had been the first word in the second sentence. The pitch accent pattern of the probe word either matched or did not match the implicit prosody presumably generated during reading. If this procedure taps into the auditory image that people generate during silent reading, we predict faster response times for matching auditory probes than for mismatching ones.

Methods

Participants

Sixty-eight adult native-English speakers participated in the study. Data from an additional two people were excluded due to too many missing values within one experimental condition. Participants were undergraduate students at The Ohio State University who received course credit for their participation.

Materials

The materials for each trial of the experiment consisted of a sentence pair and an auditory probe. Each sentence pair started with a negated statement, such as *Belinda didn't fail the test*, followed by a sentence that used lexical contrast to elaborate on which part of the first sentence was being negated, for example, *Belinda passed the test*. Here, it is the failing that is being negated and *failing* is contrasted with *passing*. Each auditory probe presented a word produced either with a rising corrective contrastive accent (L+H*) or no accent. The sentence pairs and auditory probes were combined in a 2 by 2 design to create the four experimental conditions shown in Table 1.

There were two sentence pair contrast conditions: Either the second sentence presented a correction of the subject (Subject Contrast) or the verb (Verb Contrast) of the first sentence. We will call the second sentence in each pair the target sentence (e.g. *Belinda passed the test*) and the first word of this sentence the target word (e.g. *Belinda*). When participants read the Subject Contrast sentence pairs silently, a felicitous implicit prosody would locate a corrective contrastive accent on the subject of the second sentence, i.e. on the stressed syllable of the target word (*BELINDA*). In contrast, when participants read the Verb Contrast sentence pairs silently, no implicit accent should be assigned to the target word (*BELINDA*). In this case, the verb of the second sentence (*PASSED*) should receive an implicit corrective contrastive accent. The subject noun in Verb Contrast sentences should receive no accent for two reasons: first, it was mentioned in the immediately preceding sentence, and speakers generally refrain from accenting repeated, "old" information (cf. Bolinger, 1961, 1986; Chafe, 1974, 1976; Terken & Hirschberg, 1994) and second, being at the beginning of the sentence and adjacent to a contrastively-accented word makes it a likely target for de-accenting (Hirschberg, 2008).

There were also two auditory probe conditions: Auditory probes were recordings of the target word (*Belinda*) produced either with a prosodic contour that matched or that did not match the implicit prosody appropriate for the target word. For Subject Contrast sentence pairs, an auditory probe with a L+H* accent was considered to be a Match, whereas an auditory probe with no accent was considered to be a Mismatch. The reverse was the case for Verb Contrast sentence pairs: Here, an auditory probe with no accent was considered to be a Match, and an auditory probe with a L+H* accent was considered to be a Mismatch.

Table 1: Experimental conditions with examples: CAPS indicate an auditory probe with a L+H* accent and SMALL CAPS indicate no accent.

		Visual Sentence Pair Type	
		Subject Contrast	**Verb Contrast**
Auditory Probe	**Match**	Sentence pair: *Jacquelyn didn't pass the test.* *Belinda passed the test.* Auditory probe: BELINDA	Sentence pair: *Belinda didn't fail the test.* *Belinda passed the test.* Auditory probe: BELINDA
	Mismatch	Sentence pair: *Jacquelyn didn't pass the test.* *Belinda passed the test.* Auditory probe: BELINDA	Sentence pair: *Belinda didn't fail the test.* *Belinda passed the test.* Auditory probe: BELINDA

We created four experimental lists. Each list included 32 experimental items, eight in each of the four conditions, and 32 filler items. Experimental items were rotated across lists in a Latin square. Filler trials differed from experimental trials in two ways: Filler sentence pairs contrasted either in the sentences' objects or in prepositional phrases and filler auditory probes were recordings of words other than the target word (either another word in the sentence pair or a word that was phonetically similar to the target word).

The auditory probes were created as follows: A female native-English speaker with phonetic training recorded the target sentence of each experimental sentence pair with two different prosodies, that is, either with a L+H* accent on the subject noun and deaccentuation on subsequent elements (e.g. *BELINDA passed the test*, where CAPS indicate a L+H*) or with no accent on the subject noun, a L+H* accent on the verb, and deaccentuation on subsequent elements (e.g. *BELINDA PASSED the test*, where CAPS indicate a L+H* and SMALL CAPS indicate no accent). The first word of each sentence was then extracted, so that there were two auditory probes for each target sentence: one produced with a L+H* accent (BELINDA, see Figure 1) and one produced with no accent (BELINDA, see Figure 2). All experimental target sentences started with a three-syllable proper name with main stress on the second syllable (*Belinda*) and were thus long enough to produce clear prosodic patterns that remained even after the target word was excised from the rest of the sentence. Auditory probes with no accent differed reliably from auditory probes with a L+H* accent both in duration (average: 401 ms (sd = 72) for no accent vs. 483 ms (sd = 67) for L+H*, paired t-test: $t = -12.4943$, p <

.001) and in pitch maxima of the stressed syllable (average: 182 Hz (sd = 15) for no accent vs. 289 Hz (sd = 11) for L+H*, paired t-test: *t* = -32.1541, p < .001). In addition, visual inspection of the fundamental frequency contours showed a high peak pattern toward the end of the vowel of the medial stressed syllable for L+H* accented probes, and a relatively flat, downward-sloping contour across all three syllables for probes with no accent. The same speaker also recorded the target sentence for each filler sentence pair, either as it appeared in the experiment or with a phonetically similar proper noun in subject position. From these productions, words that were not target words (for example the verb, object, noun of a prepositional phrase, or a word that phonetically resembles the target word) were extracted.

Figure 1: Spectrogram and fundamental frequency contour for the auditory probe *BELINDA*, produced with a L+H* accent on the syllable with primary stress. The x-axis shows time in milliseconds, the left-hand and right-hand y-axes show pitch measured in Hertz.

