Applied Psycholinguistics **40** (2019), 373–398
doi:10.1017/S0142716418000619

Native-like processing of prominence cues in L2 written discourse comprehension: Evidence from font emphasis

EUN-KYUNG LEE
Yonsei University

SCOTT FRAUNDORF
University of Pittsburgh

Received: June 18, 2017 Revised: July 27, 2018 Accepted: September 1, 2018

ADDRESS FOR CORRESPONDENCE
Eun-Kyung Lee, Department of English Language and Literature, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 03722, Korea. E-mail: eunkyunglee@yonsei.ac.kr

ABSTRACT
Understanding alternatives to prominent information contributes to successful native language discourse comprehension. Several past studies have suggested that the way second language (L2) learners encode and represent an alternative set in L2 speech is not exactly native-like. However, because these studies involved contrastive pitch accents in running speech, these native language–second language differences may reflect the demands of comprehending running speech in L2 rather than intrinsic deficit in discourse processing per se. Here, we tested L2 learners' discourse encoding and representation using a different cue to prominence: font emphasis in self-paced reading. We found that, in this temporally less demanding modality, L2 learners' encoding of salient alternatives became native-like. Font emphasis facilitated L2 learners' memory for the discourse by ruling out salient alternatives, just as how it facilitates native speakers'. L2 learners were also similar to native speakers in using the situation model to constrain an alternative set. The results suggest that L2 learners can show native-like processing of prominence and that previous underuse of contrastive accents in L2 comprehension could reflect cognitive demands of processing running speech in L2.

Keywords: alternative sets; discourse; font emphasis; memory; second language processing

Successful language comprehension often involves not only understanding discourse in the moment but also in later remembering what is read or spoken about. Thus, understanding the mechanisms underlying second language (L2) processing requires knowing not only how target-language sentences and discourses are processed in real time but also how they are represented in memory. However, the bulk of L2 processing research to date has examined whether and how L2 learners differ from native speakers in online processing of sentences (e.g., Felser & Roberts, 2007; Felser, Roberts, & Marinis, 2003; Frenck-Mestre & Pynte, 1997;

Jackson & Bobb, 2009; Juffs & Harrington, 1996; Lee, Lu, & Garnsey, 2013; Marinis, Roberts, Felser, & Clahsen, 2005; Papadopoulou, 2005; Papadopoulou & Clahsen, 2003; Roberts & Felser, 2011; Williams, 2006) or discourses (Felser & Cunnings, 2012; Felser, Sato, & Bertenshaw, 2009; Hopp, 2009; Pan & Felser, 2011; Pan, Schimke, & Felser, 2015; Patterson, Trompelt, & Felser, 2014; Roberts, Gullberg, & Indefrey, 2008). Little is known about whether an L2 discourse is encoded and represented in memory differently from a native language (L1) discourse, nor why any such differences might exist.

Here, we examine how L2 learners process cues to one particular influence on discourse representation: cues to a contrast between one element and its alternatives. In L1-English processing, prominence cues, such as a contrastive pitch accent in speech or font emphasis in written text, prompt comprehenders to encode a set of alternatives to the prominent element and later rule out those alternatives when remembering the true facts of the discourse (Braun & Tagliapietra, 2010; Calhoun, 2009; Fraundorf, Benjamin, & Watson, 2013; Fraundorf, Watson, & Benjamin, 2010; Reichle & Birdsong, 2014). For example, *you need to push the RED button, not the blue button* helps L1 comprehenders rule out the mentioned alternative *blue button* in their later memory, but not the unmentioned *green button*.

We explore whether L2 learners of English encode and represent an alternative set in their discourse representation in the same way as native speakers do. We also examine what kind of alternatives L2 comprehenders consider in cases where they *do* encode an alternative set in response to prominence cues. In examining these issues, we illuminate not only how L2 learners encode alternatives in discourse but also what kinds of variables constrain L2 discourse processing more broadly.

