Empirical Investigations of the Role of Implicit Prosody in Sentence Processing

Mara Breen*
Mount Holyoke College

Abstract
Recently, psycholinguistics has seen an increase in the number of empirical studies investigating the role of implicit (silent) prosodic representations in reading. The current paper reviews studies from the last several years conducted to investigate Fodor’s (2002) Implicit Prosody Hypothesis, which maintains that even during silent reading, readers generate representations of sentence intonation, phrasing, stress, and rhythm, and that these representations can affect readers’ interpretation of the text. We argue that the accumulated evidence suggests that implicit prosody can influence online sentence interpretation and explore the implications of these findings for models of sentence processing.

For over one hundred years, researchers have wondered about the nature of the inner voice during silent reading. Huey (1908/1968) was one of the first to ponder this idea, concluding: ‘The simple fact is that the inner saying or hearing of what is read seems to be the core of ordinary reading, the “thing in itself,” so far as there is such a part of such a complex process’ (p. 122). This assumption, that the inner voice is part and parcel of any normal reading, has been maintained for the majority of the 20th century. Chafe (1988) recounts the writings of Eudora Welty and Russell C. Long on the topic, concluding:

‘I am not alone in believing that writers when they write, and readers when they read, experience auditory imagery of specific intonations, accents, pauses, rhythms, and voice qualities, even if the writing itself may show these features poorly, if at all. This “covert prosody” of written language is evidently something that is quite apparent to a reflective writer or reader’ (p. 397).

An increasing interest in spoken language over the last 25 years has inspired psycholinguistic researchers to begin to critically consider the role of the inner voice during reading.

One of the main questions concerning researchers is whether the inner voice serves a purpose during reading. Is the producing, or hearing, of words and phrases during reading simply epiphenomenal—a by-product of the fact that language has been spoken far longer than it has been written (Gelb 1952), or does it enhance the reader’s processing and understanding of the written word? The goal of the current paper is to review recent psycholinguistic investigations of silent reading to begin to answer this question.

Studies of phonology’s role in reading have focused on two types of phonological representation: segmental phonology and suprasegmental phonology. Segmental phonology concerns the individual phonemes which make up words, whereas suprasegmental phonology deals with sound phenomena above the level of the word; that is, acoustic information that does not serve to distinguish one word from another, but rather conveys information about the semantic context of the word, or the attitude of the speaker. For example, the segmental features of the word ‘fire’ include the phonemes /fɪ/ /ə/ and /r/, whereas suprasegmental features determine whether the word is produced as a statement (‘Fire.’), a question (‘Fire?’),
or even an exclamation (‘Fire!’). We will briefly explore empirical investigations of implicit segmental phonological representations during silent reading before turning to suprasegmental representations, which are the focus of the current paper.

There is a long history of debate about the extent to which segmental phonological representations are activated during silent reading, particularly with respect to how these representations affect word recognition (see Frost 1998, and McCusker, Hillinger, and Bias 1981, for reviews). Those who have claimed an important role of segmental phonology in word recognition argued from a variety of findings. For example, Van Orden (1987) demonstrated that homophones are processed similarly to their sound-alike counterparts such that it’s harder for readers to reject the word rows, which sounds like the target rose, as a member of the category flower than it is to reject role, which looks like the target. Spelling and sound correspondence can also slow processing in other ways: Treiman, Freyd, and Baron (1983) showed that it’s harder for people to find a continuation for the phrase nasty hasty than the synonymous phrase mean rash, presumably due to the mismatch between the similar spelling, but dissimilar pronunciation, of nasty and hasty. Finally, McCutchen and Perfetti (1982) demonstrated that tongue twisters like (1a) are read more slowly than corresponding non-sound-alike pairs (1b).

(1)  
a. Barbara burned the brown bread badly.  
b. Samuel caught the high ball neatly.

In all of these cases, reading difficulty is due to similarity (or dissimilarity) at the level of individual phonemes. This substantial support for segmental phonological information in word recognition suggests that suprasegmental phonological information also plays a functional role in reading above the level of the word.

