

7-29-2017

Marking Stress ExPLICitly in Written English Fosters Rhythm in the Reader's Inner Voice

Jennifer Gross

Grand Valley State University, grossj@gvsu.edu

Bo Winegard

Florida State University

Andrea Plotkowski

Grand Valley State University

Follow this and additional works at: https://scholarworks.gvsu.edu/oapsf_articles

 Part of the [English Language and Literature Commons](#), and the [Psychology Commons](#)

ScholarWorks Citation

Gross, Jennifer; Winegard, Bo; and Plotkowski, Andrea, "Marking Stress ExPLICitly in Written English Fosters Rhythm in the Reader's Inner Voice" (2017). *Funded Articles*. 87.

https://scholarworks.gvsu.edu/oapsf_articles/87

This Article is brought to you for free and open access by the Open Access Publishing Support Fund at ScholarWorks@GVSU. It has been accepted for inclusion in Funded Articles by an authorized administrator of ScholarWorks@GVSU. For more information, please contact scholarworks@gvsu.edu.

Marking Stress ExPLICitly in Written English Fosters Rhythm in the Reader's Inner Voice

Jennifer Gross

Grand Valley State University, Allendale,
Michigan, USA

Bo Winegard

Florida State University, Tallahassee,
USA

Andrea R. Plotkowski

Grand Valley State University, Allendale,
Michigan, USA

ABSTRACT

Spoken English has a stress-alternating rhythm that is not marked in its orthography. In two experiments, the authors evaluated whether stylistic alterations to print that marked stress pulses fostered the rendering of rhythm (experiment 1) and stress (experiment 2) during silent reading. In experiment 1, silent readers rated the helpfulness of the stylistic alterations appearing in the last line of poems. In experiment 2, silent readers rated the helpfulness of the stylistic alterations appearing in heteronyms embedded in prose. As predicted by linguistic theories, when the stylistic alterations mapped onto the rhythmic pulses of the poems, and the lexically stressed syllables of the heteronyms, silent readers rated these alterations as more helpful compared with the incongruous conditions. In experiment 2, readers' inner voices were more tuned to the prosodic nuances of the first syllable than the second in the bisyllabic heteronyms. This prosodic tuning for the first syllable in a word was likely afforded by the strong tendency for stress to appear word-initially. In addition, the stylistically marked stress was viewed as more helpful in the early half of the sentence, when readers likely recruited more bottom-up processes. In both experiments, prior exposure to poetry was related to a refined prosodic awareness. In experiment 2, exposure to poetry predicted participants' prosody sensitivity, after controlling for the other predictors of academic achievement. The authors' ongoing studies are evaluating whether marking stress explicitly in written English might aid struggling readers and late speakers of English.

The English writing system is a printed representation of its spoken language. Written English is based on an alphabetic system that maps letters in print to sounds in spoken English. Understanding the difference between *cat* and *rat* and between *lice* and *rice* is an indicator of alphabetic literacy in English. Mastering the letter-sound correspondences (a marker of phonological awareness) is a powerful cipher when learning to read (Patterson & Coltheart, 1987; Perfetti & Bell, 1991; Van Orden & Kloos, 2005). Skilled reading reflects more than the mastery of letter-sound correspondences, however. Becoming a fluent reader requires an understanding of the prosody of English (Clin, Wade-Woolley, & Heggie, 2009; Dowhower, 1991; Holliman, Wood, & Sheehy, 2010; Miller & Schwanenflugel, 2008; Schreiber, 1991; Schwanenflugel, Hamilton, Kuhn, Wisenbaker, & Stahl, 2004; Whalley & Hansen, 2006). Prosody is the vocal properties of speech, including stress, rhythm, timing, and emphasis. For example, the difference in meaning between "A rat chased my cat?" and "A rat chased your cat!" is conveyed by prosody. Markers of prosodic fluency include reading with appropriate pausing and phrasing,

phrase-final lengthening, intonation contours, and stress assignment (Dowhower, 1991; National Reading Panel, 2000).

From early readers to adults, prosody sensitivity plays a role in reading. In 5- and 6-year-olds, prosody sensitivity uniquely predicted reading, after controlling for vocabulary knowledge, phonological awareness, and morphological awareness (Holliman et al., 2017). Third graders' knowledge of lexical prosody when reading aloud from a typical textbook was a significant predictor of oral reading fluency and good reading comprehension (Schwanenflugel & Benjamin, 2016). For example, young readers displayed knowledge of lexical prosody when they placed stress in the appropriate place for the noun and verb forms of words such as *recall* and showed syllable stress shifting when suffixes were added (e.g., *Artist* vs. *artistic*). Third graders also displayed knowledge of lexical compound contrasts when reading aloud words such as *BLACKbird* compared with *BLACK* and *BIRD*. In high school students, skill in prosodic production predicted reading comprehension among students matched on decoding ability (Breen, Kaswer, Van Dyke, Krivokapić, & Landi, 2016). Specifically, high school students with poor reading comprehension consistently produced weaker duration cues to mark the appropriate syntactic structure.

Prosodic awareness also predicted reading achievement in adult readers (Wade-Woolley & Heggie, 2015). Common measures of prosodic awareness (e.g., identifying the number of syllables in a multisyllabic word, identifying which syllable contains the main beat of the word) might be confounded with working memory. Thus, Chan and Wade-Woolley (2016) evaluated the unique contributions of working memory and prosodic awareness to word reading and reading comprehension in adults. Both executive function and prosodic awareness were related to word reading and reading comprehension. Yet, after controlling for working memory, prosodic awareness still accounted for adults' word-reading abilities (Chan & Wade-Woolley, 2016). Thus, across development, prosody sensitivity is related to reading abilities.

Although many features of spoken language (e.g., words, syntax) are easily mapped onto their written counterparts, written English does not explicitly represent the prosodic features of stress and rhythm (Fudge, 1984; Treiman & Kessler, 2005). Faced with the paucity of cues, fluent readers must infer stress and rhythm when reading aloud (Goswami et al., 2011; Nespor & Vogel, 1986; Schreiber, 1991). According to linguistic theories (Chomsky & Halle, 1968; Liberman & Prince, 1977; Selkirk, 1986, 1995a, 1995b, 2000) and behavioral data (Breen & Clifton, 2011, 2013; Kelly & Bock, 1988), the ideal English sentence has a metrical rhythm created by an alternating pattern of stressed and unstressed

syllables. In the typical English sentence, weak function words (e.g., *the, will, her*) alternate with the stronger syllables of content words (e.g., *BOAT, TIGer, BUTterfly*; Selkirk, 1986, 2000) to create the canonical rhythm of English. According to Koriat, Kreiner, and Greenberg (2002), readers use morphosyntactic cues (e.g., distinguishing between function and content words) to establish the structural and prosodic frame of a sentence, prior to analyzing the meaning of the sentence.

Historically, the rhythm of English was regarded as the simple by-product of the linear arrangements of words with static stress points. In this view, stress was regarded as an intractable feature of syllables (i.e., [+ stress], [- stress]; Chomsky & Halle, 1968). For example, because the third syllable in the word *MissisSIPpi* is stress marked, that syllable receives the most prominence when pronounced. This view assumed that stress was fixed and predictable. Assuming that stress was invariant, traditional approaches to explain stress assignment in English looked for rules that could illuminate its lawfulness (e.g., Chomsky & Halle, 1968). In some languages, stress assignment is highly predictable. For example, in Spanish, words ending in a vowel, *-n*, or *-s* receive stress on the penultimate (next to the last) syllable, as in *espinacas* and *tomate* (Hualde & Nadeu, 2014). Exceptions to these rules are unequivocally marked in the orthography by diacritical marks (e.g., *bróculi, espárragos*; Hualde & Nadeu, 2014). In Hungarian, stress always appears on the first syllable in a word (Siptár & Törkenczy, 2007). In English, stress is less lawful. In response, educators outlined the regularities that exist, while presenting a lengthy list of exceptions (Fudge, 1984). Nonetheless, approximately 85% of content words begin with a strong stress on the first syllable (Cutler & Carter, 1987). This strong, statistical tendency is used by adult listeners (Cutler & Carter, 1987; Vroomen, Tuomainen, & de Gelder, 1998) and infant listeners (Jusczyk, 1997) when parsing spoken English. It is unknown whether this statistical tendency is used in silent reading.

In contemporary linguistics, stress is viewed as dynamic rather than as an immutable property of syllables. For example, stress in English can actively shift within a word when affixes are added, as when the word *POLitic* becomes *poLITical* (Liberman & Prince, 1977; Selkirk, 1986). Stress shifts can reflect subtle differences in meaning. Consider the difference among *BLUEbird* (a specific type of bird known for its blue plumage), "that's a **BLUE** bird" (the Sesame Street character, Big Bird, after a mishap with a can of blue spray paint), and "every blue **BIRD**" (any bird that is blue). Speakers say "**BOOK**" and "**BAG**" to refer to these specific items, yet they say "**BOOK**bag" to refer to a specific type of satchel carried by school-age children. In the case of heteronyms, stress is phonemic. The stress alternations in heteronyms yield different pronunciations with unique

meanings for identically spelled words (e.g., “fresh **PRODU**ce at the store,” “need to **proDUCE** more wid-gets”). Yet, the lack of predictability on the placement of English stress is not restricted to heteronyms.

Stress assignment in English reaches across word boundaries. The rhythm rule (Lieberman & Prince, 1977; Nespor & Vogel, 1986; Selkirk, 1986) is hypothesized to be the mechanism that maintains the ideal, stress-alternating rhythm in English. Beats in a sentence can be added, moved, or deleted to prevent stress clashes (sequences of adjacent, stressed syllables). For example, the last syllable in the word *Tennes**SEE*** receives stress. The word ***AIR*** receives stress. When spoken as the phrase “Tennes**SEE** + **AIR**,” the stress pattern in *Tennes**SEE*** is reversed, yielding “**TEN**nessee **AIR**” (Lieberman & Prince, 1977). Similarly, to avoid the stress clash of two, adjacent strong syllables, the words *thir**TEEN*** and ***MEN*** spoken in isolation become “**THIR**teen **MEN**” when spoken as a phrase (Selkirk, 1986). In short, the rhythm rule extends hierarchically within and across words to preserve the canonical, alternating meter in English.

