Read Writ
DOI 10.1007/s11145-016-9690-4

That noun phrase may be beneficial and this may not be: discourse cohesion in reading and writing

Scott A. Crossley[1] · Dani Francuz Rose[1] · Cassondra Danekes[1] · Charles Wesley Rose[1] · Danielle S. McNamara[2]

Abstract This paper examines the effects of attended and unattended demonstratives on text processing, comprehension, and writing quality in two studies. In the first study, participants ($n = 45$) read 64 mini-stories in a self-paced reading task and identified the main referent in the clauses. The sentences varied in the type of demonstratives (i.e., this, that, these, and those) contained in the sentences and whether the referent was followed by a demonstrative determiner and noun (i.e., an attended demonstrative) or a demonstrative pronoun (i.e., an unattended demonstrative). In the second study, 173 persuasive essays written by high school students were rated by expert judges on overall writing quality using a standardized rubric. Expert coders manually counted the number and types of demonstratives (attended and unattended demonstratives) in each essay. These counts were used to predict the human scores of essay quality. The findings demonstrate that the use of unattended demonstratives as anaphoric references is disadvantageous to both reading time and referent identification. However, these disadvantages become advantages in terms of essay quality likely because linguistic complexity is a strong indicator of high proficiency writing. From a text processing and comprehension viewpoint, the findings indicate, then, that anaphoric reference is not always beneficial and does not always create a more cohesive text. In contrast, from a writing context, the use of unattended demonstratives leads to a more linguistically complex text, which generally equates to a higher quality text.

✉ Scott A. Crossley
scrossley@gsu.edu

[1] Department of Applied Linguistics/ESL, Georgia State University, 25 Park Place Suite 1500, Atlanta, GA 30303, USA

[2] Department of Psychology, Learning Sciences Institute, Arizona State University, Tempe, AZ, USA

Published online: 15 September 2016

Keywords Text processing · Text comprehension · Essay quality · Discourse cohesion · Unattended demonstratives · Attended demonstratives

There are a multitude of factors, both within text and within individuals that influence text comprehension, including background knowledge (McNamara, Kintsch, Songer, & Kintsch, 1996; Shapiro, 2004), reading skill (Buswell, 1922), and the linguistic features within a text (Just and Carpenter, 1987; Rayner & Pollatsek, 1994). One linguistic aspect of text that plays an important role in comprehension is cohesion, especially for challenging texts that pose knowledge demands on readers (Loxterman, Beck, & McKeown, 1994; McNamara et al., 1996; McNamara & Kintsch, 1996).

In part, the processing demands of challenging texts can be met by using cohesive devices to link text segments together either explicitly or implicitly. An example of explicit cohesive devices are connectives such as "because," "therefore," and "consequently," which act as explicit cues that inform the reader that there are relations between ideas and the nature of those relations. In contrast, an important type of implicit cohesion that links text segments is anaphoric reference. Anaphoric reference is the use of an expression, usually a personal pronoun (e.g., "he," "she," "it") or demonstrative (e.g., "this," "that," "these") that depends on a previous expression (i.e., its antecedent) for interpretation (Halliday & Hasan, 1976). For example, in the sentence "John asked Susan if she was happy," "Susan" is the antecedent for the anaphoric referent "she."

Anaphoric references that are pronominal in nature (e.g., "he," "she," "they") have been shown to be important components of text comprehensibility and readability (Amaral, 1985; Kossowska, 2004). To our knowledge and as reported in recent research (e.g., Wulff, Römer, & Swale, 2012), analogous research has not been conducted that examines corresponding effects for demonstrative pronouns (e.g., "this", "that," "those") used as anaphoric references—even though Halliday and Hasan (1976) argued that demonstrative pronouns lead to text cohesion and are one of the major cohesive devices of English, at least in the case of the demonstrative pronoun "this". The first analysis in this study addresses the role of demonstrative pronouns in reading samples by investigating differences in processing time and text comprehension for sentences that contain unattended demonstratives (i.e., "this," "that," "these," and "those" used pronominally ["This scares me"]; Geisler, Kaufer, & Steinberg, 1985) and attended demonstratives, which are used as determiners (i.e., "this," "that," "these," and "those" used before a noun ["This dog scares me"]), thus addressing a gap in our understanding of comprehension and discourse processing.

The second analysis in this study investigates the potential for attended and unattended demonstratives to predict writing quality. From a stylistic perspective, the use of unattended demonstratives has generally been discouraged in academic writing because the nature of unattended demonstratives, when compared to attended demonstratives, is more imprecise (Markel, 2004; Strunk & White, 2000). Such advice may undermine the notion that the use of unattended demonstratives leads to a cohesive text (Halliday & Hasan, 1976). This advice may also counter

empirical findings in composition studies, which have shown that higher quality writing is related to more complex linguistic features (Crossley, Allen, Kyle, & McNamara, 2014; McNamara, Crossley, & McCarthy, 2010; Witte & Faigley, 1981), which may include unattended demonstratives. To investigate these contrasting perspectives, we assess potential links between human judgments of writing quality in persuasive essays and the incidence of attended and unattended demonstratives in those essays.

The research questions that guide this study are: (a) are sentences with unattended demonstratives processed faster than sentences with attended demonstrative (i.e., are they read faster), (b) are the antecedents to unattended demonstratives identified to a greater degree than the antecedents of attended demonstratives, and (c) are attended and unattended demonstratives predictive of human judgments of writing quality?

Anaphoric reference

Anaphoric reference is likely the most common form of cohesion found in the English language (Halliday & Hasan, 1976). The clearest examples of anaphoric reference take place over the course of several sentences, although anaphoric reference can be localized within a sentence (Doyle, 1982). Most anaphoric references are deictic in nature (i.e., refer to the spatial and temporal co-ordinates of the act of an utterance; Lyons, 1977). Traditionally, both personal and demonstrative pronouns can be thought of as containing deictic information. However, the two types of pronouns differ because demonstrative pronouns indicate directed spatiality (i.e., "this" versus "that") while personal pronouns indicate the presence or absence of a specific referent in the physical discourse ("I" and "you" versus "he," "she," and "it"). Thus, the principal difference between personal and demonstrative pronouns is the distinction between pointing to something, which may be an object, an action, or an event (i.e., the demonstratum in the case of demonstrative pronouns) and referring to a specific item, which is usually an object (i.e., the referent in the case of personal pronouns; Halliday & Hasan, 1976; Webber, 1991). In both cases, the anaphoric reference refers back to some section of previous text. However, demonstrative pronouns can refer to lexically transparent items (i.e., nouns), to general classes denoted by a noun (Halliday & Hasan, 1976), and to larger segments of previous discourse (i.e., an event; Garrod & Sanford, 1977), whereas personal pronouns only refer to lexically transparent items (i.e., nouns; Gernsbacher, 1989; Halliday & Hasan, 1976). Events, unlike nouns, are more complex in that they can refer to single and multiple propositions (i.e., ideas that contain objects, but also actions, locations, and spatiality). The differences in these two types of references are illuminated in the following example (adapted from Webber, 1991).

a. John built a two-armed robot.
b. *He* had learned about robotics in class.
c. Mary taught *it* how to play the saxophone.
d. *That* took *her* three months.