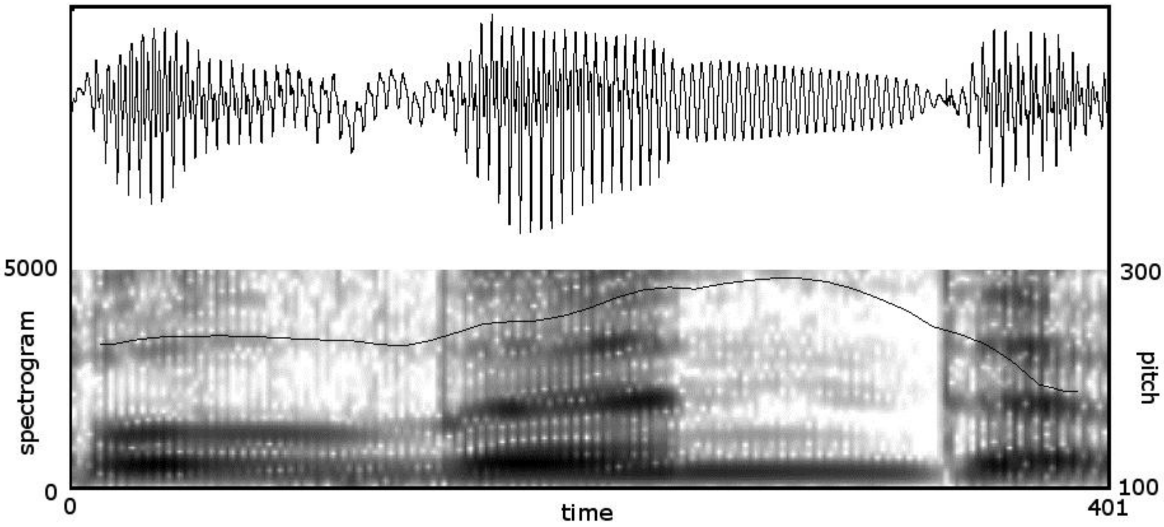
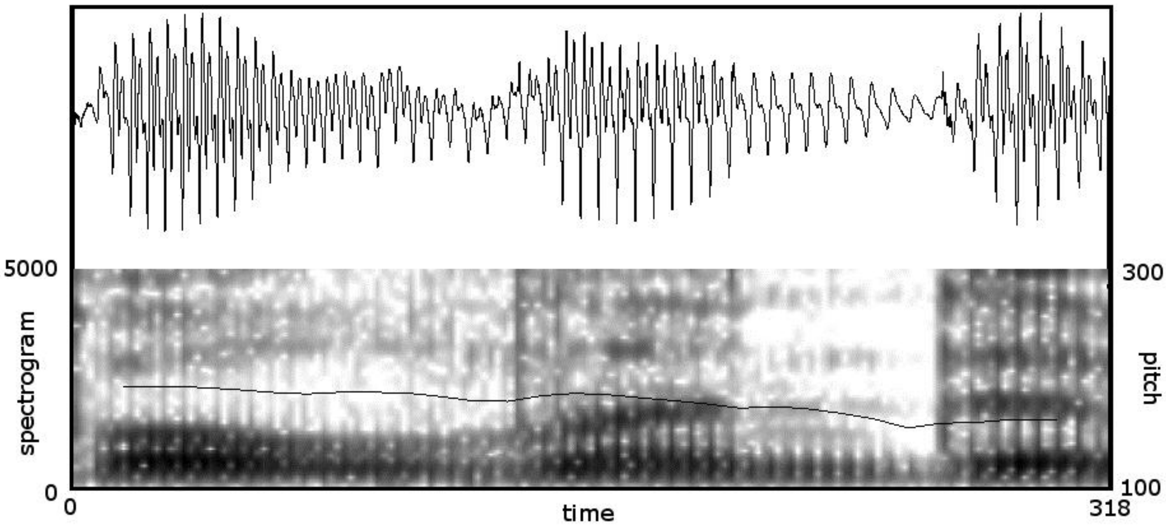


Figure 2: Spectrogram and fundamental frequency contour for the auditory probe BELINDA, produced with no accent. The x-axis shows time in milliseconds, the left-hand and right-hand y-axes show pitch measured in Hertz.



Procedure

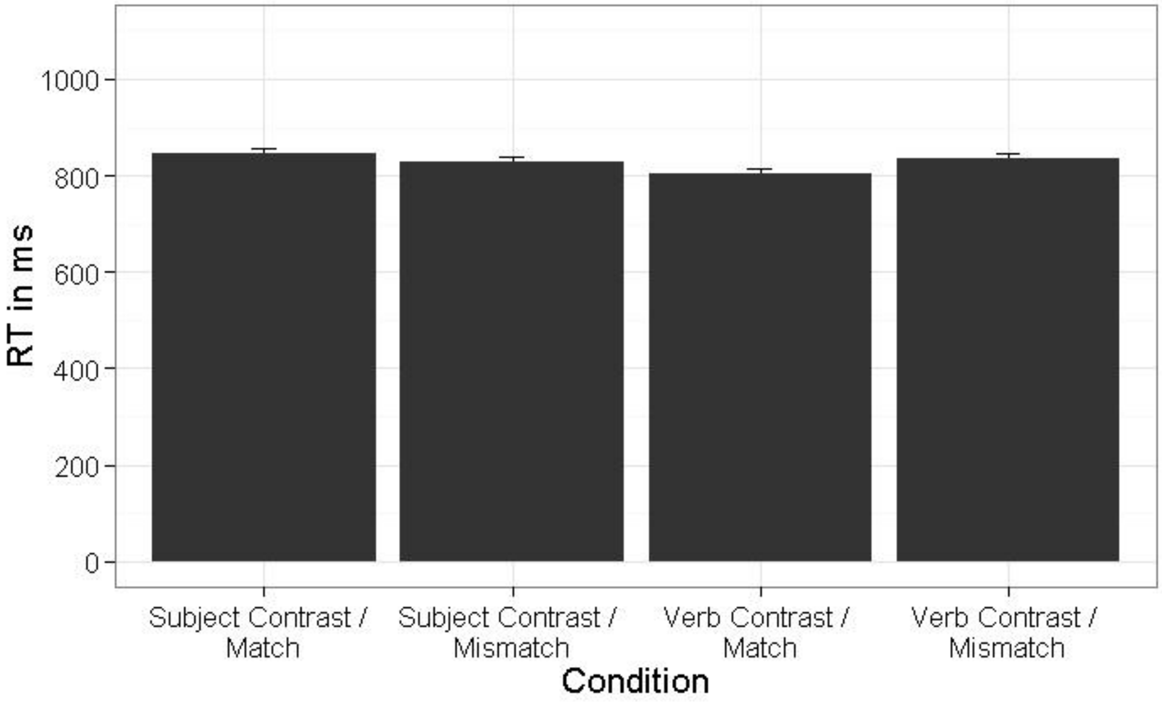
Before the start of the experiment, participants received written and oral instructions for the experimental task, performed a practice session, and had the opportunity to ask questions. Then the experiment began. Participants were seated in front of a computer screen wearing noise-cancelling headphones (used to reduce distracting sound from other participants who participated in the same room at the same time). On each trial of the experiment, participants first silently read a sentence pair. Each sentence was displayed on a separate line to highlight the contrast between the first and second sentence. After reading the pair they pressed a button, upon which the text disappeared and an auditory probe was presented. To ensure that participants were reading the sentences at normal reading speed, the auditory probe was automatically presented after three seconds if participants had not pushed the button by then. Participants were instructed to respond to the auditory probes by deciding as fast as they could whether or not the word they heard was the first word of the second sentence, i.e. the target word. To help them with this decision, the target word remained on the screen during the presentation of the auditory probe. After each decision, participants received feedback as to whether or not they had responded correctly. If participants had not responded after two seconds, they received a warning that their response was too slow. Following each response to an auditory probe, participants answered a comprehension question about the sentence pair. There was no time limit for responding to the comprehension question, and participants again received feedback as to whether or not they had responded correctly. After the experiment participations completed a language background questionnaire.