PROCESSING AND REMEMBERING PROMINENT MATERIAL IN L1 AND L2

The pattern of prominence with a discourse may imply something about how entities in the discourse relate and how they should be remembered. In English spoken sentences, one way a word can be made prominent is by producing it with a *pitch accent*, a phonological construct acoustically characterized by increased duration and intensity and by different patterns of tonal excursion. Namely, many theories of intonation (such as the ToBI system for intonational transcription of English) distinguish multiple types of pitch accents: among others, the H* accent involves a high tonal target on the stressed syllable of an accented word whereas the L+H* accent involves a high tonal target preceded by a steep rise from an initial low tone (Beckman & Ayers, 1997). The L+H* pitch accent has been argued to mark contrastive information whereas H* tends to be associated with information that is new to the discourse (Pierrehumbert & Hirschberg, 1990). The contrastive reading of the L+H* accent is supported by eye-tracking studies, which demonstrate that L1-English listeners encode an alternative set in response to an L+H* accent and use it to anticipate a referent that contrasts with a previously mentioned one (Ito & Speer, 2008; Watson, Tanenhaus, & Gunlogson, 2008; see also Weber, Braun, & Crocker, 2006, for evidence from German). Crucially, contrastive L+H* accents not only guide online processing but also

https://doi.org/10.1017/S0142716418000619 Published online by Cambridge University Press

appear to facilitate memory for discourse by strengthening a distinction between a true statement (i.e., what happened) and its alternatives (i.e., what did not happen) (Fraundorf et al., 2010).

High-proficiency L2 learners also use prominence cues to evoke a contrast between a prominent element and its alternatives. This has been observed with contrastive pitch accents in both L2-Dutch (Braun & Tagliapietra, 2011) and L2-English (Lee & Fraundorf, 2017) as well as with structural prominence cues, such as clefts, in L2-French (Reichle, 2010a; Reichle & Birdsong, 2014). For instance, Reichle and Birdsong (2014) examined English speakers' event-related potential responses while they read cleft sentences in L2-French. Each target sentence (e.g., *It is a hammer that we see on the table*) was paired with a question in which the critical noun in the focus-marking position received either noncontrastive informational focus (e.g., *What do we see on the table?*) or contrastive focus (e.g., *Is it a glass or a hammer that we see on the table?*). Just like native French speakers, high-proficiency L2 learners showed increased anterior negativity, which indexes an increase in working memory load (e.g., Cowles, 2003; Kluender & Kutas, 1993; Lee & Garnsey, 2015; van Berkum, Brown, & Hagoort, 1999), in the contrastive-focus condition as compared to the informational-focus condition. This result was interpreted as indicating that high-proficiency L2 learners evoked a contrast set to the focused word, requiring greater working memory resources to hold those alternatives active.

However, even high-proficiency L2 learners may not show fully native-like behavior in processing prominence cues (Akker & Cutler, 2003; Braun & Tagliapietra, 2011). For example, Braun and Tagliapietra (2011) had German-speaking L2 learners of Dutch listen to Dutch sentences containing a sentence-final prime word (e.g., *flamingo*), which was produced either with a contrastive double-peak contour or with a noncontrastive hat-pattern contour. As soon as the prime word was heard, a visual target word that was either contrastively related (e.g., *pelikaan*, "pelican") or unrelated (e.g., *beroemdheid*, "celebrity") to the prime word was presented. Just like native Dutch speakers, high-proficiency L2 speakers showed a priming effect for contrastively related targets after hearing the prime word with a double-peak contour. However, unlike native speakers, German-speaking L2 speakers showed the same effect even after hearing a noncontrastive hat-pattern contour. Given that both double-peak and hat-pattern contours signal contrast in German, this result suggests that high-proficiency L1-German learners of Dutch immediately encoded a set of alternatives in response to a contrastively focused word in L2 speech, but that in doing so, they transferred their L1 mapping of prosody to contrastive meaning (see Archibald, 1997, for evidence of L1 transfer in the production of L2 prosodic focus).