In this example, the personal pronouns (i.e., "he," "it," and "her") refer to specific references in the text as does the demonstrative pronoun "that." However, the demonstrative pronoun "that" refers to an entire event (i.e., teaching the robot to play the saxophone) while the personal pronouns refer to specific nouns and noun phrases (i.e., "John," "the robot," and "Mary"). In addition, if the final utterance were

e. *That* earned both of them good grades.

The demonstrative pronoun "that" could be thought of as referring to the first three events (building, learning, and teaching) in the string of clauses. Thus, while personal pronouns are generally only used to refer to specific nouns and noun phrases, demonstrative pronouns can be used to refer to both specific references in the text and to larger discourse segments such as events (Webber, 1991).

Anaphoric reference and text processing

The majority of research examining links between anaphoric reference and text processing has involved referential pronouns. Referential pronouns, especially those found in the subject position, are generally used to maintain reference to topicalized antecedents (Karmiloff-Smith, 1985) and avoid repetition. Linking text segments through the use of referential pronouns is argued to increase the grammatical cohesion in a text (Halliday & Hasan, 1976), which has been supported in a number of studies (Anderson, Garrod, & Sanford, 1983; Clark & Sengul, 1979; Ehrlich & Rayner, 1983; Hudson, Tanenhaus, & Dell, 1986; Kossowska, 2004). At least one study has also demonstrated that the referential semantic activation found in pronouns can assist in anaphoric resolution (Shillcock, 1982).

However, few, if any studies, have investigated the effects of demonstrative pronouns (i.e., unattended demonstratives) on text processing. Halliday & Hasan (1976) stated that unattended demonstratives help create cohesive text even though the use of an unattended demonstrative can include a broader sense of meaning (i.e., can refer to an event) or to a specific noun. Specifically, Halliday and Hasan argued that while it is not always easy to identify whether the referent of an unattended demonstrative is a specific noun or a larger event, the distinction is usually irrelevant and, in either case, the effect is cohesive (p. 67). In contrast, Finn (1995) argued that the use of unattended demonstratives to portray the meaning of an entire clause may lead to a more complex text while Geisler et al., (1985) argued that local cohesion is challenged in cases where unattended demonstratives refer to an event. Conversely, using an attended demonstrative before a noun may increase the clarity of the previous referent through lexical repetition (i.e., redundancy). While such redundancy would be advantageous in later recall of the referent, it may prove disadvantageous because the text would contain more symbols to express the same amount of information when compared to unattended demonstratives. These additional symbols could potentially slow down the flow of new information available to the reader (Finn, 1995). Thus, the choice between using an attended or

unattended demonstrative becomes one of economy versus clarity. In terms of economy, the unattended demonstrative can point back to a topic, bring it into focus, and discuss it in a single word. If the previous antecedent is clear, which is usually the case with shorter antecedents like nouns or noun phrases, then efficient encoding requires only the use of an unattended demonstrative (Geisler et al., 1985). However, if an attended demonstrative is used in this case, the writer risks losing the reader in inefficient and redundant prose (Finn, 1995). If the topic is unclear, as is usually the case with longer antecedents that are events and not nouns, then the use of an attended demonstrative can single out and clarify the topic of discussion (Geisler et al., 1985).

An important difference between pronominal demonstratives and demonstrative determiners in terms of antecedent recognition is the distinction between explicit and implicit referents (MacDonald & MacWhinney, 1990; Shillcock, 1982; Speelman & Kirsner, 1990). Demonstrative pronouns have been identified as less explicit than personal pronouns (Garrod, Freudenthal, & Boyle, 1994) and there is evidence that explicit anaphoric reference can aid in identifying antecedents more quickly than less explicit anaphoric references. For example, Gernsbacher (1989) found immediate activation of antecedents and suppression of non-antecedents in repeated name-recognition tasks. Dell, McKoon and Ratcliff (1983) examined definite description and found immediate activation for such antecedents as compared to non-anaphoric conditions. Together, these studies provide evidence that explicit identification leads to quicker antecedent recognition. Such findings might be transferrable to the use of attended demonstratives (i.e., "that dog") in which the demonstrative acts as an explicit identifier, likely leading to better identification of antecedents than for unattended demonstratives. However, as stated earlier, to our knowledge, the effects of unattended demonstratives on text processing and antecedent identification have not been compared directly to personal pronouns or to explicit nouns preceded by demonstrative determiners.

Anaphoric reference and writing

The predominant approach to investigating the effects of anaphoric reference on readers has been to conduct psycholinguistic studies analyzing reading times, response times, or recall accuracies. Another, less common approach to assessing the effect of anaphoric reference on readers has been to examine its relations with expert ratings of text quality (McCulley, 1985; Witte & Faigley, 1981). For instance, Witte and Faigley (1981) explored relations between features of cohesion and human judgments of persuasive writing quality. They found that, in general, high quality essays judged to be of higher quality by expert readers used more cohesive devices than low quality essays. Specifically, they found that high quality essays contained twice the level of reference cohesion (e.g., personal pronouns, unattended demonstratives, and attended demonstratives) as low quality essays. The majority of this difference, however, comprised the use of third person pronouns. In a similar study, McCulley (1985) examined relations between human judgments of persuasive writing quality and overall cohesion and between human judgments of

persuasive writing quality and specific types of reference cohesion. McCulley reported that the total number of cohesive devices in an essay positively predicted judgments of essay quality and, more specifically, that demonstrative reference (including both attended and unattended demonstratives) positively correlated with human ratings of essay quality.

While studies directly assessing links between anaphoric reference and essay quality are infrequent, stylistic advice on when and how to use anaphoric references in academic writing abounds. Stylistically, unattended demonstratives as deictic references have long been discouraged in writing because their antecedents are more obscure and their use is more imprecise (Markel, 2004; Strunk & White, 2000). As a result, writing and style guides often suggest that writers avoid the use of unattended demonstratives because they can lead to potential ambiguity in writing resulting from vague references (American Psychological Association, 2001; Axelrod & Cooper 2008; Ede, 2004; Faigley, 2007) which, in turn, may affect their processing and the effects the writing have on the reader (Swales & Feak, 2000, 2004). The reason for this, as stated earlier, is that unattended demonstratives can refer to the complete sense of a preceding sentence or clause (Strunk & White, 2000). However, such potential ambiguity does not stop many writers from using unattended demonstratives. Indeed, the percentage of sentence initial unattended demonstrative as compared to sentence initial attended demonstratives ranges from 20 to 56 % across various disciplines with the lowest frequency of unattended demonstratives found in writings about dentistry and the highest in philosophical articles (Gray & Cortes 2011; Swales, 2005).