Results

We tested whether participants responded faster to matching than to mismatching auditory probes. Such a result would suggest that participants are generating the expected implicit prosodic contours as they are reading silently and that our task taps into implicit prosody. Incorrect responses (i.e. responding ‘no’ during a target trial, 2.2% of the data points) and failures to respond within the two second limit (3% of the data points) were excluded from the data set. Incorrect responses to the auditory probes occurred in fewer than 4% of all target trials; accuracy for Verb Contrasts (proportion correct .98) showed a small but statistically significant advantage compared to that for Subject Contrasts (proportion correct .96) ($t=2.69$), but match and mismatch conditions did not differ ($t < 1$).

Response times under 200 ms and those that were two standard deviations above or below the mean for a given item and participant were also excluded from the data set. Altogether, 82 responses (8.3%) were excluded from the data set. The response times across the four conditions are shown in Figure 3. We ran mixed-effects models with response time as the dependent variable and subjects and items as simultaneous random effects. We added probe type (Match vs. Mismatch), sentence contrast (Subject Contrast vs. Verb Contrast), and the probe type x sentence contrast interaction as fixed effects to the initial model. Probe type is the fixed effect of interest for our research question.

Figure 3: Response times for auditory probes in the four critical conditions.



Redundant fixed effects were removed from the initial model until the model was minimally optimized. Random slopes were added if they improved model fit (Barr, 2013; Barr, Levy, Scheepers & Tily, 2013). The final model included sentence contrast and the probe type x sentence contrast interaction as fixed effects. Response times were shorter for probes following Verb Contrast sentence pairs than following Subject Contrast sentence pairs (*estimate* = 9.345, *t* = 2.168, *p*< .05). To explore the reliable probe type x sentence contrast interaction (*estimate* = -12.998, *t* = -3.438, *p*< .001), we fit separate mixed-effects models for Subject Contrast sentence pairs and Verb Contrast sentence pairs. Each model had response time as the dependent variable, subjects and items as simultaneous random effects, and probe type (Match vs. Mismatch) as fixed effect. Random slopes were added if they improved model fit. The results from these models reveal that participants responded marginally faster to mismatching probes than to matching ones for the Subject Contrast sentence pairs (*estimate* = -8.957, *t* = -1.663, *p*= .097), but that they responded reliably faster to matching probes compared to mismatching probes for the Verb Contrast sentence pairs (*estimate* = 16.949, *t* = 3.185, *p*< .01).

In sum, participants responded faster to simpler no accent probes compared to more complex L+H* probes. In contrast to our expectations, probe type (Match vs. Mismatch) had no effect. To understand why this was the case, we had a closer look at individual participants' data. Individual participants' response patterns were grouped into four categories, based on how they responded to the different kinds of auditory probes: Those who showed shorter response times for matches than for mismatches (shorter matches), those who showed shorter response times for mismatches than for matches (shorter mismatches), those who showed shorter response times for no accent probes compared to L+H* probes (shorter no accent), and those who showed shorter response times for L+H* probes compared to no accent probes (shorter L+H*). 41% of participants (28 out of 68) fell into the shorter no accent category, followed by shorter matches (28%, 19 out of 68), shorter mismatches (21%, 13 out of 68), and shorter L+H* (12%, 8 out of 68). Thus, the two most common response patterns were shorter no accent and shorter matches. That is, we found that an unexpectedly large percentage of participants showed a processing advantage for the no accent probe conditions. This is also reflected in the reliable probe type x sentence contrast interaction. The expected response pattern (shorter matches) was only the second most frequent pattern. One possibility for why shorter matches was not the most frequent pattern is that there are individual differences in participants' implicit prosody, such that only a portion of the participants responded to the auditory probes the way we expected. We explore this possibility in Experiment 2.

Discussion

In Experiment 1 we tested whether we could use a cross-modal priming paradigm to tap into the implicit prosodic contours that people generate when they read silently. We assumed that participants would generate an implicit L+H* accent on a correctively-contrasted subject and no implicit accent on a repeated subject. If the prosodic contour of the auditory probe word was more similar to the implicit prosody image generated for the target word, as in the Match conditions, we predicted that participants would show a processing advantage when deciding that the probe word was the same as the one at the beginning of the read target sentence. In contrast to our predictions, however, we found no effect of probe type, such that matching probes did not elicit shorter response times than mismatching probes. What we did find were reliable effects of sentence contrast, with faster responses to probes following verb contrasts than to probes following subject contrasts, and of the probe type x sentence contrast interaction, with faster responses to no accent probes compared to L+H* probes. Participants may have responded more quickly to probes following verb contrasts than to probes following subject contrasts because it is easier to confirm that the auditory probe word is the same as the target word if the target word is a repetition from the first sentence than if the target word contrasts with the first word of the first sentence. In other words, it may be easier to confirm that the auditory probe was a recording of the proper name shown twice in the written sentences than to confirm that the auditory probe was a recording of one of the two different proper names shown in the written sentences, in particular, of the proper name shown in the second written sentence. Participants may have responded more quickly to probes with no accent compared to probes with a L+H* accent because probes with no accent may require less involved processing than probes with a more complex bi-tonal pitch accent (L+H*). Similar effects have been found for edge tones: Participants in Bergmann & Speer (2007b) showed shorter response times to probes without an edge tone and probes with only a phrase accent than to probes with a boundary tone. Thus, participants were faster to respond when the probe was prosodically less complex. Alternatively, participants may have responded more quickly to probes with no accent compared to probes with a L+H* accent because no accent probes were reliably shorter in duration than L+H*. Such shorter duration may have allowed for faster recognition and thus faster response times.