A review of suprasegmental phonology’s role in reading was provided by Rayner and Pollatsek (1989, see also Rayner, Pollatsek, Ashby, and Clifton 2011), who used the term phonological coding to refer to ‘the mental representations of speech that can give rise to the experience of hearing sounds’ (p. 189). Based on reviews of empirical explorations of inner speech, which relied primarily on electromyography (i.e. recordings of articulatory muscles) and articulatory suppression during reading, they proposed two possible roles for phonological coding. First, they hypothesized that this phonological representation serves to strengthen a reader’s memory for the text (see Gathercole and Baddeley 1993). This argument was inspired by Patterson and Coltheart’s (1987) explanation of segmental phonological effects on word recognition, which applied Barnard’s (1985) reasoning that a buffer allows for the integration of cognitive processes, which may operate on different time scales. Because short-term memory representations are thought to be predominantly acoustic in nature, the generation of a phonological code strengthens a reader’s memory for individual words, giving him time to integrate them into the larger syntactic and semantic context.

Rayner and Pollatsek’s (1989) second proposal, based on the work of Slowiaczek and Clifton (1980), is that phonological coding provides prosodic structure, which in turn serves to provide cues to the organization of words. Specifically, Slowiaczek and Clifton denied these readers this phonological information through the suppression of subvocalization, implemented by having readers repeat nonsense syllables (‘colacolacola...’) while reading silently. When compared to readers who read without suppression, these readers were poorer at answering questions that required integrating propositions across sentences. The authors argue that the generation of a prosodic structure provides more cues to the structural and information status of sentences, through information about phrasing and focus.
In the sections that follow, we will explore the experimental evidence for this second proposal: The claim that readers generate prosodic representations of text during reading. However, we first need to specify what we mean by implicit prosody; that is, what are the prosodic features that are candidates for having an implicit representation?

**Prosody in Spoken Language**

Generally speaking, there are four subdivisions of spoken prosody: phrasing, stress, rhythm, and intonational contour. We will describe each of these in turn, including descriptions of their realization in spoken language.

Phrasing refers to the way words are combined perceptually into groups. There is a strong link between intonational contour and phrasing, such that words are combined into intonational phrases. Characteristic perceptual features mark the boundaries between these intonational phrases (Nespor and Vogel 1986) including the increased duration of pre-boundary words (Ferreira 1993; Lehiste, Olive, and Streeter 1976; Price, Ostendorf, Shattuck-Hufnagel, and Fong 1991; Schafer, Speer, Warren, and White 2000; Selkirk 1984; Snedeker and Trueswell 2003; Wightman, Shattuck-Hufnagel, Ostendorf, and Price 1992; Breen, Watson, and Gibson 2011), the raising or lowering of pitch prior to a boundary (Pierrehumbert 1980; Streeter 1978) and silence (Cooper and Paccia-Cooper 1980; Lehiste 1973).

Sentence phrasing is determined in part by syntactic structure, such that intonational phrase boundaries correlate with syntactic boundaries (Selkirk 1984; Schafer et al. 2000; Snedeker and Trueswell 2003; Cooper and Paccia-Cooper 1980; Watson and Gibson 2004; Breen et al. 2011; Gee and Grosjean 1983; Ferreira 1988). For example, the boundary in the spoken version of (3b) (indicated by ‘//’) (from Snedeker and Trueswell 2003) is typically interpreted by listeners as indicating the presence of a syntactic boundary; a listener would interpret the flower as the instrument used for tapping, rather than as a modifier specifying a particular frog to be tapped.

(2)

a. Tap the frog with the flower.
b. Tap the frog // with the flower.

The second component of spoken prosody is stress. Stressed syllables are more perceptually prominent than adjacent, unstressed, ones. Stresses are generally arranged in a hierarchical manner, such that they are nested within each other (Hayes 1995). Specifically, multisyllabic content words have one primary stressed syllable. Moreover, individual phrases have a main stress, or accent, which is determined in large part by the information structure of the sentence (Selkirk 1984; Gussenhoven 1983; Rooth 1996; Breen, Fedorenko, Wagner, and Gibson 2010). In (4), the location of the main accent in the sentence Damon fried an omelet will change depending on the question in (3):

(3)

a. Who fried an omelet?
b. What did Damon do to an omelet?
c. What did Damon fry?
(4)Damon fried an omelet.