Further evidence that L2 learners' processing of prominence is not entirely native-like behavior comes from Lee and Fraundorf (2017), who tested L2 learners' memory for prominent information. Lee and Fraundorf had Korean-speaking L2 learners of English listen to short discourses, such as (1) below, each of which established two alternative sets (e.g., *British–French*, *Malaysia–Indonesia*) before referring to one referent from each set, underlined in (1) for illustration purposes only. Later, participants' memory for those critical words

https://doi.org/10.1017/S0142716418000619 Published online by Cambridge University Press

was tested using a probe recognition task in which participants judged statements as *true* or *false*. Three different types of probes appeared: correct statements ([2a]), false statements about the contrastive alternative ([2b]), and false statements about an unmentioned item ([2c]).

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

(2) a. The British scientists found the endangered monkey.
b. The French scientists found the endangered monkey.
c. The Portuguese scientists found the endangered monkey.

The critical words in the original discourse were manipulated for pitch accent type (contrastive L + H* vs. noncontrastive H*). Lee and Fraundorf evaluated three different hypotheses about how contrastive accents might benefit L2 learners' memory. The *granularity account* proposes that contrastive accents improve memory by enhancing the representation of the prominent item itself, such as *British* in (1) above; this would enhance rejections of both kinds of false statements, both of which are inconsistent with *British*. By comparison, the *contrast representation account* proposes that the effect of the contrastive L + H* accent over the noncontrastive H* is to lead comprehenders to remember something about the specific salient alternative in the original discourse (i.e., that it was *not* the French who found the monkey), which would enhance rejections only of the alternative probes and not of the unmentioned probes. Finally, the *shallow representation account*[1] proposes that even when an L + H* accent evokes a set of contrasting alternatives (e.g., the set of *British* and *French*), a lack of cognitive resources for L2 processing makes it difficult for L2 learners to fully integrate into their representation of the discourse which member of the alternative set is the correct item and which is the contrast item. Thus, this account predicts that the contrastive accent would not facilitate rejections of the alternative probes and might even impair them. However, having brought to mind the contrast between the two items in the alternative set (*British* and *French*) should enhance a distinction between those items and an unmentioned item (e.g., *Portuguese*), facilitating rejections of the unmentioned statements.

Native English speakers' behavior appears to be described by the contrast representation account; that is, the L + H* accent facilitates rejections exclusively of the alternative statements (Fraundorf et al., 2010). However, high-proficiency Korean-speaking L2 learners of English showed different effects: the L + H* accent did not help them reject the contrastive alternative; rather, it *impaired* rejections of the contrastive alternative. For these learners, the L + H* accent facilitated rejections of the unmentioned items, consistent with the shallow representation account. In sum, although a contrastive pitch accent improved memory for L2 discourse to some extent by enhancing a distinction between elements inside and outside an alternative set, it failed to lead L2 learners to encode in memory the salient alternative as fully as native speakers do.

https://doi.org/10.1017/S0142716418000619 Published online by Cambridge University Press

EXPLAINING L1-L2 DIFFERENCES IN DISCOURSE PROCESSING

The results reviewed above suggest that at least some aspects of L2 discourse processing and memory may not function in a native-like manner. What might underlie these L1-L2 differences?

One possibility is that L2 learners show non-native processing of pitch accents because they fail to acquire the mapping between a specific pitch accent type in their L2 and its discourse function. This could well explain the Braun and Tagliapietra (2011) results, in which L1-German learners of Dutch appeared to inappropriately transfer their L1 prosodic knowledge. Similarly, it is possible that the absence of pitch accenting as a focus-marking device in Seoul Korean, which marks contrastive focus prosodically by placing an accentual phrase boundary before a focused word and dephrasing words following it (Jun, 1993, 2005; Ladd, 2008), may have prevented the L1-Korean learners of English in Lee and Fraundorf (2017) from grasping the discourse function of contrastive pitch accents in L2 speech. More generally, even if there is no negative transfer from L1, it is possible that L2 discourse processing may be intrinsically disadvantaged if it qualitatively differs from L1 processing, as has sometimes been proposed for other aspects of L2 processing, such as syntax (e.g., Clahsen & Felser, 2006a, 2006b).