Consistent with the rhythm rule (Lieberman & Prince, 1977; Selkirk, 1986), corpus studies of formal and conversational English, as well as of written and spoken English, have revealed a tendency toward rhythmic alternation within and across words. Kelly and Bock (1988) analyzed the rhythmic properties of a diverse set of spoken words from corpora, including parent–child interactions, trial testimony, college lectures, the Watergate tapes, and conversations between twins, couples, and adults. To understand the rhythmic properties of written text, the team analyzed corpora from Bartlett’s (1937) *Familiar Quotations*, a collection of phrases, passages, and proverbs from ancient to modern literature. The canonical rhythm of English was more likely to occur in written than spoken communication and in formal than conversational English. Providing converging experimental evidence, Kelly and Bock found that speakers adjusted the stress patterns of bisyllabic pseudowords (e.g., *colvane*) within sentences to create rhythmic alternation. Participants were more likely to place stress on the first syllable of pseudowords when they followed a word with weak stress (e.g., “Planes will **COL**vane pilots”). In contrast, when a word with strong stress preceded the pseudowords, the pseudowords were pronounced with a weak–strong pattern (e.g., “The pins **colVANE** balloons”). The speakers adjusted lexical stress in a range of pseudowords to create an alternating rhythm across an utterance.

Extending such findings to other languages, oral and silent readers of German made syntactic parsing decisions that preserved the language’s preferences for the alternation of strong and weak syllables (Kentner, 2012). Like English, German’s stress-alternating rhythm is not explicitly marked in its orthography. When unprepared

oral readers placed stress on an ambiguous German word, the placement of stress was influenced by the immediate rhythmic environment (i.e., the lexical stress of the following word). Specifically, these unprepared oral readers defaulted to avoiding stress clashes that spanned across words. In prepared reading conditions, in which oral readers had access to the disambiguating portion of the sentence before starting to read aloud, readers were unaffected by the immediate rhythmic environment and placed the stress to signal the appropriate meaning of the ambiguous word. Converging evidence was found by using eye-tracking measures, which revealed that unprepared, silent readers were incrementally building a stress-alternating prosody.

The rhythmic alternation biases observed in German oral readers (Kentner, 2012) are consistent with the construction–integration model of reading (Kintsch, 2005). In theory, reading comprehension is an ongoing, joint interaction between text-based information and the situation model formed as the reader integrates the text with relevant prior knowledge. Early in a sentence, letter decoding and word recognition draw heavily on information arriving from the sensory systems (i.e., bottom-up or data-driven processing). For example, recognizing words in a sentence activates word meanings in memory, and lexical stress may be stored as part of the words in the mental lexicon (Selkirk, 1986). By recruiting top-down processes (e.g., prior knowledge, beliefs, goals, expectations, word frequencies), the reader constructs an interpretation of the words and phrases on the page (Kintsch, 2005; Treiman, 2001). Unprepared readers appeared to be data driven as they iteratively construct a stress-alternating rhythm early in the sentence, whereas prepared readers adjusted the rhythm in light of relevant knowledge (Kentner, 2012).

Thus, how readers use stress to create the canonical rhythm of English when reading aloud is becoming well understood. Yet, do these same processes occur in silent reading? That is, do silent readers of English render stress and rhythm in their inner voices? The notion that silent reading is accompanied by a voice in the head is long-standing (Huey, 1908/1968). According to the implicit prosody hypothesis (Fodor, 1998), silent readers project a prosodic pattern onto a sentence to guide syntactic processing (Fodor, 2002). In theory, the projected prosody is presumed to be identical to the overt prosody for that sentence in a comparable context (Fodor, 2002). Consistent with the implicit prosody hypothesis, detailed phonological information is activated during silent reading (e.g., Abramson & Goldinger, 1997; Ashby & Clifton, 2005; McCutchen, Bell, France, & Perfetti, 1991; Van Orden & Kloos, 2005).

Emerging evidence suggests that silent readers represent in the inner voice the prosodic features of phrasing (Bader, 1998; Hwang & Schafer, 2009), lexical stress

(Ashby & Clifton, 2005; Ashby & Rayner, 2004), meter (Breen & Clifton, 2011, 2013), and focal stress (Gross, Millett, Bartek, Bredell, & Winegard, 2014), consistent with the implicit prosody hypothesis (Fodor, 2002). By observing eye movements, Ashby and Clifton found that polysyllabic words with two stressed syllables were read more slowly and received more fixations than did polysyllabic words with one stressed syllable. Longer reading times for words such as *FUNdaMENTal* compared with words such as *sigNIficant* are consistent with the longer pronunciation times for stressed compared with unstressed syllables in spoken English (Chomsky & Halle, 1968; Selkirk, 1986).

Breen and Clifton's (2011, 2013) silent readers formed expectations about the stress patterns of noun-verb heteronyms (e.g., *ABstract* [noun], *abSTRACT* [verb]) appearing in limericks and garden path sentences. Reading times were slower when expectations regarding prosody were not upheld. For example, reading times suffered when heteronyms appearing in garden path sentences were syntactically biased to have a noun interpretation (e.g., *ABstract*) yet had to be prosodically disambiguated as a verb (e.g., "The brilliant *abSTRACT* the..."). Compatibly, there were longer latencies in the limerick condition when the stress placement on the heteronym was inconsistent with the rhythm. For example, reading times were longer for "There once was a penniless *PEASant* [strong/weak] / who went to his master to *preSENT* [weak/strong]" than for "There once was a penniless *PEASant* [strong/weak] / who could not afford a nice *PRESent* [strong/weak]." Thus, silent readers formed a metrical representation that guided their syntactic representations.

In addition to lexical stress and meter, silent readers appear to stress-mark newsworthy content, consistent with the implicit prosody hypothesis. English speakers (Bolinger, 1978; Nava & Zubizarreta, 2010; Rooth, 1992; Schafer, Speer, Warren, & White, 2000) and expressive readers (Cooper, Eady, & Mueller, 1985; Eady & Cooper, 1986; Schwanenflugel, Westmoreland, & Benjamin, 2015) acoustically emphasize new or important content. Thus, Gross and colleagues (2014) tested how contextual newness might implicitly influence the inner voice. Silent readers were predicted to give higher helpfulness ratings for the final sentences of a paragraph when new or important content was emphatically marked and previously given information was not, compared with incongruously matched stimuli. As predicted, silent readers in experiment 1 preferred capital-emphasized, newsworthy content ("James stole the *BRACELET*") when the just read story left the reader wondering what was stolen. Readers preferred "*JAMES* stole the bracelet" when left wondering who the thief was. Experiment 2 generalized the findings to newsworthy function words and to a different behavioral measure, reaction

time. As predicted, "He *CAN*" was judged more quickly and accurately following "Can he swim?" whereas "*HE* can" was judged more quickly and accurately following "Who can swim?" The stylistic alterations used by Gross and colleagues to mark stress in print may bridge the bottom-up and top-down processes involved in reading (Kintsch, 2005; Kintsch & Mross, 1985; Rumelhart & McClelland, 1982). Moreover, the stylistic alterations that mark stress in print may be particularly helpful in the early part of sentences when bottom-up processes dominate (Kintsch, 2005).

The stylistic alterations used by Gross and colleagues (2014) could be used to test hypotheses about individual differences in rhythm and stress extraction during silent reading. There is emerging evidence for individual differences in the syntactic, and accompanying prosodic, analyses of sentences. Consider the sentence "The maid of the princess who scratched herself in public was terribly embarrassed." The key question is, Who scratched herself in public, the maid or the princess? English and Dutch readers with high-span working memory capacities chunked more information by attaching *scratching* to the maid after treating the entire subject of the sentence (up to the verb *was*) as a cohesive unit. English and Dutch readers with low-span working memory capacities attached *scratching* to the princess after forming chunks in the middle of the sentence (Swets, Desmet, Hambrick, & Ferreira, 2007). The ongoing syntactic and prosodic analyses performed during silent reading were more affected by the reader's working memory capacity than by the preferences of the reader's native language (Ferreira & Karimi, 2015; Swets et al., 2007).

In sum, spoken and written English have a canonical rhythm created by an alternating pattern of stressed and unstressed syllables. The rhythm of English is not simply the by-product of the concatenation of words with given stress points. Rather, the stress-alternating rhythm of English is dynamic and responsive to contextual conditions. Stress and rhythm are not marked in English's orthography. Skilled users of English must infer stress, and the stress-alternating preferences of its language, when writing and reading aloud. Emerging evidence reveals that silent readers render prosodic features in their inner voices. Extrapolating from linguistic theories of the stress-alternating preferences of spoken English (Chomsky & Halle, 1968; Fudge, 1984; Liberman & Prince, 1977; Selkirk, 1986, 1995a, 1995b, 2000), we evaluated whether stylistic alterations to print that mark rhythm (experiment 1) and lexical stress (experiment 2) help our silent readers hear the beat of English in their inner voices. We examined prosody sensitivity as an individual difference variable (Ferreira & Karimi, 2015; Swets et al., 2007). As previously discussed, Chan and Wade-Woolley (2016) found that after controlling for working memory, prosodic awareness

still accounted for adults' word reading. We tested hypotheses about what educational experiences promote greater awareness of rhythm and stress in written English. We speculated that greater exposure to poetry and better academic skill more generally would be linked to greater awareness of prosody.