An exploratory post-hoc analysis of individual participants' response patterns revealed that a large percentage of participants responded faster to no accent probes compared to L+H* probes. What we expected, however, was that a large percentage of participants would respond faster to matching probes compared to mismatching ones. How can we explain then that so many participants showed a different pattern of responses? As mentioned above, one possibility is that participants simply responded faster to shorter compared to longer probes or simple compared to more complex pitch accents. However, another possibility is that there are individual differences in participants' reading style, and thus in the implicit prosody they generate. In particular, it is possible that a sizeable portion of our participants read (both silently and aloud) in a rather monotone way and produce few L+H* accents during reading. If this is the case, these participants may frequently have generated implicit accent patterns unlike those instantiated in our L+H* probes. Thus, for these participants no accent probes may have always been more similar to their internal read speech, regardless of sentence contrast. It would then not be surprising to find a reliable probe type x sentence contrast interaction with faster responses for no accent compared to L+H* probes rather than a reliable effect of probe type (Match vs. Mismatch) in this study. In Experiment 2 we explore the idea that the unexpected results of this experiment were due to such individual differences.

Experiment 2

In this experiment participants read aloud a brief text after performing the same task as in Experiment 1. We added a reading aloud task since the results from Experiment 1 led us to hypothesize that participants differed in the implicit prosody they generated during silent reading and that these differences may have affected response times. The reading aloud task allows us to group people based on measurable prosodic phenomena. In particular, we can group people based on whether or not they produce L+H* accents when reading aloud. If a related mechanism is involved for generating implicit prosody during silent reading and overt prosody during reading aloud, participants who produce L+H* accents when reading aloud should be more likely to generate implicit L+H* accents during a silent reading task. And if our procedure taps into the auditory image generated during silent reading, it is these participants who should respond more quickly to matching auditory probes than to mismatching ones.

Methods

Participants

Participants were visitors to the "language pod" exhibition laboratory at the Center for Science and Industry (COSI), in Columbus, Ohio. They ranged in age from 15 to 60 years, and had educational backgrounds ranging from incomplete high school diplomas to advanced degrees. All were native speakers of Midwestern American English, had normal hearing and normal or corrected-to-normal vision. Data from twenty-seven participants were included in the study. All included participants reported hearing an auditory image of read words during silent reading. Data from an additional eight people were excluded due to too many missing values within a single experimental condition.

Materials

The materials were the same as those of Experiment 1. In addition, we used the following text passage, adapted from a 2012 New York Times article, for the reading aloud task:

> *The more automobile design has changed, the more it has remained the same. A century after the Model T Ford debuted, the vast majority of the cars on the road still feature steel bodies, chassis suspended on four wheels and four-stroke internal combustion engines. Not that would-be revolutionaries haven't tried to "improve" the automobile with a host of innovations: Bodies made of carbon-fiber. Bodies fitted with wings. Bodies that float on water. Three-wheelers. Six-wheelers. Steam engines. Jet engines. For a while, there was even talk of nuclear power. But designers don't control how cars are built, manufacturers control how cars are built.*

Procedure

The procedure was the same as in Experiment 1, with one addition: After performing the experiment participants were recorded reading aloud the above text passage. No connection between the two reading tasks was mentioned to participants.

Annotation

A coder with extensive training ToBI-annotated (Tones and Break Indices; Beckman, Hirschberg, & Shattuck-Hufnagel, 2005) the following relevant sentences from the text passage: *Not that would-be revolutionaries haven't tried to "improve" the automobile with a host of innovations* and *But designers don't control how cars are built, manufacturers control how cars are built*. These sentences were chosen because they contained words that participants most frequently pronounced with emphasis or contrast. In particular, *improve* may have received a L+H* accent because it was visually highlighted with quotes, and *manufacturers* may have received a L+H* accent because it is set in contrast to *designers*. Based on these annotations, participants were divided into three groups: L+H* users, H* users, and monotone readers. 12 participants fell into the L+H* user group. Nine of these participants produced a L+H* on *improve*. Eight of these participants produced a L+!H* and three a L+H* on *manufacturers*. These participants also deaccented repeated material. For example, all of the twelve L+H* users deaccented the repeated verb *control* that followed *manufacturers*. To illustrate this, Figure 4 shows the spectrogram and pitch track of *manufacturers control* produced by an example L+H* user. Twelve further participants fell into the H* user group. Seven of these participants produced an H* and one additional participant a H+!H* on *improve*. Nine of these participants produced an H* on *manufacturers*. These participants did not deaccent the repeated verb *control*. Instead, most of them produced *control* with a H* or !H* accent. This is illustrated in Figure 5, which shows the spectrogram and pitch track of *manufacturers control* produced by a H* user. Three participants were considered monotone: They produced no accents or a L* accent on *improve* and *manufacturers*.

Figure 4: Spectrogram and pitch track for *manufacturers control*, produced by a L+H* user. *Manufacturers* receives a L+H* accent on the syllable with primary stress and *control* is deaccented. The x-axis shows time in milliseconds, the left-hand and right-hand y-axes show formants and fundamental frequency, respectively, both measured in Hertz.