Another possibility, however, is that even if L2 learners have knowledge of a particular cue to discourse processing, for instance, that the L+H* pitch accent has a contrastive meaning, the availability of cognitive processing resources may constrain whether they actually incorporate that knowledge into their discourse representation in a native-like way. L2 processing is known to be more resource demanding than L1 processing (McDonald, 2006; Segalowitz, 2003), and a lack of cognitive resources may lead to non-native performance in various aspects of L2 development (e.g., Ardila, 2003; Harrington, 1992; Hopp, 2010; Kilborn, 1992; Sorace & Serratrice, 2009). Furthermore, even native speakers' performance may become similar to that of non-native speakers under temporally demanding conditions (e.g., Hopp, 2010; McDonald, 2006). That is, the fleeting nature of running speech may require L2 listeners to devote a majority of their cognitive resources to processing the segmental content, leaving insufficient resources for fully processing prominence cues. Thus, for instance, the L2 learners in Lee and Fraundorf (2017) may not have had sufficient opportunity to fully integrate into their discourse representation which item was the contrastive alternative.

There are many reasons to think that cognitive resource constraints may be one source of L2 learners' non-native-like processing of prominence cues in spoken discourse. First, processing a prosodic cue itself is likely computationally intense because it requires integrating information across different levels of linguistic representation such as syntax or discourse structure (Dekydtspotter, Donaldson, Edmonds, Fultz, & Petrusch, 2008; Dekydtspotter, Edmonds, Fultz, & Renaud, 2010). Second, there is evidence that it is time consuming and resource demanding to encode salient alternatives. L1 reading times suggests that although words marked as prominent are better remembered, they also require longer

https://doi.org/10.1017/S0142716418000619 Published online by Cambridge University Press

processing times in initial discourse processing (Fraundorf et al., 2013). In addition, both native speakers and high-proficiency L2 learners show event-related potential signatures of increased working memory load in conditions that evoke contrastive alternatives (Cowles, 2003; Reichle, 2010a; Reichle & Birdsong, 2014). Lastly, working memory capacity constrains the effects of prominence cues even for L1 comprehenders (Fraundorf, Watson, & Benjamin, 2012).

In this study, we test whether L2 learners are intrinsically disadvantaged in comprehending contrast in L2 discourse or whether non-native-like processing partially reflects online processing demands. We do so by testing whether L2 learners exhibit similar non-native processing and memory with a less temporally demanding cue to prominence: font emphasis when reading written text. Compared to listening, reading has been argued to require fewer computational resources for low-level processes, such as phoneme decoding and word recognition, because readers can take as long as desired (Almor & Eimas, 2008). L2 learners tend to comprehend written text better than speech (Reves & Levine, 1988). If L2 learners' non-native processing of spoken prominence reflects online processing demands, an absence of time pressure should leave more resources for drawing interferences about salient alternatives. By contrast, if there are intrinsic deficits in comprehending a discourse in L2, then a change of modality or presentation speed should be irrelevant. Thus, examining L2 learners' processing of *written* discourse can reveal the generality and driving factors of L2 learners' non-native discourse processing: can L2 learners fully specify in memory, as native speakers do, which is the salient alternative under less cognitively demanding conditions?

CONSTRAINING AN ALTERNATIVE SET

Thus far, we have considered *whether* L2 comprehenders encode and represent a set of alternatives in response to prominence cues. A related but separate question is *how* a set of alternatives, *if* encoded, is constrained. That is, what makes something constitute one of the salient alternatives worthy of remembering as part of a discourse? This question has been examined in L1 processing (e.g., Byram Washburn, Kaiser, & Zubizarreta, 2011; Fraundorf et al., 2013), but much less is known about how an alternative set is constrained in L2 comprehension and whether it proceeds based on the same constraints as for native speakers.