Although the use of unattended demonstratives is common, many reference texts contend that, in most cases, demonstratives should be followed by nouns (Johnson-Sheehan, 2005; Markel, 2004). Such advice is founded on the notion that demonstratives followed by nouns can increase awareness of given-new patterns by appropriately summing up previous phrases so that they can be commented on (Swales, 2005; Williams, 1985). The one study we found that focused on novice student writing reported that novice-writing contained fewer unattended demonstratives compared to attended demonstratives giving some credence to the notion that novice writers followed the advice found in writing and style guides (Rustipa, 2015). In contrast, the one study we found that focused on advanced writers reported that frequencies of unattended "this" increased between the final year of undergraduate writing and the third year of graduate writing (Römer and Wulff 2010) indicating that more advanced writers did not follow the advice found in writing and style guides.

Overall, studies examining reference cohesion (including attended and unattended demonstratives) show that a greater incidence of such features positively relates to higher human judgments of writing quality. In contrast, stylistic guides suggest that writers should avoid the use of unattended demonstratives and favor attended demonstratives. The first case indicates that demonstratives in general may increase text cohesion in a similar manner as personal pronouns and that their presence in an essay could be associated with estimates of higher quality writing. However, these previous studies combined attended and unattended demonstratives and did not examine them independent of one another. The second case indicates

that unattended demonstratives may render a text more linguistically complex (and potentially less cohesive) leading to lower writing quality. However, such a position should be tempered with recent studies that have demonstrated that linguistic complexity in writing samples is positively associated with human ratings of writing quality. That is to say, essays that contain more complex syntax, more sophisticated vocabulary, and fewer cohesive devices are generally scored higher (Crossley, Weston, McLain Sullivan, & McNamara, 2011; McNamara, Crossley, & McCarthy, 2010; McNamara et al., 2013; Crossley & McNamara, 2010, 2011). In total, research and stylistic advice considering the effects of demonstrative pronouns on human judgments of writing quality are contradictory. In addition, to our knowledge, no research has examined differences between attended and unattended demonstratives, which may prove to be an important distinction in terms of explaining human ratings of essay quality.

Current study

In this paper, we examine differences in the processing and recall of attended and unattended demonstratives. We also examine whether the use of attended or unattended demonstratives by writers predicts writing success as evidence by expert ratings of essay quality. In our first study, participants read 64 ministories that consisted of three sentences in a self-paced reading task. After each three-sentence ministory, participants were asked to identify the main referent. The sentences varied in the type of demonstratives (i.e., "this," "that," "these," and "those") contained in the sentences and whether the referent was followed by defined reference (i.e., used a demonstrative followed by a noun [attended demonstrative]) or undefined reference (used a demonstrative pronoun [unattended demonstrative]). The purpose of the experiment was to examine differences in processing times for attended and unattended demonstratives and to assess whether attended or unattended demonstratives led to increased identity of the main sentence referent.

There are a number of competing theories for what may result. According to Halliday and Hasan (1976), unattended demonstratives are cohesive devices (regardless of whether an unattended demonstrative refers to a specific noun or to a larger text element). Since previous studies have indicated that cohesive devices lead to faster text processing (Anderson et al., 1983; Clark & Sengul, 1979; Ehrlich & Rayner, 1983; Hudson et al., 1986; Kossowska, 2004), we could hypothesize that unattended demonstratives would be processed more quickly than attended demonstratives. However, Finn (1995) and Geisler et al. (1985) argue that unattended demonstratives can be more complex, especially when used to refer back to events. Additionally, in cases where it is difficult to pinpoint the referent to which an unattended demonstrative refers, local cohesion can be tested (Geisler et al., 1985). In this case, we could hypothesize that unattended demonstratives would be processed more slowly than attended demonstratives. In terms of referent identification, there are also competing theories. If we presume that unattended demonstratives act similarly to pronominal demonstratives, we might expect that their use could lead to greater recall of antecedents (Shillcock, 1982) because of

semantic activation of the referent by the pronoun. More likely, given that unattended demonstratives do not contain the redundancy found in attended demonstratives (i.e., they do not contain lexical repetition; McKoon & Ratcliff, 1980), they should have lower accuracy in terms of recall (Finn, 1995; Geisler et al., 1985).

In our second study, we collected 173 persuasive essays written by high school students and had expert raters judge them on overall writing quality using a standardized rubric. We then had expert coders manually count the number and types of demonstratives (attended and unattended demonstratives) in each essay and we used these incidence scores to predict the human scores of essay quality. There are two competing theories based on text cohesion and text complexity that may predict the results. On the one hand, if Halliday and Hasan (1976) are correct, a greater number of unattended demonstratives would indicate a cohesive text and would be positively associated with essay quality (McCulley, 1985; Witte & Faigley, 1981). In a similar manner, if unattended demonstratives lead to increased processing difficulty because of the complexity of the referent (Finn, 1995; Geisler et al., 1985), a greater number of unattended demonstratives may be positively associated with essay quality.

Experiment 1

The first experiment examined differences in processing speed and referent identification accuracy comparing attended and unattended demonstratives.

Method

Participants

Forty-five graduate and advanced undergraduate students from a large university in the southeastern United States either volunteered for the experiment or were given extra credit in a linguistics class to participate. Demographic data was self-reported by the participants: 78 % of the participants were female; 82 % of the participants were either native speakers of English or bilingual in English; 18 % of the participants were advanced speakers of English from a variety of first languages; 20 % were graduate students while the remaining 80 % were undergraduate students.

Materials and study design

The experimental stimuli consisted of 64 ministories containing three related sentences. All participants read the same 64 ministories and the ministories were presented in random order to control for ordering effects. The first sentence in each group set the context (e.g., "Someone closed the door quietly."). The second sentence introduced the antecedent (e.g., "The doctor ran into the door."). The third sentence contained an unattended demonstrative as an anaphoric reference (e.g., "That hurt his head") or an attended demonstrative (e.g. "That door hurt his head.").