The acoustic correlates of accents include intensity and duration, such that accented syllables are generally discriminated from unaccented ones by their greater duration (Cooper, Eady, and Mueller 1985; Eady and Cooper 1986) and greater intensity (Beckman 1986; Turk and Sawusch 1996; Kochanski, Grabe, Coleman, and Rosner 2005).
The third component of prosody that we will consider is rhythm. Hayes (1995) argues that every word has one main stressed syllable. Moreover, he argues, based in part on the claims of Selkirk (1984), that speakers distribute these stresses equally across utterances, such that they occur at regular intervals (cf. Nespor and Vogel 1989; Kelly and Bock 1988). In addition, empirical work has supported the idea that listeners perceive stresses at regular intervals, as evidenced by Pitt and Samuel’s (1990) finding that phoneme monitoring is faster for phonemes which occur in syllables predicted to be stressed based on rhythmic context. The final study we will present investigates whether readers also appear to prefer regularly-spaced stresses (Kentner 2012).

The fourth component of prosody that we will consider is intonational contour, which describes the overall tune of an utterance, including the location and direction of pitch changes. These pitch changes play an important role in specifying the semantic content of utterances (Pierrehumbert and Hirschberg 1990). For example, (5a) which is an imperative statement, would likely be produced with falling pitch on the final syllable whereas (5b), which is a yes-no question, would likely be produced with rising pitch on the final syllable.

(5)

a. You’re going to the store.
b. You’re going to the store?

Implicit Prosody Proposals

Although there has been continuous interest in the ‘inner voice’ for over 100 years, most modern investigations of implicit prosody were inspired by two related proposals by Fodor (1998) and Bader (1998). Although these two proposals were suggested to address two different questions in the sentence processing literature, they are similar in substance. We will first describe Fodor’s proposal, and then turn to Bader’s.

Fodor (1998) argued that sentence comprehension relies not only on a syntactic parser, which incrementally assigns syntactic structure to words in a sentence, but also a prosodic parser, which operates in parallel with the syntactic parser. The prosodic parser packages sentence material into phrases that are roughly balanced in length, what she terms the same-size sister constraint (cf. Grosjean, Grosjean, and Lane 1979; Gee and Grosjean 1983). If, as Fodor argues, the syntactic parser interprets only those parts of the sentence that are in a prosodic phrase, then implicit phrasing will have implications for syntactic attachments. She argues that this process is responsible for the shifting in attachment preference across the examples in (6).

(6)

a. The divorced bishop’s daughter
b. The recently divorced bishop’s daughter
c. The recently divorced bishop’s daughter-in-law

Fodor claims that readers are as likely to assume that (6a) is referring to a divorced bishop as to a divorced daughter. However, (6b) is too long to be produced in one prosodic phrase, so the prosodic packager will break it up and, as balanced phrases are preferred, the packages will be *The recently divorced* and *bishop’s daughter*, resulting in readers’ preference to assume that the daughter is divorced. Fodor argues that the preference switches in (6c), where the prosodic packages are *The recently divorced bishop’s* and *daughter-in-law* and the preference is for a divorced bishop. The same-size sister constraint falls under a larger hypothesis that Fodor (2002) termed the Implicit Prosody Hypothesis:
(7) The Implicit Prosody Hypothesis (IPH): In silent reading, a default prosodic contour is projected onto the stimulus, and it may influence syntactic ambiguity resolution. Other things being equal, the parser favors the syntactic analysis associated with the most natural (default) prosodic contour for the construction.

Fodor argues that the IPH can also account for differences in language’s relative clause (RC) attachment preferences. She suggests that those languages that favor high RC attachment are more likely to impose a prosodic break before a relative clause, thereby encouraging high attachment.

Whereas Fodor proposed an implicit prosodic theory to account, in part, for differences in cross-linguistic parsing preferences, Bader (1998) proposed the Prosodic Constraint on Reanalysis to explain why some cases of syntactic reanalysis are more difficult than others. Specifically, he claimed that it is harder to revise a syntactic misanalysis if doing so requires a concomitant prosodic reanalysis.