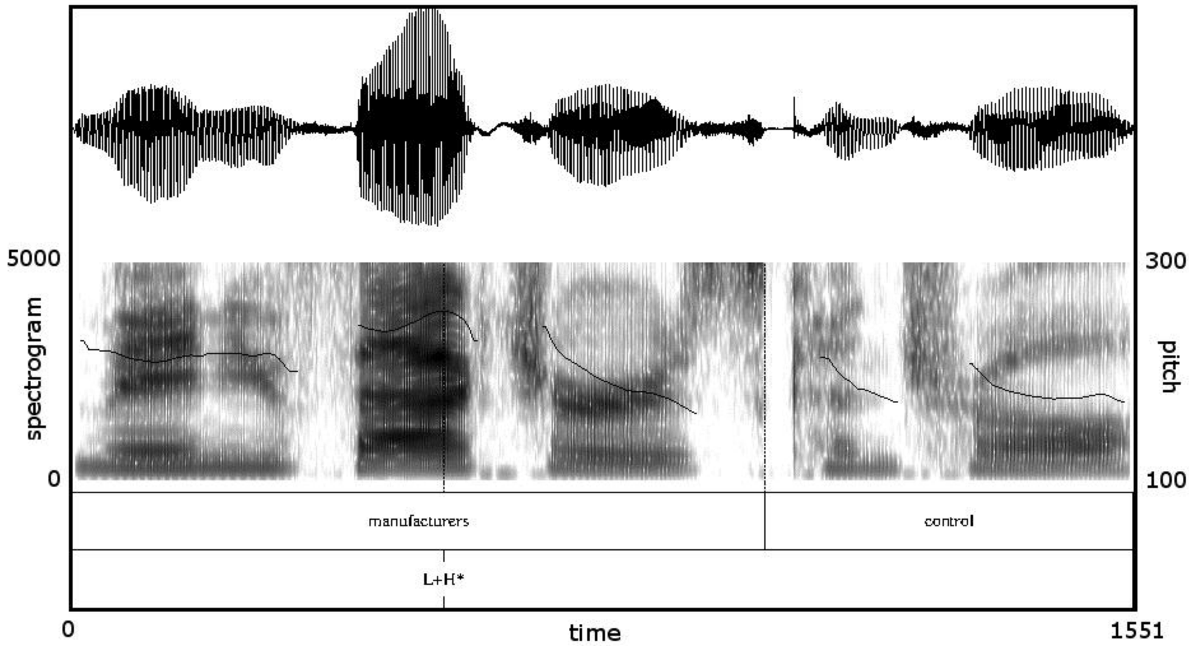
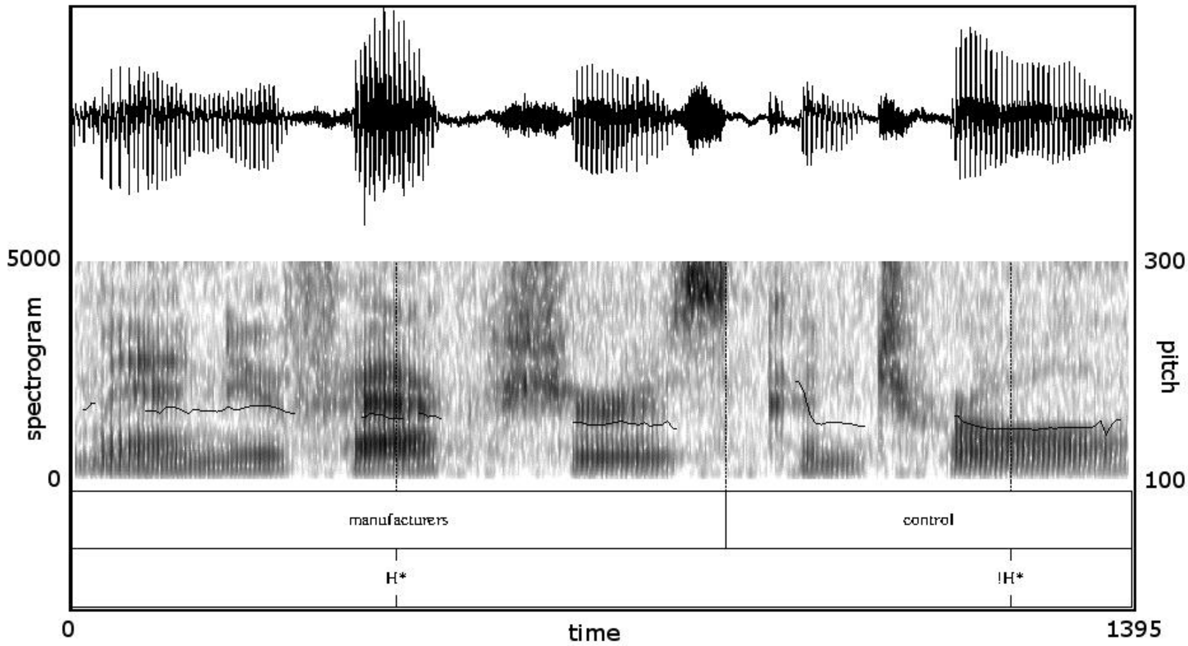


Figure 5: Spectrogram and pitch track for *manufacturers control*, produced by a H* user. *Manufacturers* receives a H* accent and *control* a !H* accent on the syllable with primary stress. The x-axis shows time in milliseconds, the left-hand and right-hand y-axes show formants and fundamental frequency, respectively, both measured in Hertz.



Results

We tested whether the L+H* users identified above responded faster to matching than to mismatching auditory probes. Notice that this group L+H*-accented a correctively contrasted noun and did not accent the repeated verb in the read-aloud passage. We thus expect that the auditory probes that we call Matches are indeed matches for this group of people, both for the Subject Contrast sentences, where the matching auditory probe has a L+H* accent, and the

Verb Contrast sentences, where the matching auditory probe has no accent. To test whether the response pattern that we expect for L+H* users is particular to this group, we also analyzed the data from the H* users. We therefore added the factor participant group (L+H* users vs. H* users) to the analyses reported below. We expect that the H* users not show a processing advantage for matching compared to mismatching probes. An analysis of data from the monotone readers is not possible due to the small sample size.

Again, failures to respond, incorrect responses, response times under 200 ms, and response times that were two standard deviations above or below the norm for a given item and participant were excluded from the data set. Incorrect responses to the auditory probes occurred in fewer than 10% of target trials for both the L+H* and H* subject groups in all four conditions, and did not differ significantly among them (all *t*s < 1.2).

The response times across the four conditions are shown in Figure 6 for the L+H* users and in Figure 7 for the H* users. We again ran mixed-effects models with response time as the dependent variable and subjects and items as simultaneous random effects. We added probe type (Match vs. Mismatch), sentence contrast (Subject Contrast vs. Verb Contrast), participant group (L+H* users vs. H* users), and all interactions as fixed effects to the initial model. Here, the probe type x participant group interaction is the fixed effect of interest for our research question.

Figure 6: Response times for auditory probes for the twelve L+H* users in the four critical conditions.