Table 1 Example stimuli from each condition

Condition	Clause 1	Clause 2	Clause 3	Question	Answer
Attended that	I love going to the theater	I watched a famous movie	That movie won many awards	What won many awards?	The movie
Unattended that	Jim can see the problem	A truck broke down	That caused a traffic jam	What caused a traffic jam?	The truck
Attended those	The sun can fade things	The schools do not look good	Those schools need to be painted	What need to be painted?	The schools
Unattended those	I like a clean house	Shoes are too dirty for inside	Those go over by the entrance	What go over by the entrance?	The shoes
Attended this	The classroom was full	One child was frowning	This child was sad	Who was sad?	The child
Unattended this	The boat was sinking quickly	They put a plug into the hole	This fixed the problem	What fixed the problem?	The plug
Attended these	Look toward the sky	There are many clouds today	These clouds are moving very fast	What are moving very fast?	The clouds
Unattended these	I really like fashion design	My shoes were made for me	These are my favorite	What are my favorite?	The shoes

In all cases, the unattended demonstratives referred back to the previous discourse (i.e., were discourse deictics; Himmelman, 1996). After the three sentences had been presented, a comprehension question appeared: "What hurt his head?" In all cases, the answer to the question could be a single noun or noun phrase (i.e., "the door"). The participants were then presented with a text box in which to type their responses. The text box was preceded by the definite article "the" in order to prime participants to answer with a noun phrase. There were eight experimental conditions: *attended that, unattended that, attended those, unattended those, attended this, unattended this, attended these,* and *unattended these*. Each experimental condition contained eight sentences. See Table 1 for example sentences from each condition.

To ensure the examples were naturalistic, all three sentences were taken from the Corpus of Contemporary American English (COCA; Davies, 2009), which includes 450 million words extracted from spoken sources, fiction pieces, popular magazines, newspapers, and academic journals from the years 1990–2012. All of the three sentences were carefully modified by three expert raters to ensure that the antecedent only occurred in the second sentence, that no irregular nouns were used as referents (i.e., people), that answers could not be guessed from context of the questions, and that the comprehension question could be answered with a single noun that was preceded by the determiner "the." Linguistically the sentences were controlled for length and word frequency (using frequency counts taken from the British National corpus and reported by the Tool for the Automatic Analysis of Lexical Sophistication [TAALES]; Kyle & Crossley, 2015) such that no differences existed among the eight experimental conditions in terms of sentence length or word frequency. In addition, the referents that comprised the answers to questions were controlled for word frequency in a similar matter such that no differences in referent frequency existed between conditions.

Apparatus and procedure

Participants were tested in separate, soundproof rooms containing desktop computers. At the beginning of the sessions, participants provided demographic information using internet-based survey software. E-Prime software was used to display the stimuli and to collect reading time recordings and referent identification. Reading times were collected for each sentence, but only the third sentence containing either an attended or unattended demonstrative was analyzed. Participants were first instructed that they would read a series of three sentences presented one sentence at a time and that they would advance through the sentences by pressing the spacebar on the computer's keyboard. After reading all three sentences, the participants were instructed to answer a question by entering a single noun into a provided textbox. After they provided an answer, they then pressed the enter key to move to the next sample. Each sentence was presented in the middle of the screen in a large font. The stimuli in the experiment were presented randomly. The session took about 30 min on average.

Statistical analysis

The data were first checked for outliers. Four participants were identified as outliers due to reading times exceeding three standard deviations below or beyond the first or third quartile. These participants were removed from the study. After removing these participants, we used a 10 % trimmed mean to control for item outliers. After controlling for all outliers, we had 2362 items for analysis from 41 participants.

For each of the variables (attended vs. unattended demonstratives and each of the eight classes of demonstratives), we used linear mixed effects (LME) modeling to evaluate baseline differences in subjects and items. For our reference baseline we used attended demonstratives when comparing differences between attended and unattended demonstratives as a whole. We used attended "this" for comparing the eight different demonstrative groups (attended and unattended "this," "that," "those," and "these"). We selected these baselines because attended demonstratives and attended "this" are more frequent in writing samples (Römer & Wulff, 2010). For the LME models, we report coefficients of the predictors based on significance at $p < 0.05$, their standard error, and derive p-values from the t-values for each of the factors in the model [standard procedure recommended by Mirman (2014)]. All analyses were carried out in R version 3.1.3 using the lme4 package (version 1.1–7; Bates, Maechler, Bolker, & Walker, 2014).

Results

The results of the mixed effects models revealed a number of statistically significant main effects for reading time and referent identification accuracy across the demonstrative types. These results are reported in Tables 2 and 3. The means and standard errors for each of the variables are reported in Table 4.

Starting with attended and unattended demonstratives, sentences with unattended demonstratives took longer to read ($B = 129.440$, $p < .001$) and reported lower

Table 2 Results of mixed-effects models comparing attended to unattended demonstratives for reading time and accuracy

	Reading time			Accuracy		
	Coeff	SE	*t* value	Coeff	SE	*t* value
Unattended demonstratives	129.440	25.390	5.098***	−0.072	0.008	−8.662***

Attended demonstratives were used as the reference baseline. Cells include coefficients, their standard errors, and level of significance estimates

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table 3 Results of mixed-effects models comparing demonstrative type for reading time and accuracy

	Reading time			Accuracy		
	Coeff	SE	t value	Coeff	SE	t value
Attended that	18.920	49.870	0.379	0.017	0.016	1.048
Attended these	125.530	49.770	2.522*	0.012	0.016	0.714
Attended those	168.590	50.420	3.343***	0.010	0.017	0.615
Unattended this	113.870	50.100	2.273*	−0.083	0.017	−5.061***
Unattended that	119.520	50.850	2.351*	−0.062	0.017	−3.698***
Unattended these	240.880	50.340	4.785***	−0.065	0.017	−3.943***
Unattended those	356.500	50.630	7.042***	−0.036	0.017	−2.172*

Attended this was used as the reference baseline. Cells include coefficients, their standard errors, and level of significance estimates

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table 4 Mean (standard deviations) for reading times and accuracy

Demonstrative type	Attended demonstratives		Unattended demonstratives	
	Accuracy	Reading time	Accuracy	Reading time
All	0.989 (0.104)	1747.490 (664.752)	0.918 (0.274)	1867.412 (708.323)
This	0.980 (0.141)	1672.162 (635.542)	0.896 (0.305)	1774.525 (635.577)
That	0.997 (0.057)	1692.757 (638.103)	0.918 (0.274)	1779.986 (697.647)
These	0.990 (0.099)	1792.319 (663.460)	0.915 (0.279)	1898.939 (716.464)
Those	0.990 (0.101)	1834.182 (710.421)	0.944 (0.229)	2017.267 (755.829)

referent identification accuracy ($B = -0.072$, $p = .001$). For specific types of demonstratives, compared to attended "this," all demonstrative types except attended "that" took longer to read in a sentence ("attended these"; $B = 125.530$, $p = .050$, "attended those"; $B = 168.590$, $p = .001$, "unattended this"; $B = 113.870$, $p = .050$, "unattended that"; $B = 119.520$, $p = .050$, "unattended these"; $B = 240.880$, $p = .001$, "unattended those"; $B = 356.500$, $p = .001$). For referent identification accuracy, compared to attended *this*, all unattended demonstrative types

demonstrated lower accuracy ("unattended this"; $B = -0.083$, $p = .001$, "unattended that;" $B = -0.062$, $p = .001$, "unattended these"; $B = -0.065$, $p = .001$, "unattended those;" $B = -0.036$, $p = .001$).