Bader bolstered his claim with on-line evidence from German, demonstrating longer reading times on the disambiguating region of a sentence when (a) the syntactic construction it signaled was dispreferred and (b) the reanalysis it signaled required a simultaneous shift in phrasal stress. In an eye-tracking study, participants read sentences like those in (8):

(8) Zu mir hat Maria gesagt,
  to me Maria has said
  ‘Maria said to me’
  a. …daß man (sogar) ihr Geld anvertraut hat.
     …that one (even) her money entrusted has
     ‘…that someone entrusted money (even) to her.’
  b. …daß man (sogar) ihr Geld beschlagnahmt hat.
     …that one (even) her money confiscated has
     ‘…that someone confiscated (even) her money.’

In (8a), ihr functions as an object pronoun, as the indirect object of entrusted, but in (8b), it functions as a possessive pronoun, specifying who the money belongs to. The latter interpretation is the preferred one, and the default phrasing of both interpretations (without sogar) is the same, namely, with stress on man and no stress on ihr. However, the default phrasing of the two sentences changes with the addition of the focus particle sogar. Specifically, Bader argues that readers first interpret ihr in (8a) as a possessive pronoun. As function words are usually not accented, the default would be for the reader to accent Geld in this production (leaving ihr unaccented). However, upon encountering anvertraut, the reader would have to reanalyze both syntactically and prosodically, first reinterpretating ihr as an object pronoun, and second, shifting the accent from Geld to ihr. Indeed, Bader (1998) observed that reanalysis in (8a) was more difficult with the presence of the focus particle than without as evidenced by longer reading times on the disambiguating region comprised of the final two words, which he argued was due to the additional cost of updating the implicit prosodic representation.

The development of these independent proposals by Fodor and Bader has inspired a new set of investigations into the reality and nature of implicit prosodic representations. Moreover, the great strides made in the field of psycholinguistics over the last 20 years mean that implicit prosody researchers have access to new experimental methods, including self-paced reading, eye-tracking, and event-related potentials, which make it possible to investigate language processing in real time.
In the next section of the paper, we will review experimental evidence for the role of implicit prosody in on-line sentence processing. The studies are organized with respect to the prosodic phenomena they address. We begin with investigations of the role of implicit phrasing, followed by stress, rhythm, and intonation.

**Implicit Prosodic Phrasing**

Implicit prosodic phrasing has been investigated more extensively than any other aspect of prosody. Most, if not all, of the studies described below take as their starting point Fodor’s (1998, 2002) proposals, described above, that implicit prosody affects attachment decisions, and that readers prefer to produce, both overtly and implicitly, phrases of similar sizes. As stated above, multiple studies have demonstrated that overt prosodic phrasing can influence syntactic attachment. For example, we know that listeners resist syntactic attachments across overt prosodic boundaries (e.g., Snedeker and Trueswell 2003; Schafer et al. 2000; Pynte and Prieur 1996). Of course, researchers cannot explicitly manipulate the presence of phrase boundaries in what participants read; therefore, they have generally followed Fodor’s (2002) four-step procedure to test for implicit prosodic effects on attachment preferences:

1. Find a factor \( F \), which can be manipulated in an experiment, and which measurably affects the OVERT prosody of a sentence.
2. Show that the overt prosodic difference caused by \( F \) measurably influences an ambiguity resolution preference in parsing.
3. Show (or claim?) that \( F \) does not affect parsing DIRECTLY.
4. Include \( F \) in a silent reading task. Is ambiguity resolution affected by \( F \) as it is in the listening task?

Using this method, effects of implicit prosodic boundary placement on attachment decisions have been reported across a wide variety of languages including Japanese (Kitagawa and Fodor 2006), English (Quinn, Abdelghany, and Fodor 2000; Swets, Desmet, Hambrick, and Ferreira 2007), German (Augurzky 2006), Croatian (Lovric 2003), Hindi (Vasishth, Agnihotri, Fernández, and Bhatt 2005), Dutch (Wijnen 2004), French (Pynte and Colonna 2000), and Korean (Hwang and Schafer 2009), among others (Jun 2003).