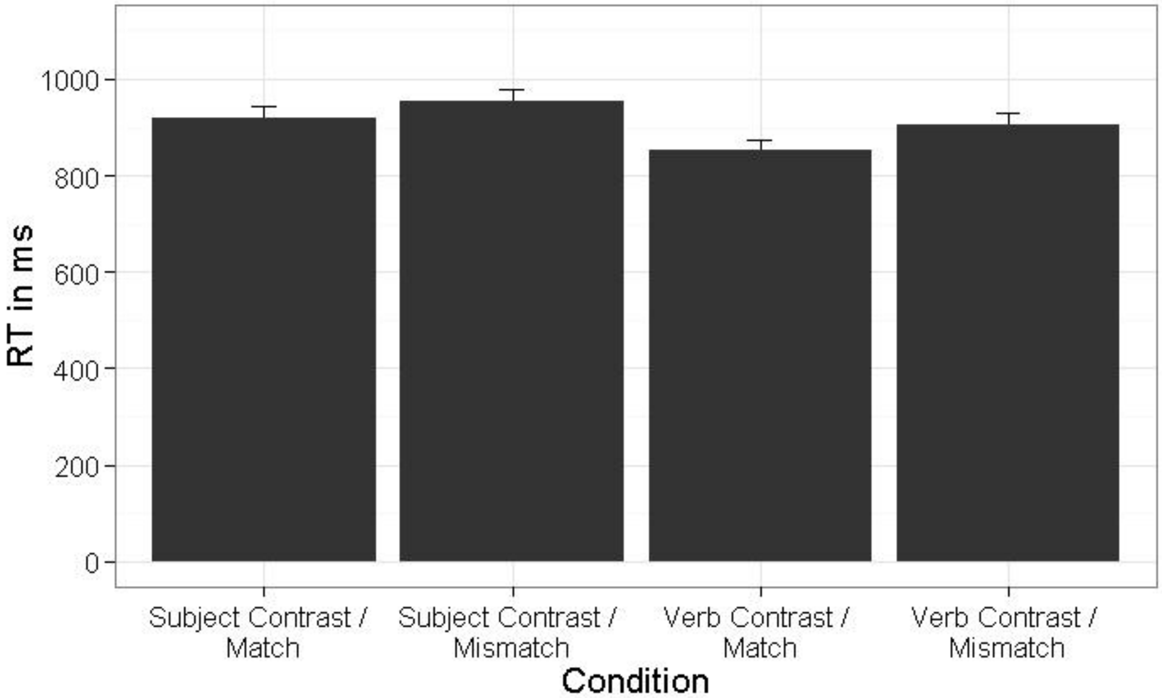
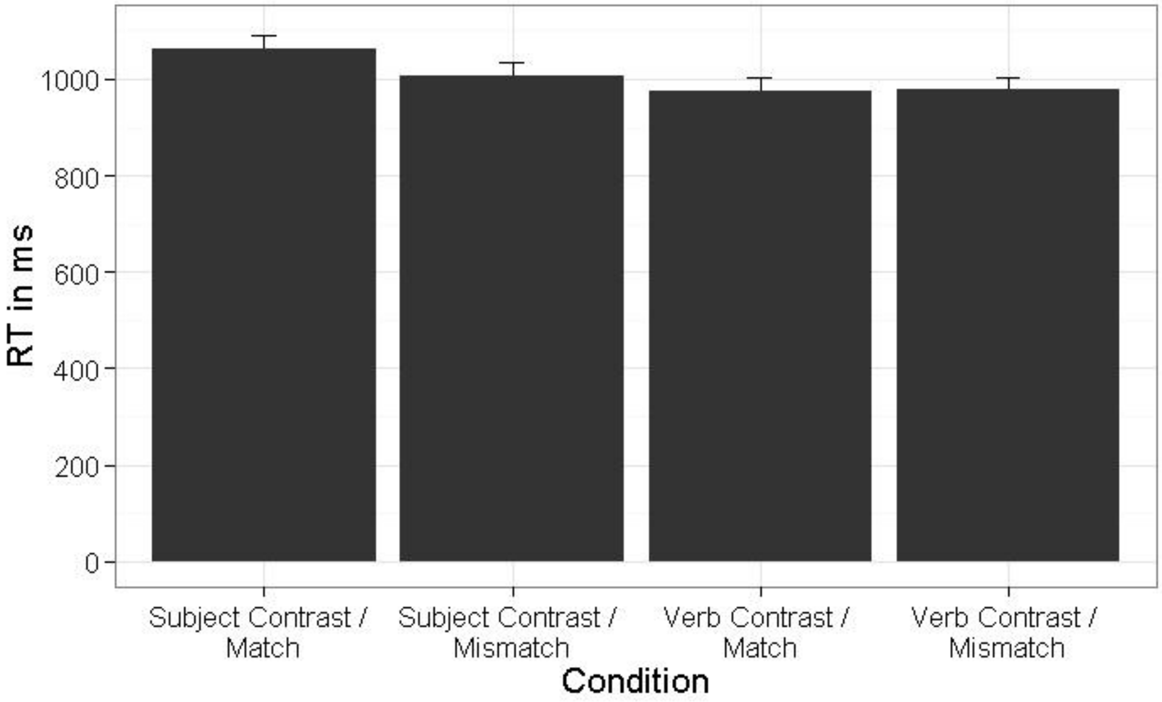


Figure 7: Response times for auditory probes for twelve H* users in the four critical conditions.



Redundant fixed effects were removed from the initial model until the model was minimally optimized. Random slopes were added if they improved model fit. The final model included sentence contrast (Subject Contrast vs. Verb Contrast) and the probe type x participant group interaction as fixed effects. As in Experiment 1, participants were faster to respond to probes following Verb Contrast sentence pairs than following Subject Contrast sentence pairs (*estimate* = -27.877, *t* = -4.148, *p*< .001). To explore the reliable probe type x participant group interaction (*estimate* = 17.677, *t* = 2.563, *p*< .05), we fit separate mixed-effects models for L+H* users and H* users. Each model had response time as the dependent variable, subjects and items as simultaneous random effects, and probe type (Match vs. Mismatch) as fixed effect. Random slopes were added if they improved model fit (Barr, 2013; Barr et al, 2013). The results from these models reveal that, importantly, L+H* users responded reliably faster to matching probes than to mismatching probes (*estimate* = 25.428, *t* = 2.954, *p*< .01). In contrast, H* users responded equally fast to matching and mismatching probes (*estimate* = -8.551, *t* = -0.808, *p*= .42).

Thus, those participants who produced L+H* accents on correctively contrasted material and no accent on repeated material when reading aloud showed the expected response pattern: A processing advantage for L+H* probes following Subject Contrast text sentence pairs, and a processing advantage for no accent probes following Verb Contrast text sentence pairs. This suggests that these readers may have generated L+H* accents and deaccentuation during silent reading, leading to shorter response times to probes that matched the generated implicit prosody than to probes that did not match. The H* users differed from the L+H* users in their response to the auditory probes: Whereas the L+H* user group's response times were shorter for Matches than for Mismatches, the H* user group showed no effect of probe type. Thus, as expected, only the L+H* user group showed faster responses for Matches than for Mismatches.