Experiment 1

The primary goal of experiment 1 was to investigate whether silent readers represent rhythm when translating print to a speech-based code. The reading material in experiment 1 was formal poems, a form of writing based on rhythms. Poetic rhythm is created by the careful arrangements of stressed and unstressed syllables into patterns (Pinsky, 1988), whereby the words and the rhythm are inseparable. Consider the arrangement of stressed and unstressed syllables into a strong-weak rhythm in the poem "Peter, Peter, pumpkin-eater, / Had a wife and couldn't keep her; / He put her in a pumpkin shell / And there he kept her very well" (Opie & Opie, 1997, p. 410).

To mark the stress pulses in print, we changed the appearance of text in accordance with Norman's (1988) four design principles. A well-designed interface (1) exploits existing knowledge, (2) is perceptually obvious, (3) gives clues to facilitate easy interaction, and (4) eases the transfer of what is known to a new context (Norman, 1988). We borrowed the use of capital letters, bolding, and enlargement of text to convey stress from existing practices, such as pronunciation in dictionaries (e.g., *aioli* = /ahy-**OH**-lee/) and comic strips (e.g., "**OUCH!**"), meeting principle 1. These stylistic¹ alterations are easily distinguished from unaltered text (principle 2). The stylistic alterations give rise to the phenomenological impression of a raised voice when reading (e.g., "**CAN YOU HEAR ME NOW?**"; principle 3) and ease the reader's rendering of stress from memory (principle 4).

Under the guise of judging a poet's tinkering with ways to improve the pleasure of reading poetry, participants were asked to rate the overall helpfulness of the stylistic alterations placed on the last line of the poems. The two- and four-line poems began in plain text (e.g., "'Twas brillig, and the slithy toves"). In the final line of each poem, the stylistic alterations either congruently mapped onto the stress pulses (e.g., "Will **BE** a **TOT**ter'd **WEED** of **SMALL** worth **HELD**"; congruent condition) or incongruently mapped onto unstressed syllables (e.g., "**WILL** be **A** tot**TER**'D weed **OF** small **WORTH** held"; incongruent condition).²

To examine the generality of prosodic rhythm effects in silent reading, three meters were used (trochaic, iambic, and anapestic) to mirror the undulating rhythms of spoken English (one or two weakly stressed syllables alternating with strongly stressed syllables; Liberman & Prince, 1977; Selkirk, 2000). The poem about pumpkin-fond Peter is written in a trochaic meter, whereby a stressed syllable is followed by an unstressed syllable ("**P**eter, **P**eter, **P**umpkin **E**ater"). An iambic meter consists of an unstressed syllable followed by a stressed syllable (e.g., "The **M**AN is **S**MALL"). An anapestic meter consists of two unstressed syllables followed by one stressed syllable (e.g., "And the **H**OUSE is the **P**LACE"). Extrapolating from linguistic theories of the regular stress patterning in English (Chomsky & Halle, 1968; Liberman & Prince, 1977; Selkirk, 1986, 1995a, 1995b, 2000) and behavioral data (Breen & Clifton, 2011, 2013; Kelly & Bock, 1988; Kentner, 2012), we predicted that silent readers would rate the stylistic alterations appearing on the last line of the poem as more helpful when it marked the rhythm (stress pulses) of the poem, compared with the dissonant, incongruous condition (for sample stimuli, see Table 1; for the complete list, see Appendix A; all of the appendixes are available as supporting information for the online version of this article).

TABLE 1
Sample Poetry Stimuli in Experiment

| Trochaic meter example | Iambic meter example | Anapestic meter example |
|---|--|--|
| "Piping down the valleys wild" | "'Twas brillig, and the slithy toves" | "Fled foam underneath us, and around us, a wandering and milky smoke" |
| " PIP ing SONGS of PLEAS ant MUS ic" (beat congruous) | "Did GYRE and GIM ble IN the WABE " (beat congruous) | "As high as the saddle-girth, covering away from our glances the tide" |
| "Pip ING songs OF pleas ANT mu SIC " (beat incongruous) | " DID gyre AND gim BLE in THE wabe" (beat incongruous) | "And those that fled, and that followed, from the foam-pale distance broke" |
| | | "The im MORT al de SIRE of im MORT als we SAW in their FACES , and SIGHED " (beat congruous) |
| | | "The im MORT al de ES ire of im MORT als WE saw in THEIR faces, AND sighed" (beat incongruous) |

The second goal was to examine prosody sensitivity as an individual difference variable. As previously discussed, individual differences in prosody sensitivity were related to differences in working memory capacity (Chan & Wade-Woolley, 2016; Ferreira & Karimi, 2015; Swets et al., 2007). Silent readers with less working memory capacity were found to be more likely to pause midsentence, and these breaks influenced the ongoing prosodic analyses (Ferreira & Karimi, 2015). After controlling for differences in executive function, prosodic awareness accounted for individual differences in adults' word-reading abilities (Chan & Wade-Woolley, 2016). Therefore, we reasoned that college readers with a refined, prosodic inner voice as measured by our experimental task would have higher scores on measures used to forecast academic achievement in college. We further reasoned that those who had more exposure to poetry would have a more refined awareness of prosody and better memory for our poetry selections.

Method

Participants

Fifty-eight students enrolled in introductory psychology courses at a public university in the Great Lakes region of the United States received course credit for their participation. Fifty-seven participants were native English speakers, and one participant was a native Spanish speaker. Seven participants reported fluency in languages in addition to English (four in Spanish, one in French, one in Arabic, and one in American Sign Language).

Stimuli

Twenty-four experimental stimuli were excerpts from published poetry (20) or poems written by the investigators (four). The stimuli included eight poems for each rhythm (iambic, trochaic, and anapestic). For each rhythm, four poems were composed of two lines, and four poems were composed of four lines. In formal poetry, there can be inconsistencies in the beat patterns (Attridge, 1995). Such rhythm irregularities were more common in our excerpts from longer poems than shorter ones. We allowed up to two deviations from the predicted meter in each poem. Only the final line of a poem was stylistically altered to be either congruous or incongruous with the poem's meter. The number of stylistically altered syllables was held constant across both the congruous and incongruous versions of a poem. Syllabification and stress assignment were verified via Dictionary.com. Participants read only one version of each poem, determined randomly, while satisfying the constraint of an equal number of congruous and incongruous trials. To compute prosody sensitivity, we subtracted each subject's

average ratings of incongruent trials from ratings of congruent trials.

Design

The study was a $3 \times 2 \times 2$ experimental design with meter type (trochaic, iambic, or anapestic), length of poem (two or four lines), and congruous versus incongruous mapping of stylistic alterations onto the implicit meter of the poems. Helpfulness ratings were the dependent variable. For the three independent variables, levels were measured within participants and stimuli.

Measures

Seven forced-choice primer trials were designed to familiarize participants with the idea of stylistically marked stress by asking them to decide which of two usages of capitalization was better (e.g., "*LABOR/laBOR*," "Sam needed to practice shooting his bow with greater *acCURacy/ACcuracy*"). No performance feedback was furnished at any time during the experiment.

The poetry memory test comprised 24 multiple-choice questions, with one question per poem. Our test should not be construed as a typical reading comprehension test. Our poetry memory test was notably challenging because it probed participants' verbatim memories for specific words or phrases. Our selection of poems did not necessarily present a complete idea or coherent message, because the poems were chosen on the basis of their meter rather than content. Some of the poetry selections may have been familiar to the readers, yet many selections were likely unfamiliar to them (see Appendix B).

As a proxy of prior exposure to poetry, we created a poet recognition test. Like its inspiration, the author recognition test (Stanovich & West, 1989), the poet recognition test is a checklist of 60 names. Thirty-two of the names were targets (famous poets), and 28 were foils (culturally salient figures, such as actors, political figures, and comedians). Participants were instructed to identify the poets from nonpoets and warned that guessing could be easily detected. Our poet recognition test, like the author recognition test, was scored as the number of correct identification of authors (hits) minus the number of incorrect identification of foils (see Appendix C).

The demographic questionnaire asked participants to report their score on a U.S. college standardized admissions screening test known as the ACT. ACT scores served as a proxy of individual differences in academic achievement. The demographic questionnaire also asked participants to report their cumulative grade point average (GPA), and the number of poems they read in the last year as an index of exposure to poetry. The number of poems read in the last year was coded as 1 for *10 or fewer*, 2 for *11–20*, 3 for *21–30*, and 4 for *31 or more*.

Procedure

A story introduced the task. Participants were told that a creative, prize-winning poet decided to stylistically highlight parts of words to enhance the reading of poetry, yet it took a little bit of tinkering for the author to decide where to use this new technique. For example, the new technique seemed to work in “The ITsy BITsy SPIder...” In contrast, the new technique did not seem to work in “HumpTY DumpTY sat ON...” The finished book featuring the new technique to enhance the reading experience was sent to the editor. Yet, during electronic transit, a computer virus corrupted the book by inserting random stylistic changes to the poems. Playing the role of the editor, the participants’ job was to sort out the mess by judging the virus-corrupted, unhelpful alterations from the intentional, helpful alterations of the poems by using the 5-point scale with the endpoints labeled “unhelpful” and “helpful.” Participants were cautioned to read all poems carefully in preparation for the poetry memory test at the end of the session.

After consenting to participate and reading the cover story, participants completed the tasks in the following order: the primer trials, the poems, the poet recognition test, the poetry memory test, and the demographic questionnaire.

Results and Discussion

Our findings are consistent with the experimental hypothesis derived from linguistic theory: Stylistic alterations of the text that congruously mapped onto the stress-alternating rhythm of poems were rated as more helpful, compared with poems in the incongruously mapped condition, $F(1, 57) = 83.5, p < .0001, \eta_p^2 = .23$ by subjects, and $F(1, 18) = 112.7, p < .0001, \eta_p^2 = .84$ by items. There were no main effects for type of meter or poem length and no interactions involving any factors (see Table 2).