Fodor’s procedure was adopted by Hirose (2003), who explored implicit prosodic effects on relative clause attachment in Japanese. She manipulated the length of the subject in sentence fragments like (10). Length influences overt boundary placement in that speakers are more likely to place boundaries after longer constituents than shorter ones (Watson and Gibson 2004; Watson, Breen, and Gibson 2006; Breen et al. 2011; Ferreira 1993; Cooper and Paccia-Cooper 1980; Gee and Grosjean 1983).

(10)
(Hoso’kawa-to) Mori’sita-ga si’nyaku-o kokoro’kara (Hoso’kawa-and) Morisita-Nom new medicine-Acc truly sinyoosita yuujii’tntati-ni… trusted friends-Dat…

In these fragments, the interpretation of the object of the verb of a relative clause shinyooshita (trusted) is ambiguous between yuujii’nntati-ni (friends) and si’nyaku-o (medicine).
When the subject was a long conjoined noun phrase, readers were more likely to provide a continuation indicating their interpretation of the object of the verb as *medicine*. Hirose interpreted this pattern as indicating that, in the version with the short subject, readers placed an implicit prosodic boundary after *medicine*, thereby blocking an attachment between *medicine* and *trusted*. Conversely, in the long subject condition, readers placed a boundary after the subject and before the object, thereby allowing an attachment between *medicine* and *trusted*.

Sato, Kobayashi, and Miyamoto (2007) provided further support for the claim that Hirose’s observed results are not due to memory load or information status (see Fodor’s 3rd step). Specifically, they demonstrated that deaf readers, most deaf since birth, were not subject to the same length effects as hearing readers. Under the assumption that deaf readers are not generating implicit prosodic phrasing during silent reading, these results support the claim that Hirose’s (2003) effects are indeed prosodic in nature.

A second research path exploring the role of prosody in syntactic attachment that has received some empirical support has been inspired by Fodor’s (1998) *same size sister constraint*, which, as described above, predicts that the desire for prosodic balance can influence parsing decisions. Augurzky (2008) demonstrated a higher probability of a boundary in cases where boundary placement resulted in balanced phrases (see also Hirose 1999).

Despite the confluence of cross-linguistic findings supporting a role of implicit prosody in syntactic ambiguity resolution, some data call this relationship into question. Results from Jun (2010) and Bergmann, Armstrong, and Maday (2008) suggest that implicit prosodic phrasing may not be fully consistent with explicit phrasing. Jun (2010) demonstrated that speakers’ overt prosodic phrasing of out-of-the-blue productions of sentences with relative clause ambiguities did not match the implicit prosodic phrasing that would have led to the same speakers’ interpretations of the ambiguities. She argues that these differences may be due to readers’ slower, more deliberate productions of unfamiliar material during overt reading, as compared to faster silent reading. Jun suggests that, rather than focusing on the results of off-line attachment decisions as evidence for implicit prosody, researchers should instead look to on-line methods, like event-related potentials (ERPs) and eye-tracking.

In fact, many researchers have investigated implicit prosody using such online methods. For example, ERP studies on the topic have utilized Steinhauer, Alter, and Friederici’s (1999) discovery of an ERP component which seems maximally responsive to overt prosodic phrasing. This component, the Closure Positive Shift (CPS), is a large, widespread, positive deflection, observed in response to overt intonational phrase boundaries. Steinhauer (2003) demonstrated that the CPS elicited by overt prosodic boundaries is very similar in timing and morphology to the one elicited by commas in the same sentences read silently. Moreover, Steinhauer and Friederici (2001) had participants read locally ambiguous sentences, which could be disambiguated by the presence of a prosodic boundary. Before reading, participants listened to a filtered version of an overt production of the target sentence, which maintained the prosodic contour but not the lexical material, and were directed to apply the prosodic contour to the subsequently read sentence. The results revealed a CPS in the locations where the prosodic contour would have induced readers to postulate an implicit prosodic boundary during silent reading, an effect the authors interpreted as strong support for the fact that the CPS was reflecting implicit prosody.