Discussion

In this experiment, we looked for evidence of the implicit prosody readers were induced to generate for sentences with corrective contrasts by testing whether differences in the prosody participants used when reading aloud co-occurred with differences in their pattern of responding in our cross-modal priming task. We found that participants who produced L+H* accents and deaccentuation when reading aloud responded faster to matching auditory probes than to mismatching auditory probes. Since a reliable effect of probe type (Match vs. Mismatch) relies on participants generating implicit L+H* accents on contrasting items and no accent on repeated items in the experimental task, the data from this experiment provide some evidence that participants who produced measureable L+H* accents and deaccentuation in the read-aloud task also generated implicit L+H* accents on contrast items and no accent on repeated items during the silent reading task. In contrast, participants who produced H* accents on contrast items during the read-aloud task showed no reliable differences in responding to the probe type (Match vs. Mismatch) manipulation. The absence of a reliable three-way interaction also suggests that, unlike the participant group in Experiment 1, the H* users did not respond reliably faster to no accent probes compared to L+H* probes (even though, as shown in Figure 7, their responses were numerically faster for no accent than for L+H* probes). Thus, the H* user group seemed to show no sensitivity to the auditory probe manipulation at all. One possible explanation for this result is that neither the L+H* nor the no accent auditory probes matched the implicit prosody that the H* users generated during the silent reading task. Instead, H* probes might have been needed in order to constitute "matches" for this group of participants. If so, since the experiment did not contain any target auditory probes with a H* accent, neither probe was primed by the implicit prosody of the silently read text for this group.

We draw two conclusions from this experiment. First, the results suggest that there is some similarity between the prosody produced while reading aloud and that produced while reading silently: It was only the participants who produced measurable L+H* accents on the corrective contrast word and deaccentuation on repeated words when reading aloud that showed a priming response for matching auditory probes in the cross-modal task. Participants who produced H* and !H* accents when reading corrective contrast material aloud showed no sensitivity to the differences between auditory probes. Thus, the L+H* user and the H* user groups behaved differently from each other in both tasks. In addition, both groups behaved consistently across the two tasks, i.e. their read-aloud prosody predicted their performance in cross-modal priming in the silent reading task. This suggests that participants' propensity to use L+H* accents for corrective contrasts and deaccentuation for repeated items was similar for reading aloud and for reading silently. This is a potentially important result. Since we can neither hear nor phonetically measure implicit prosody, we can’t know how readers' implicit prosody "sounds," nor can we know if the auditory image generated during reading is the same for every reader. The results from this experiment do give us some insight into how we might assess the "sound" of readers’ implicit prosody. In particular, the results suggest that there may not be a one-to-one match between overt and implicit prosody, but that more general characteristics of speech read aloud are also found in silently read speech. One of these characteristics is one’s propensity to produce salient L+H* accents for corrective contrast words and no accent for repeated words. The second conclusion that we draw is that the current cross-modal priming paradigm in combination with a read-aloud diagnostic task may be well suited to study the sound of implicit prosody.

General Discussion

This paper presents findings from two visual-to-auditory cross-modal priming experiments designed to investigate whether the implicit prosody generated during silent reading can prime an appropriately intonated auditory probe. Our results indicate a qualified 'yes' to this question: We found evidence that, for speakers who prosodically marked corrective contrastives and orthographically marked words with a salient rising pitch accent (L+H*) followed by a deaccented region in oral reading, an appropriately intonated probe word could be primed by a corrective contrast in preceding silently read text.

While our first experiment showed no effect of whether the prosody of the probe was consistent with the location of the corrective contrast in the visual sentence pair, it did show effects of the corrective contrast location, with shorter participant responses for verb contrasts, which involved repetition of the sentence-initial target word. In addition, Experiment 1 showed longer response times for auditory probes with L+H* accents than for those with no accent. This overall pattern of responding was repeated in Experiment 2 for the H* subject group, who showed significantly longer response times for Verb Contrast trials, and numerically longer times for L+H* probes within contrast types. The L+H* participant group in Experiment 2 provides initial evidence that a well-known prosodic pattern, that for corrective contrast, can be evoked by a sentence pair presented in text and used to prime a subsequent auditory pitch accent pattern. These results are consistent with previous findings that silently read statements and questions can speed processing of subsequently presented auditory words with falling or rising intonation, respectively (Abramson, 2007). However, we needed to resort to an overt reading task to provide information about the implicit prosody we might expect to induce from individual readers. This suggests that individual differences may obscure results from the priming paradigm, especially when there are multiple potential grammatical prosodies available for a particular read text.

Such individual differences may be the reason why studies involving word-stress manipulations have so far yielded more consistent results than studies involving sentence-level prosodic phrasing regarding both the existence of an implicit prosodic contour generated during silent reading and information about what this implicit prosody may sound like. While there are some words whose stress pattern is affected by the sentential context (e.g. *He’s sixTEEN* vs. *SIXteen candles*) or is subject to individual differences (e.g. the noun *address* can be pronounced with stress either on the first or the second syllable), English stress is a word-level phenomenon, i.e. stress is a property of individual lexical items. As such, there is little room for individual differences when it comes to the stress patterns of particular lexical items. In contrast, there are numerous felicitous pitch accent patterns for the sentence contrasts that readers experienced in the silent reading task in this study. Even though Bock & Mazzella (1983) found a processing advantage for the pattern that we hypothesized to be readers’ most common implicit prosodic contour (with a L+H* accent on corrective contrasts and no accent on repeated material), the prosodies produced in our overt reading task suggested that there may be two approximately equally common patterns that readers generated during silent reading, along with a far less likely 'monotone' pattern. That is, we found both the hypothesized contour and one with a H* on corrective contrast words and a !H* on repeated material. Interestingly, the rate of L+H* no accent vs. H* !H* production for corrective contrast across participants in our production study seems comparable to that found previously for the spontaneous production of contrastive adjective sequences (Ito & Speer, 2006). Any study investigating sentence-level implicit prosody will likely have to deal with such individual differences. Thus, it may not be possible to study sentence-level implicit prosody without recourse to participants’ overt prosody. The advantage of the visual-to-auditory cross-modal priming paradigm presented here is that a very brief and simple diagnostic allows grouping participants based on certain prosodic phenomena, so that implicit prosody can then be studied without comparison of implicit and overt prosody on a sentence-by-sentence basis. Indeed, our results suggest that such comparisons, as have been done in previous work (Bergmann & Ito, 2009; Foltz et al., 2011; Hwang & Schafer, 2009), may not be useful: rather, measurable general characteristics and tendencies found in overt prosody from reading aloud can be used to predict implicit prosodic behavior and group participants based on these predictions to then see if the cross-modal priming paradigm may be used to confirm the predictions. However, since the current experiments differ from the previous work in that they test cross-sentential pitch accent patterns rather than implicit prosodic phrasing, this conclusion may be premature.