Using Pearson correlational analyses, we explored the relations between performance on the experimental task (a measure of prosody sensitivity) and the poet

TABLE 2
Descriptive Statistics for Helpfulness Ratings for Poetry Stimuli in Congruous and Incongruous Conditions in Experiment 1

| Incongruous | | | | Congruous | | | |
|------------------------|------|------------|------|--|------|------------|------|
| By subjects | | By stimuli | | By subjects | | By stimuli | |
| M | SD | M | SD | M | SD | M | SD |
| <i>Trochaic meter</i> | | | | <i>Example: “PIPing SONGS of PLEASant MUSIC”</i> | | | |
| <i>Two-line poems</i> | | | | <i>Example: “Did GYRE and GIMble IN the WABE”</i> | | | |
| 2.28 | 1.27 | 2.60 | 0.32 | 4.00 | 0.89 | 3.96 | 0.38 |
| <i>Four-line poems</i> | | | | | | | |
| 2.56 | 1.21 | 2.56 | 0.30 | 3.58 | 1.13 | 3.63 | 0.40 |
| <i>Iambic meter</i> | | | | <i>Example: “The imMORTal deSIRE of imMORTals”</i> | | | |
| <i>Two-line poems</i> | | | | | | | |
| 2.51 | 1.06 | 2.52 | 0.36 | 3.57 | 1.04 | 3.49 | 0.26 |
| <i>Four-line poems</i> | | | | | | | |
| 2.50 | 1.07 | 2.49 | 0.19 | 3.20 | 1.46 | 3.66 | 0.36 |
| <i>Anapestic meter</i> | | | | | | | |
| <i>Two-line poems</i> | | | | | | | |
| 2.56 | 1.11 | 2.71 | 0.33 | 3.64 | 1.03 | 3.53 | 0.70 |
| <i>Four-line poems</i> | | | | | | | |
| 2.31 | 1.08 | 2.39 | 0.32 | 3.63 | 0.93 | 3.59 | 0.40 |

Note. The minimum for each is 1, and the maximum for each is 5.

recognition test, the frequency of reading poems, the poetry memory test, self-reported ACT scores, and self-reported cumulative GPA (see Table 3). Our correlational analyses were likely underpowered because of the small sample size ($n = 58$).

We hypothesized that exposure to poetry would cultivate a refined awareness of prosody. Consistent with this prediction, participants with greater prosody sensitivity read more poems in the last year and had higher GPAs. Suggesting that our poet recognition test has construct validity, those who read more poems in the last year performed better on the poet recognition test and the poetry memory test. Moreover, performing better on the poet recognition test was correlated with performing better on the poetry memory test. Higher ACT scores were correlated with better performance on the poet recognition test, consistent with the findings of Stanovich and West (1989). Prosody sensitivity, unexpectedly, was not related to higher ACT scores. Because of the small sample size, the results of the correlational analyses should be viewed as tentative.

In summary, as predicted from linguistic theories (Chomsky & Halle, 1968; Liberman & Prince, 1977; Selkirk, 1986, 1995a, 1995b, 2000) and behavioral data (Breen & Clifton, 2011, 2013; Kelly & Bock, 1988; Kentner, 2012), when the stylistic alterations mapped onto the beat of the poems, participants rated these alterations as more helpful, compared with alterations in the incongruous condition. These findings robustly generalized across all three meters. Correlation analyses revealed that a refined sensitivity to prosody was correlated with reading more poems in the last year and higher GPAs. Reading more poems in the last year was also related to better performance on the poet recognition test and the poetry

memory test. Recognizing more poets hidden among foils was correlated with better performance on the poetry memory test and higher ACT scores.

Experiment 2

The first goal of experiment 2 was to investigate whether adult readers represent lexical stress in their inner voices when translating print to a speech-based code. As previously discussed, third graders showed an understanding of lexical prosody when reading aloud (Schwanenflugel & Benjamin, 2016). Under the guise of helping an author's tinkering with ways to improve the pleasure of reading prose, participants were asked to judge the helpfulness of the stylistic alterations. In the congruous condition, the stylistic alterations of print mapped onto the stressed syllables in the heteronyms (e.g., "the science **PRO**JECT," "pro**J**ECT the film"). In the incongruous condition, the stylistic alterations mapped onto the unstressed syllables in the heteronyms (e.g., "the science pro**J**ECT," "**PRO**ject the film"). In theory, helpfulness ratings should be higher when the stylistic alterations mapped onto the stressed syllables in the heteronyms (e.g., "To con**T**EST his grade, Sam questioned his instructor about a response scored incorrectly"), compared with the incongruous condition (e.g., "To generate a reputation as a re**B**EL, Mark added scandalous content to his novel"). As suggested by these examples, the stylistic alterations appeared as bolded capitals in experiment 2.

Because of the statistical tendency for stress to appear word-initially (Cutler & Carter, 1987), our readers' inner voices may be tuned to the prosodic nuances of the first syllable (e.g., "fresh **P**ROduce") more than the

TABLE 3
Correlations in Experiment 1

| Variable | 1 | 2 | 3 | 4 | 5 | 6 |
|---|-------|-------|------|------|------|------|
| 1. Prosody sensitivity | – | | | | | |
| 2. Poems read last year | .28* | – | | | | |
| 3. Poetry memory test | .18 | .35* | – | | | |
| 4. Poet recognition test (Hits – Foils) | .14 | .36** | .32* | – | | |
| 5. ACT scores | –.06 | .11 | –.01 | .29* | – | |
| 6. GPA | –.30* | –.17 | –.14 | –.11 | .20 | – |
| <i>M</i> | 1.17 | 1.21 | 13.4 | 6.78 | 24.9 | 3.10 |
| <i>SD</i> | 1.07 | 1.0 | 3.09 | 3.40 | 3.35 | 0.44 |
| Minimum | –2.0 | 0 | 4 | 0 | 20 | 2 |
| Maximum | 3.50 | 4 | 19 | 16 | 34 | 3.89 |

* $p < .05$. ** $p < .001$.

second syllable (e.g., “pro**DUCE** steam”). Therefore, our second goal was to evaluate this tuning difference by assessing whether silent readers gave a larger rating difference between the congruent and incongruent conditions when making judgments about the first syllable compared with the second syllable in the bisyllabic heteronyms.

Our third goal was to explore whether prosody sensitivity interacted with the bottom-up and top-down processes in text comprehension (Kintsch, 2005; Kintsch & Mross, 1985; Rumelhart & McClelland, 1982). Kentner (2012) found that early in a sentence, unprepared readers were data driven as they iteratively constructed a stress-alternating rhythm. In our experiment, participants read stress-alternating heteronyms appearing in sentences without the benefit of prior context. Thus, we reasoned that our readers’ inner voices would be more tuned to the benefit afforded by stylistic alterations of text in the early part of sentences when bottom-up processes dominate, compared with the later part of sentences when top-down processes dominate. Thus, the difference between ratings in the congruent and incongruent conditions should be larger early in a sentence compared with late in the sentence.

The fourth goal of experiment 2 was to replicate and extend the findings of individual differences in prosody sensitivity. Correlational analyses examined relations between prosody sensitivity and other individual differences, including exposure to poetry and measures predictive of academic achievement (i.e., reading comprehension, ACT scores, vocabulary scores). Extrapolating from Schwanenflugel and Benjamin’s (2016) finding that children with good lexical prosody had good reading comprehension abilities, we reasoned that our college readers with a refined inner voice for the stress

alternations in heteronyms would have higher reading comprehension scores and higher scores on the tests used to forecast academic achievement, when compared with students who performed less well on our experimental task. As revealed in experiment 1, we speculated that prosody sensitivity would be related to exposure to poetry.

The research team conducted two replications (studies 1 and 2) of the experiment to check for the robustness of the findings.

Method

Participants

All enrolled in introductory psychology courses at a public university in the Great Lakes region of the United States, 482 students received course credit for their participation. Seventeen participants were late speakers of English.

Stimuli

The experimental stimuli were 20 bisyllabic heteronyms appearing in sentences. Across the experimental conditions, care was taken to control for factors that could influence the reading difficulty of the sentences. Thus, the sentences in which heteronyms occurred were matched on number of syllables in a sentence, number of words in a sentence, and reading level (see Table 4). The Flesch–Kincaid grade level (Kincaid, Fishburne, Rogers, & Chissom, 1975) built in to Microsoft Office Word was used to determine the reading level. This measure renders a reading level score to indicate the appropriateness of the text for specific grade levels in the United States. Thus, a score of 8

TABLE 4
Sample of a Heteronym’s Sentences Matched on Words, Syllables, and Reading Level Across the Experimental Conditions of Syllable Location (First or Second), Place in the Sentence (Early or Late), and Stress Marking (Congruous or Incongruous)

| Heteronym: <i>conduct</i> (15 words, 28 syllables, reading level of 12.2) | |
|---|---|
| First | Second |
| <i>Early</i> | |
| <i>Congruous</i> : “Because of their CON duct, Robert and Timothy are obligated to perform community service this week.” | <i>Congruous</i> : “Because of the leader’s failure to properly con DUCT, the violinists trailed behind the percussion section.” |
| <i>Incongruous</i> : “Because of their con DUCT, Robert and Timothy are obligated to perform community service this week.” | <i>Incongruous</i> : “Because of the leader’s failure to properly CON duct, the violinists trailed behind the percussion section.” |
| <i>Late</i> | |
| <i>Congruous</i> : “Robert and Timothy are obligated to perform community service this week because of their CON duct.” | <i>Congruous</i> : “The violinists trailed behind the percussion section because of the leader’s failure to properly con DUCT.” |
| <i>Incongruous</i> : “Robert and Timothy are obligated to perform community service this week because of their con DUCT.” | <i>Incongruous</i> : “The violinists trailed behind the percussion section because of the leader’s failure to properly CON duct.” |

indicates that the text is appropriate for eighth-grade readers and above. The reading levels of the experimental stimuli on the Flesch–Kincaid grade-level test ranged from 9.9 to 12.6.