Most recently, Hwang and Steinhauer (2011) had participants read locally ambiguous Korean sentences, which are effectively disambiguated by the presence of an overt phrase boundary after an initial noun phrase (NP). The authors varied the length of the initial NP, and observed a CPS only after the long sentence-initial NP. Furthermore, subsequent explicit disambiguation to the less preferred interpretation of the sentence elicited a smaller ERP marker of syntactic difficulty (the P600) when the NP was long. The authors argue that, when the initial NP was long, readers imposed an implicit prosodic boundary, which
served to ameliorate subsequent garden path effects (see Liu, Wang, and Zhixing 2010 for similar results from Chinese).

Researchers have also begun to explore the question of implicit phrasing by manipulating comma presence in syntactically ambiguous and unambiguous sentences. Several researchers have demonstrated that comma presence ameliorates garden path effects (Clifton 1993; Mitchell and Holmes 1985). Hill and Murray (2000), for example, reported results from eye-tracking studies of several well-known garden-path sentences and demonstrated that although comma presence resulted in slower reading times at the location of the comma, they led to faster reading times following it. However, commas are not only useful in cases of syntactic ambiguity, and some have argued that reading is generally facilitated by commas in certain syntactic positions. For example, Staub (2007) provided evidence that commas facilitate reading even in unambiguous sentences, demonstrating shorter reading times on the vet and and his assistant in (11) when there was a comma before the vet.

(11)

a. When the dog arrived(,) the vet and his assistant went home.
b. When the dog arrived at the clinic(,) the vet and his assistant went home.

Importantly, the manipulation of the subordinate clause did not interact with comma presence; the presence of a comma facilitated reading of (11b), even though the temporary ambiguity that is arguably present in (11a) is not present in (11b). Staub (2007) interpreted this result as suggesting that a comma is generally preferred at the boundary between the subordinate clause and main clause, for reasons that are unrelated to syntactic ambiguity.

Further evidence for the claim that commas facilitate parsing whenever they coincide with clause boundaries comes from Hirotani, Frazier, and Rayner (2006), who compared reading times for simple and complex sentences presented with and without commas, as in (12). They observed shorter whole-sentence reading times for sentences with commas than without, suggesting that commas facilitated reading. Moreover, similar to the results of Staub (2007), they found that the length of the sentence did not interact with the comma effect. They interpreted these results as indicating that, irrespective of the length of the adjacent material, readers dwell on the material before commas to do any necessary integrative processing before moving on to the next phrase or clause.

(12)

a. The freshmen(,) who attended(,) were polite and courteous.
b. The freshmen(,) who in the end attended(,) were polite and courteous.

Both Staub (2007) and Hirotani et al. (2006) suggest that these results are due to the fact that commas are the written manifestation of implicit prosodic boundaries, such that correctly-placed commas correspond to the location of prosodic boundaries that the reader would produce were she reading the sentence aloud. Corroborative evidence for this claim comes from Ren and Yang (2010), who conducted an eye-tracking study of the role of commas in written Chinese, in which word boundaries are not explicitly marked. They observed shorter reading times after commas only when the comma marked a clause boundary, and not when the comma marked a word or phrase boundary.

**Implicit Stress Assignment**

Several studies have now explored the role of implicit stress assignment in silent reading. These studies have explored the extent to which lexical stress (Wilkenfeld 1985; Ashby
and Clifton 2005; Breen and Clifton 2011, 2013) and accents (Bader 1998; Stolterfoht, Friederici, Alter, and Steube 2007) are represented in silent reading.

Ashby and Clifton (2005) demonstrated longer reading times for four-syllable words with two stressed syllables (e.g., *UltiMAtum*) than for frequency-matched and length-matched four-syllable words with one stressed syllable (e.g., *inSANity*). They interpreted the result as evidence that readers routinely assign lexical stress patterns to silently read words. Additional evidence for the recruitment of implicit word-level metrical representations in reading comes from Wilkenfeld (1985), who used written words to prime spoken targets. The targets were stress-alternating homographs (e.g., *object*), which can have a strong-weak (*Object*) or weak-strong (*objECT*) pattern depending on syntactic category, but which have a preferred metrical pattern (i.e., *Object*). Participants could be induced to produce the dispreferred metrical pattern of the target if they first silently read a list of eight words which also had the dispreferred pattern. Wilkenfeld interpreted this result as evidence that the implicit prosodic representation of the eight preceding words effectively primed the stress pattern of the target.