Theories of discourse (e.g., Chafe, 1974; Prince, 1981) delineate multiple constraints on the salience or activation of referents. Here, we focus on one potential constraint: plausibility as an alternative. In discourse (1) above, the intended set of alternatives is fairly unambiguous: not only are the *British* and *French* scientists the only groups searching for the monkey, they are the only scientists mentioned in the discourse. Now consider discourse (3) below, which establishes a contrastive pair (e.g., new *Mexican* and *Indian* restaurants) but also mentions another, noncontrastive item (e.g., an existing *Italian* restaurant) before referring to one of the contrasting items, the *Mexican* restaurant. Although both *Indian* and *Italian* are members of the same semantic category as *Mexican* and

https://doi.org/10.1017/S0142716418000619 Published online by Cambridge University Press

were established in the discourse, only *Indian* is contextually plausible as an alternative to *Mexican* (i.e., the critic could have visited the Indian restaurant instead, but not the Italian restaurant). If mere mention is enough to establish an item in the alternative set, then emphasizing *Mexican* should lead comprehenders to remember both *Indian* and *Italian* as salient alternatives. By contrast, if the set of alternatives is constrained based on contextual plausibility, then only *Indian* should be remembered as a salient alternative.

(3) A new Mexican and a new Indian restaurant had recently opened. Both were waiting to hear whether or not the local food critic would like their desserts and specials. They were nervous because the critic was notoriously harsh and disliked even the popular entrees at the local Italian restaurant. The critic originally planned to dine at both new restaurants during the week. But he caught a cold and could only visit the Mexican restaurant, where he awarded the specials a favorable review.

Prior work suggests that L1 comprehenders constrain the set of alternatives based on contextual plausibility (Byram Washburn et al., 2011; Fraundorf et al., 2013), rather than semantic category membership (e.g., Blok & Eberle, 1999) or prior mention (Fraundorf et al., 2013). For example, Fraundorf et al. (2013, Experiment 3 tested memory for discourses like (3) and found that, for native English speakers, emphasizing the critical word (e.g., Mexican) in the continuation of the original discourse facilitated later rejections only of the contrast item, and not of the merely mentioned item. This finding suggests that the set of alternatives encoded in response to prominence was narrowly constrained to include only alternatives plausible in the situation model of the discourse (Zwaan & Radvansky, 1998) rather than every referent mentioned (see also Byram Washburn et al., 2011).

However, it is yet unclear whether L2 comprehenders constrain alternative sets using the same or different criteria as native speakers. Lee and Fraundorf (2017) provide some general evidence that Korean-speaking L2 learners of English do not rely on semantic category membership in defining alternative sets: a contrastive accent facilitated rejections of an unmentioned item (e.g., *Portuguese*) but not of a contrastive alternative (e.g., *French*) even though both were members of the same semantic category as the correct item (e.g., *British*). However, because only a plausible alternative was mentioned in the original discourse, their design does not reveal whether L2 learners restricted this alternative set based on plausibility as an alternative or on mere mention in discourse. We distinguish these possibilities in the current work using a different set of test probes that separately test a plausible alternative and a merely mentioned item.

PRESENT STUDY

In the present study, we test how a prominence cue affects memory for written discourse among L1-Korean speakers with moderate or high proficiency in

https://doi.org/10.1017/S0142716418000619 Published online by Cambridge University Press

English. Our goals were (a) to determine what underlies L2 learners' non-native encoding of salient alternatives in spoken language comprehension by examining whether this phenomenon generalizes to written text and (b) to test how L2 learners constrain a set of alternatives.

Self-paced reading of written text allows comprehenders to slow down as much as necessary to process the discourse. Thus, if processing limitations underlie L2 learners' non-native-like processing of a prominence cue in speech, our self-paced reading paradigm should make it easier for L2 learners to process the literal text, and as a result, allow them to have more resources for representing and remembering salient alternatives to prominent information. Consequently, they should obtain a more native-like discourse representation; namely, prominence should facilitate L2 learners' rejections specifically of the contrastive alternative on the memory test, as seen with L1 learners. By contrast, if L2 learners' non-native performance stems from more intrinsic deficits in L2 discourse processing, prominence cues would affect L2 learners' memory differently from native speakers even in self-paced reading. Specifically, to mark prominence in written text, we use font emphasis.[2] Fraundorf et al. (2013) have demonstrated that emphasizing words with capitals or italics has similar effects on later memory as contrastive pitch accents. Here, we use capitalization exclusively (see also Fraundorf et al., 2013, Experiment 3) to make our study comparable to Lee and Fraundorf (2017) insofar as each involves prominence cues unavailable in L1-Korean: while italicization is available for changing the appearance of letters in Korean, the Korean writing system does not distinguish between lowercase and uppercase letters.