Design

A partial Latin square design was used to place the 20 heteronyms into the $2 \times 2 \times 2$ factorial design with the independent variables of (1) syllable location (first or second in the heteronym), (2) place in sentence (early or late), and (3) stress marking (congruous or incongruous mapping). For the three independent variables, levels were measured within participants and stimuli.

Table 5 shows how the experimental conditions were distributed across studies 1 and 2 and sessions 1 and 2. The design of the experiment was constrained by the small number (i.e., 20) of bisyllabic heteronyms available for study. Twenty stress-alternating heteronyms yielded 40 stimuli. Both meanings of each heteronym were presented in each replication study, with only one meaning presented at each session, to minimize effects of prior exposure. For example, if *pro' duce* (noun) appeared at session 1, then *produce'* (verb)

TABLE 5
The Schedule of How the Experimental Conditions of Syllable Location (First or Second), Place in the Sentence (Early or Late), and Stress Marking (Congruous or Incongruous) Were Distributed Across Studies and Sessions

| Study | Session | Conditions |
|-------|---------|----------------------------|
| 1 | 1 | First, early, congruous |
| | | Second, late, incongruous |
| | | First, early, incongruous |
| | | Second, late, congruous |
| | 2 | Second, early, incongruous |
| | | First, late, congruous |
| | | Second, early, congruous |
| | | First, late, incongruous |
| 2 | 1 | Second, early, incongruous |
| | | First, late, congruous |
| | | First, early, incongruous |
| | | Second, late, congruous |
| | 2 | First, early, congruous |
| | | Second, late, incongruous |
| | | Second, early, congruous |
| | | First, late, incongruous |

appeared at session 2. A fully crossed design contains eight experimental conditions, and in both experiments, all eight experimental conditions were represented. Yet, at each session, only four of the eight conditions were represented. Five stimuli were presented in each of the eight conditions. Heteronyms were randomly assigned to experimental conditions at session 1. Because of the constraints just described, the assignment of a stimulus in session 1 constrained where the stimulus was placed in session 2. The pairwise comparisons for three independent variables (syllable location, place in sentence, and stress marking) had 10 stimuli, resulting in nine degrees of freedom. Thus, in the analyses reported here, there is some unavoidable confounding of words with conditions (for a complete list of stimuli, see Appendix D).

Students participated in the experiments on two occasions (sessions 1 and 2), separated by several months. The substantial delay between sessions was chosen to minimize the impact of judgments made at session 1 on judgments made at session 2. Temporal separation has been shown to reduce carryover effects and increase the independence of guesses (Vul & Pashler, 2008).

Participants rated the stylistic alterations on a 7-point scale, with 7 as *helpful* and 1 as *unhelpful*. The change from a 5-point scale in experiment 1 to a 7-point scale in experiment 2 was motivated by the presumption that judging stress alternations in heteronyms would be more difficult than judging stress alternations in poems. The 7-point scale gave participants a greater range to express their judgments of helpfulness for the more difficult task.

Measures

A 20-item multiple-choice reading comprehension test was administered. The questions were based on the heteronym sentences, with the correct answer as one of the four options (see Appendix E). The reading comprehension test was designed to measure higher level comprehension rather than simple, rote recognition. To minimize rote recognition, the comprehension questions mostly contained synonyms of the heteronyms in lieu of the original words.

As in experiment 1, the poet recognition test was administered, and participants were asked to indicate the names of poets hidden among foils.

To measure individual differences in academic achievement, we selected two 18-item true/false vocabulary tests from MyVocabulary.com. The tests can be used to prepare for the SAT, which is a standardized test commonly used in college admissions in the United States. MyVocabulary.com grants permission to use the vocabulary tests for educational purposes. On the tests, participants had to decide whether each vocabulary

word was used properly in the sentence (e.g., “Mark needs to legally **PREVARICATE** when he signs documents before a notary,” “Tom gave his uncle a **FALLACIOUS** statement when he yawned nine times in a row but said he had slept well”; see Appendix F).

The demographic questionnaire asked participants to report whether English was their first language, and their scores on the ACT (a widely used measure of academic ability).

Procedure

The story that introduced the experimental task in experiment 1 was modified for experiment 2. Participants were told that a prize-winning author had decided to try a new technique in her latest book. The author believed that **CAP**-highlighting parts of the text might enhance the reading experience, yet it took a little bit of tinkering for the author to decide where to use this new technique. The use of capital letters in “It’s a **BIRD**, it’s a **PLANE**, it’s **SUPERMAN!**” seemed to work. In contrast, the use of caps in “Get **OUT OF** my way!” did not seem to work. The finished book featuring the new technique to enhance the reading experience was sent to the editor. Yet, during electronic transit, the book was corrupted by a virus that inserted nonintentional instances of capitals. Playing the role of the editor, participants were asked to sort out the mess by judging virus-corrupted, nonintentional caps from intentional, helpful caps. Participants were told that they would complete a reading comprehension test based on the author’s book at the end of the

session. No performance feedback was furnished at any time.

Similar to experiment 1, nine forced-choice training stimuli were used to familiarize participants with the stylistic alterations by asking participants to decide which one of the two usages was better.

At session 1, during the initial weeks of the semester, students consented to participate in the experiment, read a cover story, completed the forced-choice training stimuli, judged the helpfulness of the placement of the bolded caps in 20 heteronyms in sentences, and recognized the names of poets hidden among foils on the poet recognition test. At session 2, students consented to participate in the experiment, read a cover story, completed the forced-choice training stimuli, judged the helpfulness of the placement of the bolded caps in 20 heteronyms in sentences, completed the reading comprehension test, completed the SAT vocabulary tests, and answered a brief demographic questionnaire.

Results and Discussion

The results reported here combine the data sets from the two replications (studies 1 and 2). Consistent with linguistic theory, there was a main effect of congruency, $F(1, 481) = 157.6, p < .0001, \eta_p^2 = .25$ by subjects, and $F(1, 9) = 247.8, p < .0001, \eta_p^2 = .97$ by items. When there was a match between the stylistic alterations and the stress pulse in the heteronym, the stylistic alterations were viewed as more helpful (mean [M] = 4.4, standard deviation [SD] = 0.8), compared with the mismatched stimuli ($M = 3.7, SD = 1.0$; see Table 6).

TABLE 6
Descriptive Statistics for Heteronym Stimuli in Congruous and Incongruous Conditions in Experiment 2

| Experimental condition | By subjects | | | | By stimuli | | | |
|--|-------------|-----------|---------|---------|------------|-----------|---------|---------|
| | <i>M</i> | <i>SD</i> | Minimum | Maximum | <i>M</i> | <i>SD</i> | Minimum | Maximum |
| <i>Congruous condition</i> | | | | | | | | |
| Early in sentence/first-syllable stress | 4.6 | 1.2 | 1 | 7 | 4.6 | 0.3 | 4.2 | 5.3 |
| Early in sentence/second-syllable stress | 4.4 | 1.3 | 1 | 7 | 4.4 | 0.2 | 4.1 | 4.7 |
| Late in sentence/first-syllable stress | 4.5 | 1.1 | 1 | 7 | 4.5 | 0.3 | 4.0 | 5.0 |
| Late in sentence/second-syllable stress | 4.2 | 0.9 | 1 | 7 | 4.1 | 0.2 | 3.7 | 4.4 |
| <i>Incongruous condition</i> | | | | | | | | |
| Early in sentence/first-syllable stress | 3.5 | 1.2 | 1 | 7 | 3.5 | 0.2 | 3.2 | 3.8 |
| Early in sentence/second-syllable stress | 4.1 | 1.3 | 1 | 7 | 4.1 | 0.2 | 3.8 | 4.4 |
| Late in sentence/first-syllable stress | 3.5 | 1.3 | 1 | 7 | 3.5 | 0.3 | 3.0 | 3.9 |
| Late in sentence/second-syllable stress | 3.9 | 1.3 | 1 | 7 | 3.9 | 0.4 | 3.3 | 4.7 |

Because most English words have a stressed, initial syllable (Cutler & Carter, 1987; Vroomen et al., 1998), we expected larger differences in helpfulness ratings between congruous and incongruous conditions for the first syllable (e.g., “demanding lecture **CON**tent”) than for the second syllable (e.g., “Amy feels con**TENT**”). As predicted, there was a main effect of syllable location (first vs. second) in the heteronyms, $F(1, 481) = 37.5, p < .0001, \eta_p^2 = .07$ by subjects, and $F(1, 9) = 9.4, p < .01, \eta_p^2 = .51$ by items; and an interaction between stress marking and syllable location (first vs. second), $F(1, 481) = 93.8, p < .0001, \eta_p^2 = .16$ by subjects, and $F(1, 9) = 15.9, p < .01, \eta_p^2 = .64$ by items. Readers rated the congruous, stylistic alterations as more helpful when they occurred in the first syllable ($\mu = 4.5, \sigma = 1.0$) than when they occurred in the second syllable ($\mu = 4.3, \sigma = 0.9$), $F(1, 481) = 29.7, p < .01$ by subjects, and $F(1, 9) = 8.1, p < .05$ by items. This tuning of the inner voice robustly generalized to the mismatched conditions. Our readers gave significantly lower ratings of helpfulness when the incongruous mapping occurred in the first syllable ($\mu = 3.5, \sigma = 1.1$) than in the second ($\mu = 4.0, \sigma = 1.1$), $F(1, 481) = 125.5, p < .0001$ by subjects, and $F(1, 9) = 20.9, p < .001$ by items.