Discussion

Our first study examined the effects of using demonstratives as either pronouns ("This scares me") or determiners ("This dog scares me") on reading time and referent identification. Participants' reading times for sentences with unattended demonstratives were slower than sentences that contained attended demonstratives. Also, unattended demonstratives led to significantly lower referent identification when compared to attended demonstratives. These findings held for individual types (i.e., "this," "that," "these," "those") when compared to the most frequent attended demonstrative ("attended this"). The findings provide support for the hypothesis that unattended demonstratives can lead to increased processing times because of the potential for the referent to be more complex than the referent in attended demonstratives and that the redundancy found in attended demonstratives in the form of lexical repetition leads to greater referent identification (Finn, 1995; Geisler et al., 1985). We do not find support for the notion that, like other cohesive devices, unattended demonstratives lead to faster text processing, especially when compared to attended demonstratives (cf. Anderson et al., 1983; Clark & Sengul, 1979; Ehrlich & Rayner, 1983; Hudson et al., 1986; Kossowska, 2004) at least when compared to attended demonstratives. The findings have important theoretical implications for how attended and unattended demonstratives affect text comprehension processes.

Demonstratives and text cohesion

Previous research has indicated that the use of personal pronouns as anaphoric references leads to more cohesive text that is processed more quickly. Such findings have been reported for recent antecedents as compared to distant antecedents (Clark & Sengul, 1979; Ehrlich & Rayner, 1983), for topicalized antecedents (Anderson, Garrod, & Sanford, 1983), and for antecedents that occur in the subject as compared to the object position (Hudson, Tanenhaus, & Dell, 1986; McKoon & Ratcliff, 1980). However, similar effects for unattended demonstratives, which are argued to be cohesive devices (Halliday & Hasan, 1976), were not found in this study, at least when compared to attended demonstratives.

Sentences containing unattended demonstratives had slower reading times than those containing nouns preceded by demonstrative determiners even when these sentences were controlled for length and word frequency and when a specific noun, as compared to an event, was a plausible referent. Thus, the use of unattended demonstratives does not decrease reading times, indicating that a sentence containing an unattended demonstrative may be more complex as well as less cohesive than one containing redundant information (i.e., an attended demonstrative). This finding indicates a potential difference in how anaphoric pronominal and demonstrative references are processed. It appears that pronominal referents that refer to a specific

noun lead to a more cohesive text, especially when the antecedent is recent (Clark & Sengul, 1979; Ehrlich & Rayner, 1983), topicalized (Anderson et al., 1983), and in the subject position (Hudson et al., 1986; McKoon and Ratcliff, 1980. However, unattended demonstratives, even those that can refer to a specific noun, appear to interfere with text processing likely because the referent can be more obscure or complex (Markel, 2004; Strunk & White, 2000). This holds true for demonstrative pronouns as a class as well as to individual demonstrative pronoun types (i.e., "this," "that," "these," and "those"). In addition, researchers have indicated that the repetition found in attended demonstratives may also slow down the flow of information (Finn, 1995; Geisler et al., 1985) even though repeating a noun as found in the attended demonstrative condition could provide additional clarity through redundancy. Considering the reading time results, it appears that repetition of the noun can increase reading speed, at least when attended demonstratives are compared to unattended demonstratives.

Demonstratives and referent identification

Personal pronouns used as anaphoric referents can assist in anaphoric resolution (Shillcock, 1982), especially if the antecedents are topicalized (Speelman & Kirsner, 1990). However, this effect may not hold with unattended demonstratives because they are argued to be less explicit than personal pronouns since they can refer to both specific referent and also to segments of text (Garrod et al., 1994). In contrast, repeating the referent through the use of attended demonstratives may afford referent identification based on text redundancy.

Our analysis supports the notion that attended demonstratives lead to greater referent identification than unattended demonstratives. In addition, the low response accuracy reported in the unattended demonstratives condition is telling. Over 8 % of the unattended demonstratives were either identified incorrectly or not identified at all by the participants in this study (as compared to 1 % for the attended demonstratives). This finding indicates that the referents for unattended demonstratives are more difficult to identify, or at the least, the referent is more ambiguous leading to lower identification.

Experiment 2

The second experiment examined relations between the incidences of attended and unattended demonstratives and expert ratings of essay quality.

Method

Participants

For this study, we recruited 87 students from public high schools in the metro Phoenix area. Students ranged in age from 14 to 19 ($M = 15.9$, $SD = 1.3$) and

ranged in grade level from 9 to 12 ($M = 10.2$, $SD = 1.0$). Each student wrote two persuasive essays over the course of a 2–4 week period in the Writing Pal Automatic Writing Evaluation (AWE) system (Crossley, Roscoe, & McNamara, 2013; McNamara et al., 2013). The essays were written on two prompts (on the value of competition and on the role of images/impressions) counterbalanced across collection points. Of the 65 participants, 70.8 % were female and 29.2 % were male. Of the 87 participants, complete data for 86 of the participants was available. One participant did not complete the second essay. The final corpus contained 173 essays.

Essay scoring

Each essay was read and scored by two trained raters on overall quality (i.e., a holistic score). The holistic grading scale was based on a standardized rubric commonly used in assessing SAT[1] essays. The trained raters had Master's degrees in English and each rater had at least 3 years of experience teaching university level composition classes. The raters were informed that the distance between each score was equal. The raters were first trained to use the rubric with 20 practice essays and after the raters had reached an inter-rater reliability (IRR) of at least $r = .70$, the raters then scored the 173 essays independently. After scoring was completed, differences between raters were calculated. If the difference in ratings was less than two points, an average score was computed for that essay feature (e.g., a score of a 3 and a 4 would be averaged to a score of 3.5). If the difference was greater than two points, a third expert rater adjudicated the final rating and the average score between the third rater and the rating from the other two raters that was closest was averaged (e.g., if the initial scores were a 3 and a 5 and the third rater scored the essay a 3, the score given to the essay would be a 3).