The use of written text also provides an additional check on the role of processing limitations in L2 discourse memory. In order for this account to be plausible, L2 learners, like native speakers, should take longer to process emphasized words than nonemphasized words. Otherwise, it is implausible that the processing-time demands of emphasized words account for any L1-L2 differences.

If L2 learners are able to encode a set of alternatives, a second goal of the present study is to query how that set is constrained. Native speakers appear to construct a set of alternatives based on plausibility in the situation model (Byram Washburn et al., 2011; Fraundorf et al., 2013); if L2 learners do the same, font emphasis should affect rejections *only* of the contrast items, not of the merely mentioned items. By contrast, if font emphasis helps L2 learners reject any type of item mentioned in the discourse, it would suggest that L2 learners define the alternative set based on mere mention in the discourse. Investigating how L2 learners define an alternative set may thus further indicate whether L2 learners differ from native speakers in the way they use prominence cues to encode and represent a discourse in memory.

METHOD

Participants

Sixty native Korean speakers participated in the experiment for payment. All participants had learned English as an L2 and resided in Seoul at the time of

https://doi.org/10.1017/S0142716418000619 Published online by Cambridge University Press

testing. None of the participants had stayed in an English-speaking country for longer than 6 months before the age of 15.

We assessed participants' English proficiency using the same multiple-choice test as in Lee and Fraundorf (2017). In the test, participants chose one of three response alternatives for each of 40 blanks in a passage. The participants also self-rated their English proficiency on a 6-point scale (6 = *native-like*), in reading, writing, speaking, and listening. The proficiency test had moderate reliability (Kuder–Richardson coefficient of reliability = .65), and each participant's test score was significantly correlated with the average self-rating, $r = .46$, F (1, 58) = 15.4, $p < .01$.

Based on a previously used criterion (Lee & Fraundorf, 2017), 7 participants whose multiple-choice test scores were less than 29 (out of 40) were considered as "low proficiency" and were excluded from analysis. We excluded these participants to reduce the possibility that any differences we observed in participants' memory performance simply reflected an inability to process the literal text. Four additional participants were excluded due to technical problems. This left 49 L1-Korean speakers with moderate or high proficiency in English (10 males, 39 females) for analysis. Table 1 summarizes their biographical information, proficiency test scores, and self-rating scores.

Materials

In the study phase, participants read 36 written discourses (available in Appendix C of Fraundorf et al., 2013). Each discourse began with a *context passage* that (a) established two pairs of contrasting alternatives and (b) introduced another *merely mentioned* item for each pair of alternatives. This merely mentioned item was from the same semantic category but was mentioned in a context that made it an implausible alternative for the contrasting pair. For example, in (3) above, the

Table 1. *Summary of the participants' biographical information (means and standard deviations)*

	Mean	SD
Age	23.1	2.84
Age of first exposure	8.3	1.80
Months in English-speaking countries	2.4	4.22
Percentage of daily English use	20.6	23.58
Proficiency test score (/40)	33.6	2.91
Self-rating (/6):		
Reading	4.3	0.70
Writing	3.4	0.84
Speaking	3.3	0.82
Listening	3.9	0.92

Note: The percentage of daily English use was self-computed by each participant relative to all daily language activities including reading.

https://doi.org/10.1017/S0142716418000619 Published online by Cambridge University Press

food critic plans to visit both the *Mexican* and the *Indian* restaurants (one pair of contrasting alternatives) but is not planning to visit the *Italian* restaurant because he has already reviewed it (a merely mentioned item). Similarly, the critic plans to review the *desserts* and *specials* (another pair of contrasting alternatives) but has already reviewed *entrees* (the corresponding merely mentioned item). The inclusion of the merely mentioned item allows us to test how an alternative set is defined in L2 processing; does it include all mentioned referents, or just the alternatives that are plausible in the discourse?