Extrapolating from the construction–integration model of reading (Kintsch, 2005), we reasoned that our silent readers’ inner voices would be more tuned to the impact of the stylistic alterations when placed early in the sentence when bottom-up processes dominate, compared with late in the sentence when top-down processes dominate. The interaction between place in the sentence (early or late) and syllable location (first vs. second) provided support for our prediction, $F(1, 481) = 6.2, p < .01, \eta_p^2 = .01$ by subjects, and $F(1, 9) = 5.5, p < .05, \eta_p^2 = .38$ by items. The main effect of place in the sentence was significant in the by-subjects analyses,

$F(1, 481) = 12.1, p < .001, \eta_p^2 = .02$; and marginally significant in the by-items analyses, $F(1, 9) = 3.8, p = .08, \eta_p^2 = .30$. Planned comparisons revealed that stylistically marked lexical stress early in the sentence was rated as more helpful ($\mu = 4.5, \sigma = 1.1$) than when late in the sentence ($\mu = 4.35, \sigma = 0.8$), $F(1, 481) = 7.4, p < .01$ by subjects, and $F(1, 9) = 6.7, p < .05$ by items. Yet, our readers rated the incongruous conditions equally unhelpful in the early ($\mu = 3.7, \sigma = 1.1$) and late in the sentence placements ($\mu = 3.8, \sigma = 1.0$), $F(1, 481) = 7.4, p < .01$ by subjects, and $F(1, 9), p < .30$ by items. Perhaps not surprising, stress-inducing text alterations that mapped onto unstressed syllables were consistently viewed as unhelpful regardless of where they appeared in the sentence.

A three-way interaction among syllable location, place in sentence, and stress marking was significant by subjects, $F(1, 481) = 8.8, p < .05, \eta_p^2 = .02$; but not by items ($p = .7$). Therefore, it was not analyzed further.³

Correlational analyses explored the relation between prosody sensitivity (performance on our experimental task) and the other measures that commonly correlate with academic achievement (self-reported ACT score, SAT vocabulary test, and reading comprehension). Prosody sensitivity was calculated separately for the two experimental sessions (1 and 2). As predicted, students with a refined awareness of prosody had higher reading comprehension scores, higher self-reported ACT scores, and higher SAT vocabulary scores, when compared with students with less refined awareness (see Table 7).

Replicating study 1, performance on the poet recognition test was related to prosody sensitivity measured at both sessions. Additionally, performance on the poet recognition test was strongly related to vocabulary skills and reading comprehension.

TABLE 7
Correlations in Experiment 2

| Variable | 1 | 2 | 3 | 4 | 5 | 6 |
|--|-------|-------|-------|-------|-------|------|
| 1. Prosody sensitivity at session 1 | – | | | | | |
| 2. Poet recognition test at session 1 | .19** | – | | | | |
| 3. Reading comprehension test at session 2 | .16** | .21** | – | | | |
| 4. SAT vocabulary test at session 2 | .24** | .26** | .38** | – | | |
| 5. ACT score at session 2 | .30** | .24** | .19** | .35** | – | |
| 6. Prosody sensitivity at session 2 | .63** | .22** | .31** | .28** | .34** | – |
| <i>M</i> | 0.67 | 8.9 | 10.4 | 25.4 | 24.0 | 0.70 |
| <i>SD</i> | 1.2 | 5.7 | 4.0 | 4.7 | 3.7 | 1.6 |
| Minimum | –2.1 | –6.0 | 4 | 8 | 15 | –3.5 |
| Maximum | 6.0 | 32 | 20 | 36 | 35 | 6.0 |

** $p < .001$.

We evaluated whether performance on the poet recognition test was uniquely related to prosody sensitivity when measures of individual differences in academic achievement were controlled statistically. After controlling for self-reported ACT scores ($\beta = 0.21, p < .0001$), reading comprehension ($\beta = 0.08, p = .10$), and SAT vocabulary scores ($\beta = 0.15, p < .01$), the poet recognition test still accounted for a significant variance in prosody sensitivity ($\beta = 0.09, R^2 \text{ change} = .01, p < .05$) at session 1. The full regression equation at time 1 was significant, with an R^2 of .127, $F(4, 458) = 17.7, p < .0001$. At session 2, after controlling for self-reported ACT scores ($\beta = 0.24, p < .0001$), reading comprehension ($\beta = 0.22, p < .0001$), and SAT vocabulary scores ($\beta = 0.10, p < .05$), the poet recognition test marginally significantly accounted for variance in prosody sensitivity ($\beta = 0.09, R^2 \text{ change} = .01, p = .052$). The full regression equation was significant, with an R^2 of .201, $F(4, 458) = 28.7, p < .0001$. Thus, we found some support that exposure to poetry may cultivate an inner voice for the rhythm of English, including the stress-alternating rhythm in heteronyms.

General Discussion

The canonical rhythm of English is created by the succession of one or two weakly stressed syllables alternating with strongly stressed syllables (Lieberman & Prince, 1977; Selkirk, 2000). Infant-directed speech is an exaggerated example of the singsong prosody of the English language (Bryant & Barrett, 2007). Although the English writing system is a printed representation of its spoken language, the orthography of English does not explicitly represent the prosodic features of rhythm and stress (Fudge, 1984; Treiman & Kessler, 2005). When confronted with the paucity of cues, the expressive reader must successfully render these prosodic features from memory informed by context. Just as English speakers and fluent oral readers hear a beat created by alternations between stressed and unstressed syllables, we investigated whether silent readers preserved this rhythm in their inner voices.

Extrapolating from linguistic theories of the stress-alternating preferences of spoken English (Chomsky & Halle, 1968; Fudge, 1984; Lieberman & Prince, 1977; Selkirk, 1986, 1995a, 1995b, 2000), we evaluated whether stylistic alterations to print that mark rhythm (experiment 1) and lexical stress (experiment 2) could help our silent readers hear the beat of English in their inner voices. To saliently, and intuitively, mark stress pulses in print, we consulted Norman's (1988) interface design principles before deciding on capital letters, bolding, and enlargement. In experiment 1, we focused on alternating rhythms of

English that span phrases and sentences across three poetic meters (iambic, trochaic, and anapestic). In experiment 2, we focused on the stress alternations of English that subtly occur within bisyllabic heteronyms. We suspected that perceiving the local stress alternation within one word would be more challenging than discerning meter across at least two lines of text. In both experiments, when the stylistic alterations mapped onto the beat of the poems (experiment 1) and the lexically stressed syllable in the heteronyms (experiment 2), participants rated these alterations as more helpful, compared with the incongruous conditions. Our findings suggest that marking stress explicitly in written English helped readers render a rhythmic inner voice.

The results of experiment 2 offer some additional hints about the potential impact of marking stress explicitly in written English. We predicted, and found, that our readers' inner voices were more tuned to the prosodic nuances in the first syllable (e.g., "SUBject of a research paper") compared with the second syllable (e.g., "I must obJECT"). This tuning difference likely results from the fact that most English words have a stressed, initial syllable (Cutler & Carter, 1987; Vroomen et al., 1998). Additionally, consistent with the construction-integration model of reading (Kintsch, 2005), silent readers' inner voices were more attuned to the impact of the stylistic alterations that congruously marked syllable stress early in the sentences when bottom-up processes dominate, compared with late in the sentence when top-down processes dominate. Our findings complement Kentner's (2012) findings that stress-alternating prosody is incrementally built in German readers. A reviewer of an earlier draft of this article suggested an alternative explanation for our findings: Prosody sensitivity lessens as a sentence unfolds because spoken prosody (pitch and volume) diminishes as the speaker runs out of air. Of course, this speech mechanism would have to generalize to silent reading.

The current experiments also revealed insights into individual differences in prosody sensitivity in that the canonical rhythm of written English was better detected by some readers than by others. Exposure to poetry forecasted a refined sensitivity to prosody in both experiments. Exposure to the rhythms in formal poetry may promote a polished inner voice for the stress-alternating preferences in English, including the dynamic rhythm of heteronyms. In experiment 2, the poet recognition test predicted participants' sensitivity to prosody, even after controlling for other strong predictors of prosody sensitivity, such as indicators of academic achievement. Of course, there are alternative explanations. Individuals with a cultivated sensitivity to the rhythm of English may be drawn to poetry. The

underlying mechanism for these correlates could be individual differences in working memory (Chan & Wade-Woolley, 2016), which was not measured in our studies. Correlational data do not reveal causal mechanisms, nor do they reveal the influence of third variables (e.g., volume of reading, reading acumen).

Our poet recognition test could potentially serve as a quick screen for individual differences in exposure to poetry. We modeled our test after the author recognition test (Stanovich & West, 1989), which is a measure of print exposure that is less vulnerable to social desirability effects than is asking participants about their reading habits. Stanovich and West found that performance on the author recognition test predicted word-reading skills in adults, after controlling for other predictors of word recognition (e.g., orthographic and phonological abilities). Compatibly, we found that performance on the poet recognition test was correlated with ACT scores in both experiments 1 and 2. Moreover, the poet recognition test was related to reading more poems in the last year and performing well on the poetry memory test in experiment 1, suggesting that our poet recognition test has construct validity.

There are potential limitations to our research. First, we can never know for certain whether we manipulated the inner voice that is believed to accompany silent reading (Huey, 1908/1968). Fodor (2002) noted that the role of prosody in silent reading is hard to prove. Silent readers are thought to project prosodic patterns onto text, and these projections can only be indirectly observed (via helpfulness ratings, eye tracking, reaction time, and event-related potentials). Second, the integrity of our findings assumes a direct link between emphatic stylistic alterations and implicit emphasis in the reader's inner voice. It is possible that our readers were responding to other linguistic features that systematically covaried with our experimental manipulations of prosodic meter. Nonetheless, our claims about a rhythmic inner voice are strengthened by testing specific predictions drawn from linguistic theory, by exploiting different stimuli (e.g., poems with different meters, heteronyms in prose), and by generalizing our previous research in silent reading on focus prosody (Gross et al., 2014) to prosodic meter and lexical stress. Third, the design of our experiments cannot tease apart interference and facilitation effects, because the experimental design lacks a plain text control condition. We do not have a plain text condition because our primary dependent variable, helpfulness ratings, cannot be applied to plain text. Nonetheless, the lack of a plain text control condition opens the possibility that the observed differences in helpfulness ratings might reflect only interference in the incongruent condition, only facilitation in the congruent condition, or both. That our effects might only reflect interference is plausible, as rhythmic

irregularities in natural speech were particularly costly (interfering) to ongoing language processing by listeners (Bohn, Knaus, Wiese, & Domas, 2013).