Breen and Clifton (2011, 2013) provided evidence of an on-line role for implicit metrical representations in silent reading by embedding stress-alternating noun–verb homographs in garden-path sentences, as in (13):

(13)

a. The brilliant *abstract* the best ideas from the things they read.

b. The brilliant *report* the best ideas from the things they read.

c. The brilliant *abstract* was accepted at the prestigious conference.

d. The brilliant *report* was accepted at the prestigious conference.

Results demonstrated longer reading times on *from the things they read* in (13a, 13b) than *at the prestigious conference* in (13c, 13d), demonstrating that readers had to shift their syntactic representation of both *abstract* and *report* from noun to verb. Moreover, the reading time cost was greater for (13a) than (13b), which Breen and Clifton interpret as evidence that in (13a) readers also must shift the stress pattern of *abstract*, from *ABstract* to *abSTRACT*.

In addition to evidence that lexical stress is part of the default representation of silent reading, there is also evidence that phrasal stress is assigned to read material. As described above, Bader (1998) demonstrated a reading time cost when syntactic reanalysis required a concurrent reanalysis of sentence stress, such that an accent needed to be shifted from one word to another. In a related ERP experiment, Stolterfoht, Friederici, Alter, and Steube (2007) observed a negative-going wave 450–650 ms after presentation of a word which signaled the need for a shift in the location of the nuclear accent. Importantly, this effect was distinct from a late positive wave elicited by the need to revise the focus structure of the sentence. Finally, Kitagawa, Tamaoka, and Tomioka (2013) demonstrated higher acceptability judgments for Japanese sentences where an explicit phrase accent disambiguated focus location than for sentences without the phrase accent where, presumably, readers were silently assigning a default focus accent to a non-focused element.

Implicit Linguistic Rhythm

Kentner (2012) conducted a series of experiments investigating the extent to which rhythmic constraints influence on-line parsing decisions. Specifically, speakers of English and German seem to prefer producing speech with alternating strong and weak syllables (Hayes 1995; Selkirk 1984). As such, they tend to avoid stress clashes, in which two adjacent syllables are stressed. This avoidance of stress clash has been observed in overt production (Kelly
and Bock 1988; Anttila, Adams, and Speriosu 2010), which begs the question of whether stress clash is also avoided in silent reading. To test this hypothesis, Kentner employed stimuli like those in (14) where a temporary syntactic ambiguity can be disambiguated by stress. The ambiguity arises because *mehr* is ambiguous up to the final phrase of the sentence; it is either part of the temporal adverbial *nicht mehr* (14a) or a comparative quantifier (14b).

(14) Der Polizist sagte, dass man . . .  
The policeman said that one . . .

a. . . . nicht mehr NACHeNweisen/erMITteln kann, wer der Täter war.  
. . . couldn’t prove/determine anymore who the culprit was.

b. . . . nicht MEHR NACHeNweisen/erMITteln kann, als die Tatzeit.  
. . . couldn’t prove/determine more than the date of the crime.

In an unprepared reading task, where participants began reading aloud without first silently reading the sentence, readers were more likely to accent *mehr* when it preceded *ermitteln* than when it preceded *nachweisen*, a result that Kentner attributes to readers’ avoidance of a stress clash between *mehr* and the first syllable of *nachweisen*. This effect was also evident in a silent reading task, such that reading times were longer on the disambiguating final phrase of the sentence in (14b) for *ermitteln* than *nachweisen*. Kentner argues that the implicit rhythmic structure of *mehr nachweisen* would lead to an initial interpretation of *mehr* as an (unstressed) temporal adverbial, an interpretation that would have to be reanalyzed upon encountering the final phrase, which signals the need for reanalysis of *mehr* as a quantifier.

As observed for phrasing and stress assignment, Kentner’s results provide additional evidence for the on-line role of implicit prosodic representations in sentence comprehension. Moreover, Kentner’s method may reveal implicit rhythmic effects in other stress-timed languages whose speakers (and readers) are likely to avoid stress clashes.