Coding of demonstratives

Two raters who were advanced linguistic students at a large university in the southeastern United States were trained to classify the incidence of demonstratives found in each essay. Specifically, they were trained to code each demonstrative in each essay as being attended or unattended and to annotate the type of demonstrative (i.e., *this, these, that* and *those*). In addition, for undefined demonstratives, the raters located the initial referent and annotated the distance between the referent and the unattended demonstrative (e.g., the referent was in the preceding sentence or the referent was two sentences back). In practice, however, the vast majority of all referents were in the previous sentence. Initial IRR in terms of simple agreement between the two raters was high (99 %). Differences, where they existed, were adjudicated between the two raters.

[1] The SAT is a college entrance exam commonly administered in the United States. An important component of the exam is a writing section in which test-takers are required to produce an essay based on general knowledge within a 25-min time frame.

From the human coding, 10 feature counts were extracted: attended demonstratives counts for *this, these, that,* and *those* along with a total count for attended demonstratives and unattended demonstratives for *this, these, that,* and *those* along with a total count for unattended demonstratives. These counts were then divided by the number of words in the essay to provide an average count for each feature within each essay.

Results

For the essay analysis, we used the demonstrative counts as predictor variables and the human scores for overall essay quality as the criterion variables. Demonstrative counts that lacked normal distributions were removed. Six variables were removed because of high skewness and kurtosis (> 3) as a result of numerous zero counts in the data. These zero counts precluded transforming the data. Prior to analysis, the corpus was divided into training and test sets using a 67/33 split (Witten, Frank, & Hall 2011). Using the training set, correlations were then calculated to determine whether there was a statistical ($p < .05$) and meaningful (at least a small effect size, $r > .10$) relation between the demonstrative counts and the human scores. All features were also checked for multicollinearity with other features ($r > .90$). The features that showed statistical and meaningful effects and were not multicollinear were included as predictor variables in a stepwise multiple regression to explain the variance in the human scores. The model from the stepwise regression was then used to predict the variance in the human scores for the essays in the test set.

Correlations

Correlations were conducted between the selected features that did not demonstrate multicollinearity and were normally distributed ($n = 4$). Of these variables, two demonstrated significant ($p < .050$) and positive correlations with the human scores as well as at least small effect size ($r > .100$). These two indices were *unattended this* and *all unattended demonstratives* (see Table 5 for correlations and descriptive statistics) although, if controlling for multiple comparisons ($p < .0125$) only the former would meet the criteria.

Regression

A stepwise regression analysis using the two significant indices as the independent variables to explain the variance in the human scores of essay quality yielded a

Table 5 Descriptive statistics and correlations between selected features and essay quality

Feature	Mean (SD)	*r*	*p*
Unattended this	0.343 (0.376)	0.219	0.004
All unattended demonstratives	0.634 (0.412)	0.173	0.023
Attended these	0.17 (0.279)	0.125	0.103
Unattended that	0.256 (0.373)	−0.065	0.393

Table 6 Stepwise regression analysis and significance values for demonstrative types predicting essay scores

Entry	Index added	*r*	Total R^2	*B*	B	*SE*	*t*
Entry 1	Unattended demonstrative this	0.218	0.048	0.377	0.150	0.218	2.511*

B = unstandardized β; B = standardized; *SE* standard error. Estimated constant term is 2.836

* $p < .05$

significant model, $F(1, 126) = 6.303$, $p < .050$, $r = .218$, $R^2 = .048$. One of the features was included as a significant predictor of the essay scores: *unattended this*. The model demonstrated that the one variable explained 5 % of the variance in the human scores of essay quality for the 127 essays in the training set (see Table 6 for additional information). When the model was applied to the test set, the model yielded $r = .224$, $R^2 = .501$, indicating that the variable explained 5 % of the variance in the human scores for the 46 essays in the test set and that the model is stable.

Discussion

The findings from the essay corpus analysis demonstrate that unattended demonstratives are positively associated with greater essay quality in terms of human ratings supplied by high knowledge, expert readers. This finding counters previous studies that associated unattended demonstratives with greater overall cohesion and linked this increased cohesion to greater essay quality (McCulley, 1985; Witte & Faigley, 1981). It is likely that low sample sizes ($N = 10$ in the case of Witte & Faigley, 1981) and combining attended and unattended demonstratives into the same class (in the cases of McCulley, 1985; Witte & Faigley, 1981) might account for these differences.

In light of Experiment 1, which demonstrated that unattended demonstratives did not lead to faster text processing or greater referent recall when compared to attended demonstratives, we presume that the unattended demonstratives found in the persuasive essays led to greater linguistic complexity in the text (and likely to less cohesive text). As previous studies have demonstrated, human judgments of writing quality are generally positively associated with linguistic complexity. Specifically, writing proficiency is linked with linguistic features that make the text harder to process not easier. For instance, from a lexical perspective, higher quality essays contain longer words (Crossley et al., 2014), less frequent words (Crossley et al., 2011, 2014; McNamara et al., 2010, 2013), less frequent bigrams (Burstein, Chodorow, & Leacock, 2004; Crossley, Cai, & McNamara, 2012) and include greater lexical diversity (Crossley et al., 2011; McNamara et al., 2010, 2013). Higher quality essays also tend to be more syntactically complex. For example, they tend to have more T-units per sentence (i.e., a main clause and dependent clause combined; Yang, Lu, & Weigle, 2015), T-units with more words (Witte & Faigley, 1981), and more clauses per T-unit (Yang et al., 2015). In addition, higher quality

advanced essays tend to have more modifiers per word phrase, more modifiers per noun phrase (Crossley et al., 2011), include more words before the main verb (McNamara et al., 2010), more prepositions, more instances of subordination, more passives (Connor, 1990), and use more negations (Crossley et al., 2014). Importantly, in consideration of text cohesion, higher quality essays tend to include less local cohesion (i.e., cohesion that links short segments of texts like sentences together). For instance, higher quality essays tend to have less lexical overlap between sentences and use fewer connectives used to link sentences (Crossley et al., 2011; Crossley, Kyle, & McNamara, 2015).

When considered within this research framework, the notion that more unattended demonstratives equate to higher quality essays seems reasonable. Indeed, since unattended demonstratives are more difficult to process and lead to greater comprehension difficulties (see Experiment 1), their use would doubtlessly make an essay more complex. In addition, the use of unattended demonstratives would make a text less cohesive and less cohesive text is generally equated with higher writing quality (Crossley et al., 2011; Crossley, Kyle, & McNamara, in press). Thus, the findings from this analysis support the basic notion that the use of more complex linguistic structures leads expert raters to judge an essay as being of higher quality.