Our ongoing experiments are evaluating whether beginning readers, struggling readers, or late speakers of English might benefit from marking stress explicitly in written English. We recently found that the proper melody of heteronyms when silently reading was undetected by some college students who had low ACT scores or learned English as a second language (Gross, Plotkowski, & Winegard, 2015). Beginning or struggling readers might benefit from reading lessons that draw a more explicit link between the rhythm in speech and the rhythm in writing. Stress-alternating rhythms are easily found in children's books and nursery rhymes, although inconsistencies in the prevalent meter are common. Consider the appearance of the trochaic meter in a popular book by Dr. Seuss (1960): "ONE fish / TWO fish / RED fish / BLUE fish" and "SOME have TWO feet / and SOME have FOUR. / SOME have SIX feet / and SOME have MORE" (n.p.). A trochaic rhythm also appears in the nursery rhyme "TWINKle, TWINKle, LITtle STAR." The dactylic and iambic meters co-occur in a childhood favorite: "HICKory, DICKory, DOCK [dactylic] / The MOUSE ran UP the CLOCK [iambic]" (Opie & Opie, 1997, p. 244). Dr. Seuss (1958) wrote the book *Yertle the Turtle* in anapestic meter (e.g., "And toDAY, the great YERtle, that MARvelous HE, / Is KING of the MUD. That is ALL, he can SEE" (p. 39).

For our next project, we will evaluate whether young readers' prosodic awareness might benefit from seeing the stress pulses of the different rhythms materialized in their literature. Our prosody training manipulation will begin with simple rhythms (iambic and trochaic) to scaffold the learning of more difficult rhythms (dactylic and anapestic). A similar prosody training intervention is underway with late speakers of English and struggling adult readers.

To conclude, the interface of the English orthography is underspecified with regard to rhythm and stress. A long-term goal is to investigate the feasibility of developing a free application for use with smart devices that would transform ordinary text to stylistically enhanced text. Can YOU iMAGine an APP that HELPS you HEAR the RHYTHm of TEXT?

NOTES

[The copyright line for this article was updated on November 13, 2017, after first online publication on July 29, 2017.]

¹ The stylistic alterations to text should not be confused with an alternate meaning of *stylistic*, the linguistic and tonal style of text.

² To mark congruent and incongruent stress pulses in the stimuli used in experiment 1, font enlargement from 12 point to 16 point with bolding was used. Because *Reading Research Quarterly's*

design template could not accommodate larger font sizes in the examples furnished in this article, capital letters are used in lieu of enlargement. See the accompanying appendixes in the online version of this article to see how the stimuli were rendered in our experiments.

³ When three independent replications of the experiment were analyzed, the three-way interaction among syllable location, place in sentence, and stress marking was not significant, $F(1, 598) = 0.59$, $p = .4$, $\eta_p^2 = .001$ by subjects, and $F(1, 14) = 0.51$, $p = .5$, $\eta_p^2 = .03$ by items.

REFERENCES

- Abramson, M., & Goldinger, S.D. (1997). What the reader's eye tells the mind's ear: Silent reading activates inner speech. *Perception & Psychophysics*, 59(7), 1059–1068. <https://doi.org/10.3758/BF03205520>
- Ashby, J., & Clifton, C., Jr. (2005). The prosodic property of lexical stress affects eye movements during silent reading. *Cognition*, 96(3), B89–B100. <https://doi.org/10.1016/j.cognition.2004.12.006>
- Ashby, J., & Rayner, K. (2004). Representing syllable information during silent reading: Evidence from eye movements. *Language and Cognitive Processes*, 19(3), 391–426. <https://doi.org/10.1080/01690960344000233>
- Attridge, D. (1995). *Poetic rhythm: An introduction*. New York, NY: Cambridge University Press.
- Bader, M. (1998). Prosodic influences on reading syntactically ambiguous sentences. In J.D. Fodor & F. Ferreira (Eds.), *Studies in theoretical psycholinguistics: Vol. 21. Reanalysis in sentence processing* (pp. 1–46). Dordrecht, The Netherlands: Kluwer Academic.
- Bartlett, J. (1937). *Familiar quotations* (11th ed.). Boston, MA: Little, Brown.
- Bohn, K., Knaus, J., Wiese, R., & Domas, U. (2013). The influence of rhythmic (ir)regularities on speech processing: Evidence from an ERP study on German phrases. *Neuropsychologia*, 51(4), 760–771. <https://doi.org/10.1016/j.neuropsychologia.2013.01.006>
- Bolinger, D. (1978). Intonation across languages. In J. Greenberg (Ed.), *Universals of human language, Vol. 2: Phonology* (pp. 471–524). Stanford, CA: Stanford University Press.
- Breen, M., & Clifton, C., Jr. (2011). Stress matters: Effects of anticipated lexical stress on silent reading. *Journal of Memory and Language*, 64(2), 153–170. <https://doi.org/10.1016/j.jml.2010.11.001>
- Breen, M., & Clifton, C., Jr. (2013). Stress matters revisited: A boundary change experiment. *Quarterly Journal of Experimental Psychology*, 66(10), 1896–1909. <https://doi.org/10.1080/17470218.2013.766899>
- Breen, M., Kaswer, L., Van Dyke, J.A., Krivokapić, J., & Landi, N. (2016). Imitated prosodic fluency predicts reading comprehension ability in good and poor high school readers. *Frontiers in Psychology*, 7, 10–26. <https://doi.org/10.3389/fpsyg.2016.01026>
- Bryant, G.A., & Barrett, H.C. (2007). Recognizing intentions in infant-directed speech: Evidence for universals. *Psychological Science*, 18(8), 746–751. <https://doi.org/10.1111/j.1467-9280.2007.01970.x>
- Chan, J.S., & Wade-Woolley, L. (2016). Explaining phonology and reading in adult learners: Introducing prosodic awareness and executive functions to reading ability. *Journal of Research in Reading*. Advance online publication. <https://doi.org/10.1111/1467-9817.12083>
- Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York, NY: Harper & Row.
- Clin, E., Wade-Woolley, L., & Heggie, L. (2009). Prosodic sensitivity and morphological awareness in children's reading. *Journal of Experimental Child Psychology*, 104(2), 197–213. <https://doi.org/10.1016/j.jecp.2009.05.005>
- Cooper, W.E., Eady, S.J., & Mueller, P.R. (1985). Acoustical aspects of contrastive stress in question–answer contexts. *The Journal of the Acoustical Society of America*, 77(6), 2142–2156. <https://doi.org/10.1121/1.392372>
- Cutler, A., & Carter, D.M. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer Speech & Language*, 2(3/4), 133–142. [https://doi.org/10.1016/0885-2308\(87\)90004-0](https://doi.org/10.1016/0885-2308(87)90004-0)
- Dowhower, S.L. (1991). Speaking of prosody: Fluency's unattended bedfellow. *Theory Into Practice*, 30(3), 165–175. <https://doi.org/10.1080/00405849109543497>
- Eady, S.J., & Cooper, W.E. (1986). Speech intonation and focus location in matched statements and questions. *The Journal of the Acoustical Society of America*, 80(2), 402–415. <https://doi.org/10.1121/1.394091>
- Ferreira, F., & Karimi, H. (2015). Prosody, performance, and cognitive skill: Evidence from individual differences. In L. Frazier & E. Gibson (Eds.), *Explicit and implicit prosody in sentence processing: Studies in honor of Janet Dean Fodor* (pp. 119–132). Cham, Switzerland: Springer.
- Fodor, J.D. (1998). Learning to parse? *Journal of Psycholinguistic Research*, 27(2), 285–319. <https://doi.org/10.1023/A:1023258301588>
- Fodor, J.D. (2002). Prosodic disambiguation in silent reading. In M. Hirotani (Ed.), *Proceedings of the 32nd conference of the North East Linguistic Society* (Vol. 1, pp. 113–132). Amherst: Graduate Linguistic Student Association, University of Massachusetts.
- Fudge, E. (1984). *English word-stress*. Boston, MA: Allen & Unwin.
- Goswami, U., Wang, H.-S., Cruz, A., Fosker, T., Mead, N., & Huss, M. (2011). Language-universal sensory deficits in developmental dyslexia: English, Spanish, and Chinese. *Journal of Cognitive Neuroscience*, 23(2), 325–337. <https://doi.org/10.1162/jocn.2010.21453>
- Gross, J., Millett, A.L., Bartek, B., Bredell, K.H., & Winegard, B. (2014). Evidence for prosody in silent reading. *Reading Research Quarterly*, 49(2), 189–208. <https://doi.org/10.1002/rrq.67>
- Gross, J., Plotkowski, A.R., & Winegard, B. (2015, May). *In the melody of the reader's inner voice, heteronyms are tricky*. Poster session presented at the 27th annual meeting of the Association for Psychological Science, New York, NY.
- Holliman, A.J., Gutiérrez Palma, N., Critten, S., Wood, C., Cunnane, H., & Pillinger, C. (2017). Examining the independent contribution of prosodic sensitivity to reading and spelling in early readers. *Reading and Writing*, 30(3), 509–521. <https://doi.org/10.1007/s11145-016-9687-z>
- Holliman, A.J., Wood, C., & Sheehy, K. (2010). Does speech rhythm sensitivity predict children's reading ability 1 year later? *Journal of Educational Psychology*, 102(2), 356–366. <https://doi.org/10.1037/a0018049>
- Hualde, J.I., & Nadeu, M. (2014). Rhetorical stress in Spanish. In H. van der Hulst (Ed.), *Word stress: Theoretical and typological issues* (pp. 228–252). New York, NY: Cambridge University Press.
- Huey, E.B. (1968). *The psychology and pedagogy of reading: With a review of the history of reading and writing and of methods, texts, and hygiene in reading*. Cambridge, MA: MIT Press. (Original work published 1908)
- Hwang, H., & Schafer, A.J. (2009). Constituent length affects prosody and processing for a dative NP ambiguity in Korean. *Journal of Psycholinguistic Research*, 38(2), 151–175. <https://doi.org/10.1007/s10936-008-9091-1>
- Jusczyk, P.W. (1997). Finding and remembering words: Some beginnings by English-learning infants. *Current Directions in Psychological Science*, 6(6), 170–174. <https://doi.org/10.1111/1467-8721.ep10772947>
- Kelly, M.H., & Bock, J.K. (1988). Stress in time. *Journal of Experimental Psychology: Human Perception and Performance*, 14(3), 389–403. <https://doi.org/10.1037/0096-1523.14.3.389>
- Kentner, G. (2012). Linguistic rhythm guides parsing decisions in written sentence comprehension. *Cognition*, 123(1), 1–20. <https://doi.org/10.1016/j.cognition.2011.11.012>