References Cited

Abramson, M. (2007). The written voice: Implicit memory effects of voice characteristics following silent reading and auditory presentation. *Perceptual and Motor Skills*: 105, pp. 1171-1186.

Ashby, J., & Clifton, C. Jr., (2005). The prosodic property of lexical stress affects eye movements during silent reading. *Cognition, 96,* B89–B100.

Barr, D. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in Psychology,* 4, 1-2.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language,* 68, 255–278.

Beckman, M.E., Hirschberg, J., & Shattuck-Hufnagel, S. (2005). The Original ToBI System and the Evolution of the ToBI Framework. In S.-A. Jun (Ed.), *Prosodic Typology: The Phonology of Intonation and Phrasing*, pp. 9-54. Oxford: Oxford University Press.

Bergmann, A., & Ito, K. (2007). Attachment of Ambiguous RCs: A Production Study. *Talk given at the 13th Annual Conference on Architectures and Mechanisms for Language Processing (AMLaP), Turku, Finland.*

Bergmann, A., & Ito, K. (2009). Production and Comprehension of Interpretation-driven versus Input-driven Speech. *Poster presented at 22nd Annual CUNY Conference on Human Sentence Processing, Davis, CA.*

Bergmann, A., Ito, K., & Maday, K. (2008). Order Effects in Production and Comprehension of Prosodic Boundaries. *Talk given at TIE3: The Third TIE Conference on Tone and Intonation, Lisbon, Portugal.*

Bergmann, A., & Speer, S. R. (2007a). On Priming by Implicit Prosody. *Poster presented at the 20th Annual CUNY Conference on Human Sentence Processing, La Jolla, CA.*

Bergmann, A., & Speer, S. R. (2007b). More on Priming by Implicit Prosody. *Poster presented at the 13th Annual Conference on Architectures and Mechanisms for Language Processing (AMLaP), Turku, Finland.*

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

Bolinger, D. (1961). Contrastive accent and contrastive stress. *Language, 37,* 83-96.

Bolinger, D. (1986). *Intonation and its parts.* Edward Arnold, London.

Breen, M., & Clifton, C.E. Jr. (2011). Stress matters: Effects of anticipated lexical stress on silent reading. *Journal of Memory and Language, 64*, 153-170.

Breen, M., & Clifton, C. Jr., (2013). Stress matters revisited: A display change experiment. *Quarterly Journal of Experimental Psychology, 66*, 10, 1896-1909.

Chafe, W. (1974). Language and consciousness. *Language, 50*, 111-133.

Chafe, W. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and points of view. In C. Li (Ed.), *Subject and topic.* pp. 25-56. New York: Academic Press.

Fernández, E. M. (2003). *Bilingual sentence processing: Relative clause attachment in English and Spanish.* Amsterdam, The Netherlands: John Benjamins.

Fernández, E. M., & Bradley, D. (1999). Length effects in the attachment of relative clauses in English. *Poster presented at the 12th annual CUNY conference on human sentence processing, New York.*

Fodor, J. D. (1998). Learning to parse. *Journal of Psycholinguistic Research, 27*, 2, 285-319.

Fodor, J. D. (2002). Prosodic disambiguation in silent reading. *NELS, 32*, 113-32.

Foltz, A., Maday, K., & Ito, K. (2011). Order Effects in Production and Comprehension of Prosodic Boundaries. In S. Frota, G. Elordieta, & P. Prieto (Eds.), *Prosodic categories: production, perception and comprehension*, pp. 39-68. Dordrecht: Springer.

Hirose, Y. (1999). Resolving reanalysis ambiguity in Japanese relative clauses. Unpublished doctoral dissertation, CUNY Graduate Center, New York.

Hirschberg, J. (2008). Pragmatics and intonation. In L.R. Horn & G. Ward (eds.) *The Handbook of Pragmatics*. Blackwell Publishing Ltd, Oxford, UK.

Howell, P., & Kadi-Hanifi, K. (1991). Comparison of prosodic properties between read and spontaneous speech. *Speech Communication, 10*, 163-169.

Hwang, H., & Schafer, A. J. (2009). Constituent length affects prosody and processing for a dative NP ambiguity in Korean. *Journal of Psycholinguistic Research, 38*, 151-175.

Ito & Speer, S.R. (2006). Using interactive tasks to elicit natural dialogue. In P. Augurzky & D. Lenertova (Eds.), *Methods in Empirical Prosody Research*, Mouton de Gruyter.

Ito, K., & Speer, S. R. (2008). Anticipatory effects of intonation: Eye movements during instructed visual search. *Journal of Memory and Language, 58*, 541-573.

Jun, S.-A. (2010). The implicit prosody hypothesis and overt prosody in English. *Language and Cognitive Processes, 25*, 7, 1201-1233.

Jun, S.-A., & Kim, S. (2004). Default phrasing and attachment preferences in Korean.

Proceedings of INTERSPEECH-ICSLP, Jeju, Korea.

Jun, S.-A., & Koike, C. (2003). Default prosody and RC attachment in Japanese. Talk given at the13th Japanese-Korean Linguistics Conference, Tucson, AZ. [Published in *Japanese-Korean Linguistics 3,* 41_53, CSLI, Stanford, in 2008]

Schafer, A.J., Speer, S.R., Warren, P., & White, D. (2000). Intonational disambiguation in sentence production and comprehension. *Journal of Psycholinguistic Research, 29,* 169-182.

Speer, S.R., Warren, P., & Schafer, A.J. (2011). Situationally independent prosody. *Laboratory Phonology, 2,* 1, 35–98

Swets, B., Desmet, T., Hambrick, D. Z., & Ferreira, F. (2007). The role of working memory in syntactic ambiguity resolution: A psychometric approach. *Journal of Experimental Psychology: General*, 136, 64-81.

Terken, J., & Hirschberg, J. (1994). Deaccentuation of words representing Given information: contributions of persistence of grammatical function and surface position. *Language and Speech, 37*, 125-145.

Vasishth, S., Agnihotri, R. K., Fernández, E. M., & Bhatt, R. (2004). Noun modification preferences in Hindi. In *The Proceedings of Seminar on Construction of Knowledge,* Vidya Bhawan Education Resource Centre, Udaipur, India, pp. 160_171.

Wijnen, F. (2004). The implicit prosody of jabberwocky and the relative clause attachment riddle. In H. Quené & V. van Heuven (Eds.), On speech and language. *Studies for Sieb G. Nooteboom,* pp. 169-178. Utrecht: Landelijke Onderzoeksschool Taalwetenschap. (LOT Occasional Series 2).