Demonstratives and text stylistics

The findings from the second study are problematic from a text stylistics perspective. Historically, the use of unattended demonstratives in writing has been discouraged (American Psychological Association, 2001; Axelrod & Cooper, 2008; Ede, 2004; Faigley, 2007). The reason for this, as discussed earlier, is that the antecedents for unattended demonstratives are more obscure and their use is considered imprecise (Markel, 2004; Strunk & White, 2000). These elements are argued to cause potential ambiguity in writing which can affect both the processing of a text and the effect the writing has on the reader (Swales & Feak, 2000, 2004). However, the complexity found in unattended demonstratives, which likely caused the imprecision in labeling antecedents and slower text processing we saw in Experiment 1, is likely the same complexity that indicates to the reader that the writing is of higher quality.

Thus, while style guides argue that the use of unattended demonstratives should be avoided, this study provides some empirical evidence to counter this notion. When the findings from this study are aligned with evidence supporting the common use of unattended demonstratives in academic writing (Gray & Cortes 2011; Swales, 2005), a case could be made supporting their use as a stylistic feature. In addition, as demonstrated in Table 5, the majority of attended demonstrative features (including the count for all attended demonstratives) showed negative, albeit not significant, correlations with human ratings of essay quality. Thus, the use of attended demonstratives, could negatively impact a reader's impression of an essay such that more cohesive text may actually be judged as lower quality. This contrasts with many reference texts that contend that demonstratives should be followed by nouns (Johnson-Sheehan, 2005; Markel, 2004).

Conclusion

This study combines a traditional psycholinguistic approach to understanding referent identification and processing speed with a corpus-based approach to understanding how discourse features function in a real-world setting. The findings demonstrate that the use of unattended demonstratives as anaphoric references is disadvantageous to both text processing speed and referent identification likely because of the complexity associated with unattended demonstratives. However, these disadvantages become advantages in terms of essay quality because linguistic complexity is a strong indicator of high proficiency writing. From a text processing and comprehension viewpoint, the findings indicate, then, that the use of anaphoric references (i.e., unattended demonstratives) is not always beneficial. This conclusion is predicated on the notion that the while use of personal pronouns as anaphoric references does appear to be beneficial for both text processing and referent identification (Kossowska, 2004), the same does not hold for unattended demonstratives. In contrast, from a writing context, the use of unattended demonstratives leads to a more complex text, which generally equates to a higher quality text. Thus, the adages of stylebooks may fall short, especially within the context of academic writing.

Missing from this study is direct evidence comparing the use of personal pronouns as anaphoric references to demonstrative pronouns. Such a study would further our understanding of how these two types of referents differ. In addition, while the study allows generalizations about how demonstrative pronouns interact with text processing and comprehension, the method used in the study does not allow for direct extension of the findings to longer text segments (such as reading texts or academic writing samples). Additional studies are warranted to investigate how demonstrative pronouns affect processing in other naturalistic settings such as assessments of text readability. Lastly, it is possible that the referent identification task in the first experiment may have biased the reading times for the final sentence because the participants were anticipating the identification task. Clearly, the task itself is an important component of the experiment in order to examine referent identification across conditions. However, as noted earlier, it is unlikely that referent identification task influenced the reading times more so for unattended than for attended demonstratives. Nonetheless, follow-up studies should consider using general comprehension questions or assessing reading times without the presentation of comprehension questions. However, some sort of comprehension question is generally necessary to ensure that the participants attend to the meaning of the sentences.

Acknowledgments This research was supported in part by the Institute for Education Sciences (IES R305A080589 and IES R305G20018-02). Ideas expressed in this material are those of the authors and do not necessarily reflect the views of the IES. We thank our expert raters for their assistance in scoring the essays used in this study.

References

Amaral, M. P. (1985). On the categories of textual cohesion and text complexity. *Letras De Hoje, 18*(60), 29–40.

American Psychological Association. (2001). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: American Psychological Association.

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

Axelrod, R. B., & Cooper, C. R. (2008). *The St Martin's guide to writing* (8th ed.). Boston: Bedford/St Martin's.

Bates, D., Maechler, M., & Bolker, B. (2014). lme4: Linear mixed-effects models using s4 classes.

Burstein, J., Chodorow, M., & Leacock, C. (2004). Automated essay evaluation: The Criterion online service. *AI Magazine, 25*, 27–36.

Buswell, G. T. (1922). *Fundamental reading habits: A study of their development*. Chicago: University of Chicago Press.

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

Connor, U. (1990). Linguistic/rhetorical measures of international persuasive student writing. *Research in the Teaching of English, 24*, 67–87.

Crossley, S. A., Allen, L. K., Kyle, K., & McNamara, D. S. (2014). Analyzing discourse processing using the simple natural language processing tool (SiNLP). *Discourse Processes, 51*(5–6), 511–534.

Crossley, S. A., Cai, Z., & McNamara, D. S. (2012). Syntagmatic, paradigmatic, and automatic n-gram approaches to assessing essay quality. In P. M. McCarthy & G. M. Youngblood (Eds.), *Proceedings of the 25th international Florida Artificial Intelligence Research Society (FLAIRS) conference*. (pp. 214–219). Menlo Park, CA: The AAAI Press.

Crossley, S. A., Kyle, K., & McNamara, D. S. (2015). To aggregate or not? Linguistic features in automatic essay scoring and feedback systems. *Journal of Writing Assessment, 8*(1).

Crossley, S. A., Kyle, K., & McNamara, D. S. (in press). The Tool for the Automatic Analysis of Text Cohesion (TAACO): Automatic assessment of local, global, and text cohesion. *Behavior Research Methods.*

Crossley, S. A., & McNamara, D. S. (2010). Cohesion, coherence, and expert evaluations of writing proficiency. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd annual conference of the cognitive science society* (pp. 984–989). Austin, TX: Cognitive Science Society.

Crossley, S. A., & McNamara, D. S. (2011). Text coherence and judgments of essay quality: Models of quality and coherence. In L. Carlson, C. Hoelscher, & T. F. Shipley (Eds.), *Proceedings of the 29th annual conference of the cognitive science society* (pp. 1236–1241). Austin, TX: Cognitive Science Society.

Crossley, S. A., Roscoe, R., & McNamara, D. S. (2013). Using automatic scoring models to detect changes in student writing in an intelligent tutoring system. In McCarthy, P. M. & Youngblood G. M., (Eds.), *Proceedings of the 26th international Florida Artificial Intelligence Research Society (FLAIRS) conference*. (pp. 208–213). Menlo Park, CA: The AAAI Press.

Crossley, S. A., Weston, J., McLain Sullivan, S. T., & McNamara, D. S. (2011). The development of writing proficiency as a function of grade level: A linguistic analysis. *Written Communication, 28*(3), 282–311.

Davies, M. (2009). The 385+ million word Corpus of Contemporary American English (1990–2008+): Design, architecture, and linguistic insights. *International Journal of Corpus Linguistics, 14*(2), 159–190.