- Kincaid, J.P., Fishburne, R.P., Jr., Rogers, R.L., & Chissom, B.S. (1975). *Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy-enlisted personnel* (Research Branch Report No. 8-75). Millington: Naval Technical Training, U.S. Naval Air Station, Memphis, TN.
- Kintsch, W. (2005). An overview of top-down and bottom-up effects in comprehension: The CI. *Discourse Processes*, 39(2/3), 125–128. <https://doi.org/10.1080/0163853X.2005.9651676>
- Kintsch, W., & Mross, E.F. (1985). Context effects in word identification. *Journal of Memory and Language*, 24(3), 336–349. [https://doi.org/10.1016/0749-596X\(85\)90032-4](https://doi.org/10.1016/0749-596X(85)90032-4)
- Koriat, A., Kreiner, H., & Greenberg, S.N. (2002). The extraction of structure during reading: Evidence from reading prosody. *Memory & Cognition*, 30(2), 270–280. <https://doi.org/10.3758/BF03195288>
- Liberman, M., & Prince, A.S. (1977). On stress and linguistic rhythm. *Linguistic Inquiry*, 8(2), 249–336.
- McCutchen, D., Bell, L.C., France, I.M., & Perfetti, C.A. (1991). Phoneme-specific interference in reading: The tongue-twister effect revisited. *Reading Research Quarterly*, 26(1), 87–103. <https://doi.org/10.2307/747733>
- Miller, J., & Schwanenflugel, P.J. (2008). A longitudinal study of the development of reading prosody as a dimension of oral reading fluency in early elementary school children. *Reading Research Quarterly*, 43(4), 336–354. <https://doi.org/10.1598/RRQ.43.4.2>
- National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups* (NIH Publication No. 00-4754). Washington, DC: National Institute of Child Health & Human Development, National Institutes of Health.
- Nava, E. & Zubizarreta, M.L. (2010). Deconstructing the nuclear stress algorithm: Evidence from second language speech. In N. Erteschik-Shir & L. Rochman (Eds.), *The sound patterns of syntax* (pp. 291–316). New York, NY: Oxford University Press.
- Nespor, M., & Vogel, I. (1986). *Prosodic phonology*. Dordrecht, The Netherlands: Foris.
- Norman, D.A. (1988). *The design of everyday things*. New York, NY: Doubleday.
- Patterson, K.E., & Coltheart, V. (1987). Phonological processes in reading: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 421–448). Hillsdale, NJ: Erlbaum.
- Perfetti, C.A., & Bell, L. (1991). Phonemic activation during the first 40 ms of word identification: Evidence from backward masking and priming. *Journal of Memory and Language*, 30(4), 473–485. [https://doi.org/10.1016/0749-596X\(91\)90017-E](https://doi.org/10.1016/0749-596X(91)90017-E)
- Pinsky, R. (1988). *The sounds of poetry: A brief guide*. New York, NY: Farrar, Straus and Giroux.
- Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics*, 1(1), 75–116. <https://doi.org/10.1007/BF02342617>
- Rumelhart, D.E., & McClelland, J.L. (1982). An interactive activation model of context effects in letter perception: Part 2. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review*, 89(1), 60–94. <https://doi.org/10.1037/0033-295X.89.1.60>
- Schafer, A.J., Speer, S.R., Warren, P., & White, S.D. (2000). Intonational disambiguation in sentence production and comprehension. *Journal of Psycholinguistic Research*, 29(2), 169–182. <https://doi.org/10.1023/A:1005192911512>
- Schreiber, P.A. (1991). Understanding prosody's role in reading acquisition. *Theory Into Practice*, 30(3), 158–164. <https://doi.org/10.1080/00405849109543496>
- Schwanenflugel, P.J., & Benjamin, R.G. (2016). Lexical prosody as an aspect of oral reading fluency. *Reading and Writing*, 30(1), 1–20. <https://doi.org/10.1007/s11145-016-9667-3>
- Schwanenflugel, P.J., Hamilton, A.M., Kuhn, M.R., Wisenbaker, J.M., & Stahl, S.A. (2004). Becoming a fluent reader: Reading skill and prosodic features in the oral reading of young readers. *Journal of Educational Psychology*, 96(1), 119–129. <https://doi.org/10.1037/0022-0663.96.1.119>
- Schwanenflugel, P.J., Westmoreland, M.R., & Benjamin, R.G. (2015). Reading fluency skill and the prosodic marking of linguistic focus. *Reading and Writing*, 28(1), 9–30. <https://doi.org/10.1007/s11145-013-9456-1>
- Selkirk, E. (1986). *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.
- Selkirk, E. (1995a). Sentence prosody: Intonation, stress, and phrasing. In J.A. Goldsmith (Ed.), *Handbook of phonological theory* (pp. 550–569). Cambridge, MA: Blackwell.
- Selkirk, E. (1995b). The prosodic structure of function words. In J. Beckman, L.W. Dickey, & S. Urbanczyk (Eds.), *Papers in optimality theory* (pp. 439–470). Amherst: Graduate Linguistic Student Association, University of Massachusetts.
- Selkirk, E. (2000). The interaction of constraints on prosodic phrasing. In M. Horne (Ed.), *Prosody: Theory and experiment* (pp. 231–261). Dordrecht, The Netherlands: Kluwer.
- Siptár, P., & Törkenczy, M. (2007). *The phonology of Hungarian*. Oxford, England: Oxford University Press.
- Stanovich, K.E., & West, R.F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly*, 24(4), 402–433. <https://doi.org/10.2307/747605>
- 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(1), 64–81. <https://doi.org/10.1037/0096-3445.136.1.64>
- Treiman, R. (2001). Reading. In M. Aronoff & J. Rees-Miller (Eds.), *Blackwell handbook of linguistics* (pp. 664–672). Oxford, England: Blackwell.
- Treiman, R., & Kessler, B. (2005). Writing systems and spelling development. In M. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 120–134). Oxford, England: Blackwell.
- Van Orden, G.C., & Kloos, H. (2005). The question of phonology and reading. In M.S. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 61–78). Malden, MA: Blackwell.
- Vroomen, J., Tuomainen, J., & de Gelder, B. (1998). The roles of word stress and vowel harmony in speech segmentation. *Journal of Memory and Language*, 38(2), 133–149. <https://doi.org/10.1006/jmla.1997.2548>
- Vul, E., & Pashler, H. (2008). Measuring the crowd within: Probabilistic representations within individuals. *Psychological Science*, 19(7), 645–647. <https://doi.org/10.1111/j.1467-9280.2008.02136.x>
- Wade-Woolley, L., & Heggie, L. (2015). Implicit knowledge of word stress and derivational morphology guides skilled readers' decoding of multisyllabic words. *Scientific Studies of Reading*, 19(1), 21–30. <https://doi.org/10.1080/10888438.2014.947647>
- Whalley, K., & Hansen, J. (2006). The role of prosodic sensitivity in children's reading development. *Journal of Research in Reading*, 29(3), 288–303. <https://doi.org/10.1111/j.1467-9817.2006.00309.x>

LITERATURE CITED

- Opie, I., & Opie, P. (1997). *The Oxford dictionary of nursery rhymes* (2nd ed.). Oxford, England: Oxford University Press.
- Seuss, Dr. (T.S. Geisel). (1958). *Yertle the Turtle and other stories*. New York, NY: Random House.
- Seuss, Dr. (T.S. Geisel). (1960). *One fish two fish red fish blue fish*. New York, NY: Random House.

Submitted September 14, 2016
Final revision received May 12, 2017
Accepted May 22, 2017

JENNIFER GROSS (corresponding author) is a professor of psychology at Grand Valley State University, Allendale, Michigan, USA; e-mail grossj@gvsu.edu.

BO WINEGARD is a graduate student of psychology at Florida State University, Tallahassee, USA; e-mail winegard@psy.fsu.edu.

ANDREA R. PLOTKOWSKI was a student at Grand Valley State University, Allendale, Michigan, USA, when this research was conducted. She is now a clinical audiologist at Ear, Nose & Throat Consultants of East Tennessee, Knoxville, USA; e-mail andreaplotkowski@gmail.com.

Supporting Information

Additional supporting information may be found in the online version of this article:

- Appendix A: Stimuli in Experiment 1
- Appendix B: Poetry Memory Test Used in Experiment 1
- Appendix C: Poet Recognition Test With Poets Highlighted
- Appendix D: Stimuli in Experiment 2
- Appendix E: Reading Comprehension Test Used in Experiment 2, With the Correct Answers as Option A
- Appendix F: True/False SAT Vocabulary Tests Used in Experiment 2