Implicit Intonation

There has thus far been fairly little in the way of investigation of implicit intonational contours. One example, however, comes from Abramson (2007), who demonstrated that subjects had better memory for auditorily-presented words which matched the intonational contour of a visually-presented study passage. Subjects read study sentences in which a target word was presented as the last word in either a declarative or an interrogative sentence, as in (15):

(15) a. He said: ‘I want to open the package.’  
b. He said: ‘Do you want to open the package?’

When tested five minutes after study, subjects performed faster lexical decision for old targets (e.g., *package*) that were presented in a tonally consistent way (i.e., with rising or falling intonation). This single study suggests that the investigation of implicit intonational contours is a fruitful avenue for future study.

The Function of Implicit Prosody

The preceding sections have reviewed empirical evidence for a role for implicit prosody in each of four subdivisions of prosody: phrasing, stress, rhythm, and intonational contour.
The weight of the evidence supports the claim that suprasegmental representations are not epiphenomenal. Moreover, they are not simply a post-processing translation of written material into more durable acoustic representations for the purpose of strengthening memory for the text and extending the time window for interpretive processing. Rather, implicit prosody plays a functional role in on-line language comprehension. That is, the generation of an implicit prosodic representation of written words can serve to direct readers’ interpretations of ambiguous structures and can facilitate processing of written language, or, conversely, can impair comprehension of an intended message.

This conception of implicit prosody puts it on par with other sources of information known to influence on-line language comprehension, including knowledge about word and category frequency, discourse structure, world knowledge, and explicit prosody. Future study of implicit prosody should focus on understanding how these representations interact with other sources of information in language comprehension.

Future research should also address individual differences in implicit prosodic representations. Jun (2003), for example, has demonstrated that individual speakers’ R.C attachment decisions are consistent with their overt prosodic phrasing. Moreover, results from Swets et al. (2007), suggest that these individual differences in attachment decisions may be due to differences in working memory capacity. Further studies should explore to what extent implicit prosodic representations differ across individuals, and how these differences might impact not only attachment decisions, but also the interpretation of accent and discourse.

Finally, there is preliminary evidence of a strong connection between children’s prosodic fluency and their reading comprehension skill such that readers who produce fluent prosody are also better comprehenders (Schwanenflugel, Hamilton, Kuhn, Wisenbaker, and Stahl 2004). For example, Miller and Schwanenflugel (2008) demonstrated, in a longitudinal study, that children who produced ‘adult-like’ prosody at the end of first and second grade were found to be more fluent readers, and better comprehenders, at the end of third grade. Importantly, prosodic fluency has been shown to independently predict reading comprehension above and beyond other factors known to influence comprehension, such as reading rate and decoding skill (Miller and Schwanenflugel 2006), particularly for more difficult texts (Benjamin and Schwanenflugel 2010). Future studies could determine whether implicit prosody is mediating this effect, thereby increasing understanding of the processes underlying skilled reading.

In summary, the research reviewed here has demonstrated the effective application of the tools of psycholinguistics to the study of implicit prosody, demonstrating its importance in sentence comprehension. In addition, this work highlights the many remaining questions to be explored about the details of implicit prosodic representations, their interaction with other sources of knowledge, and their variability across individuals and development.

Short Biography

Mara Breen’s research explores the features of spoken language as well as how these features are processed and interpreted by listeners. In addition, her work explores the ways in which speech production and comprehension interface with attention and memory processes. She has co-authored papers on these topics for Journal of Memory and Language, Language and Cognitive Processes, Cognition, Attention Perception & Psychophysics, and Corpus Linguistics and Linguistic Theory. Prior to teaching in the Psychology and Education Department of Mount Holyoke College, she taught and conducted research at the University of Massachusetts Amherst. She holds a BA in Cognitive Science from Hampshire College and a PhD in Cognitive Science from the Massachusetts Institute of Technology.
Note

*Correspondence address: Mara Breen, Department of Psychology and Education, Mount Holyoke College, South Hadley, MA 01075. E-mail: mbreen@mtholyoke.edu

Works Cited


Augurzky, P. 2006. Attaching relative clauses in German: the role of implicit and explicit prosody in sentence processing. MPI Series in Human Cognitive and Brain Sciences 77. Leipzig, Germany.