Dell, G. S., McKoon, G., & Ratcliff, R. (1983). The activation of antecedent information during the processing of anaphoric reference in reading. *Journal of Verbal Learning and Verbal Behavior, 22*, 121–132.

Doyle, A. E. (1982). The limitations of Cohesion. *Research in the Teaching of English, 16*(4), 390–395.

Ede, L. (2004). *Work in progress* (6th ed.). Boston: Bedford/St. Martins.

Ehrlich, K., & Rayner, K. (1983). Pronouns assignment and semantic integration during reading: Eye-movements and immediacy of processing. *Journal of Verbal Learning and Verbal Behavior, 22*, 75–87.

Faigley, L. (2007). *Writing: A guide for college and beyond*. New York: Pearson/Longman.

Finn, S. (1995). Measuring effective writing: Cloze procedure and anaphoric "this". *Written Communication, 12*(2), 240–266.

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

Garrod, S., & Sanford, A. (1977). Interpreting anaphoric relations: The integration of semantic information while reading. *Journal of Verbal Learning and Verbal Behavior, 16*, 77–90.
Geisler, C., Kaufer, D., & Steinberg, E. (1985). The unattended anaphoric "this": When should writers use it? *Written Communication, 2*(2), 129–155.
Gernsbacher, M. A. (1989). Mechanisms that improve referential access. *Cognition, 32*, 99–156.
Gray, B., & Cortes, V. (2011). Perception vs. evidence: An analysis of this and these in academic prose. *English for Specific Purposes, 30*(1), 31–43.
Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.
Himmelman, N. P. (1996). Demonstratives in narrative discourse: A taxonomy of universal uses. In B. Fox (Ed.), *Studies in Anaphora* (pp. 205–254). Amsterdam: John Benjamin.
Hudson, S. B., Tanenhaus, M. K. & Dell, G. S. (1986). The effect of discourse center on the local coherence of a discourse. In *Proceedings of the eighth annual conference of the cognitive science society*, (pp. 96–101). Mahwah, NJ: Lawrence Erlbaum Associates
Johnson-Sheehan, R. (2005). *Technical communication today*. New York: Pearson.
Just, M. A., & Carpenter, P. A. (1987). *The psychology of reading and language comprehension*. Boston: Allyn & Bacon Inc.
Karmiloff-Smith, A. (1985). Language and cognitive processes from a developmental perspective. *Language and Cognitive Processes, 1*(1), 61–85.
Kossowska, A. (2004). The anaphora-cataphora switch using pronouns, determinatives, and adverbs. *Studia Linguistica, 23*, 29–48.
Kyle, K., & Crossley, S. A. (2015). Automatically Assessing Lexical Sophistication: Indices, Tools, Findings, and Application. *TESOL Quarterly*, *49*(4), 757–786.
Loxterman, J. A., Beck, I. L., & McKeown, M. G. (1994). The effects of thinking aloud during reading on students' comprehension of more or less coherent text. *Reading Research Quarterly, 29*, 353–367.
Lyons, J. (1977). *Semantics*. Cambridge: Cambridge University Press.
MacDonald, M., & MacWhinney, B. (1990). Measuring inhibition and facilitation from pronouns. *Journal of Memory and Language, 29*, 469–492.
Markel, M. (2004). *Technical communication* (7th ed.). Boston: Bedford/St. Martins.
McCulley, G. A. (1985). Writing quality, coherence, and cohesion. *Research in the Teaching of English, 19*(3), 269–282.
McKoon, G., & Ratcliff, R. (1980). The comprehension processes and memory structures involved in anaphoric reference. *Journal of Verbal Learning and Verbal Behavior, 19*, 668–682.
McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010). The linguistic features of quality writing. *Written Communication, 27*(1), 57–86.
McNamara, D. S., Crossley, S. A., & Roscoe, R. (2013). Natural language processing in an intelligent writing strategy tutoring system. *Behavior Research Methods, 45*(2), 499–515.
McNamara, D. S., & Kintsch, W. (1996). Learning from text: Effects of prior knowledge and text coherence. *Discourse Processes, 22*, 247–288.
McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction, 14*, 1–43.
Mirman, D. (2014). *Growth curve analysis and visualization using R*. New York: Taylor & Francis.
Rayner, K., & Pollatsek, A. (1994). *The psychology of reading*. Englewood Cliffs, NJ: Prentice Hall.
Römer, U., & Wulff, S. (2010). Applying corpus methods to writing research: Explorations of MICUSP. *Journal of Writing Research, 2*(2), 99–127.
Rustipa, K. (2015). The use of demonstrative pronoun and demonstrative determiner *this* in upper-level student writing: A case study. *English Language Teaching, 8*(5), 158–167.
Shapiro, A. M. (2004). How including prior knowledge as a subject variable may change outcomes of learning research. *American Educational Research Journal*, *41*(1), 159–189. doi:10.3102/00028312041001159.
Shillcock, R. (1982). The on-line resolution of pronominal anaphora. *Language and Speech, 25*, 385–402.
Speelman, C. P., & Kirsner, K. (1990). The representation of text-based and situation-based information in discourse comprehension. *Journal of Memory and Language, 29*, 119–132.
Strunk, W., & White, E. B. (2000). *The elements of style* (4th ed.). Boston: Allyn & Bacon.
Swales, J. (2005). Attended and unattended "this" in academic writing: A long and unfinished story. *ESP Malaysia, 11*, 1–15.
Swales, J. M., & Feak, C. B. (2000). *English in today's research world: A writing guide*. Ann Arbor, MI: University of Michigan Press.

Swales, J. M., & Feak, C. B. (2004). *Academic writing for graduate students* (2nd ed.). Ann Arbor, MI: University of Michigan Press.

Webber, B. L. (1991). Structure and ostension in the interpretation of discourse deixis. *Language and Cognitive Processes, 6*, 107–135.

Williams, J. M. (1985). *Style: Ten lessons in clarity and grace* (2nd ed.). Glenview, IL: Scott Foresman.

Witte, S. P., & Faigley, L. (1981). Coherence, cohesion, and writing quality. *College Composition and Communication, 32*(2), 189–204.

Witten, I. A., Frank, E., & Hall, M. A. (2011). *Data mining*. San Francisco, CA: Elsevier.

Wulff, S., Römer, U., & Swale, J. (2012). Attended/unattended this in academic student writing: Quantitative and qualitative perspectives. *Linguistics and Linguistic Theory, 8*(1), 129–157.

Yang, W. W., Lu, X., & Weigle, S. C. (2015). Different topics, different discourse: Relationships among writing topic, measures of syntactic complexity, and judgments of writing quality. *Journal of Second Language Writing, 28*, 53–67.