The context passage was then followed by a *continuation passage* that critically referred to one item from each contrasting pair; for example, in (3), the critic visited only the *Mexican* restaurant (but not the Indian restaurant) and reviewed the *specials* (but not the desserts). Half of the critical words received font emphasis and half did not. The assignment of words to emphasis conditions was counterbalanced across experimental lists.

The written discourses were constructed so that the two critical words in the continuation (e.g., *Mexican* and *specials*) never appeared as the first or last word in a punctuation-marked clause because clause boundaries have a large effect on reading times and could overwhelm the effects of interest (Reichle, Warren, & McConnell, 2009). In addition, the two critical words never appeared as the first or last word in a line of text when the stimuli were displayed in the self-paced reading task because line breaks could similarly distort reading times (Rayner, 1998).

Two other controls concerned how we established the two contrasting alternatives in the context passage. If the first of these two items was always the one mentioned in the continuation passage, readers might be able to predict the outcome of the continuation even before reading it. Thus, the stimuli were balanced so that in half of the stimuli, the critical word mentioned in the continuation was the one that appeared first in the context passage (e.g., *Mexican* appears before *Indian* in [3]), and in half the critical word appeared second in the context passage. Similarly, the two contrasting alternatives within a pair (e.g., *desserts* and *specials*) were also matched in orthographic length so that the reader could not anticipate which would be mentioned in the continuation passage simply based on the width of the mask in the self-paced moving window task.

For the subsequent recognition memory test, two sets of memory probes were constructed for each discourse. One set of probes tested one of the critical words (e.g., *Mexican*) and one tested the other (e.g., *specials*), such as (4) and (5). For each critical word, three possible probes were presented by varying one word in the probe sentence to refer to either the correct item (true), the contrastive alternative (false), or the merely mentioned item (false).

(4) a. Because the food critic caught a cold, he only visited the Mexican restaurant. (correct probe)
b. Because the food critic caught a cold, he only visited the Indian restaurant. (contrast probe) c. Because the food critic caught a cold, he only visited the Italian restaurant. (merely mentioned probe)

(5) a. The food critic gave a favorable review to the specials at one of the new restaurants. (correct probe)
b. The food critic gave a favorable review to the desserts at one of the new restaurants. (contrast probe)
c. The food critic gave a favorable review to the entrees at one of the new restaurants. (merely mentioned probe)

Each participant saw only one of the three memory probes for a given fact; the assignment of items to probe types was counterbalanced across lists and was fully crossed with whether the critical word received font emphasis at the time of study. (No font emphasis ever appeared in the memory probes.) This resulted in a $2 \times 3 \times 2$ design of font emphasis × probe type × group (L1 vs. L2), with the first two variables varying within subjects and the last between subjects.

Our primary interest was comparing the two types of false probes: the contrastive alternative (e.g., *Indian* in the above example) and the merely mentioned item (e.g., *Italian*). To ensure that this comparison was independent of lexical properties such as frequency or imageability, we counterbalanced across experimental lists which word was the contrastive alternative and which was the merely mentioned item (i.e., a different participant would instead see *Italian* as the contrastive alternative and *Indian* as the merely mentioned item). As our primary interest was in participants' ability to reject the false statements, we did not counterbalance the word used as the correct probe in order to avoid a combinatorial explosion of lists.

Procedure

The experiment consisted of a study phase followed by a test phase. Participants were instructed in English that they would read 36 stories and that their memory of these stories would be tested after reading all of them; the exact format of the memory test was not described in advance. In the study phase, participants read the stories one word at a time at their own pace, in a noncumulative moving-window self-paced reading paradigm. Each trial began with the first word displayed at the left side of the screen and with the other words masked by lines matched in length. Pressing the space bar caused the next word to be displayed and the previous word to be reverted to a line. There was no visual separation between the context passage and the continuation passage of each discourse. The order of the stories was randomized.

The test phase was administered after participants had read all of the 36 stories. Participants saw 72 statements about the stories that they had read (2 statements about each story). The participants' task was to judge whether each probe was true or false based on what they had read; participants were instructed to choose "false" if they thought that any part of the statement was false. The test probes were presented one at a time; after the participant made the "true" or "false" judgment, the next probe was immediately presented. The test probes were presented in a different random order.

RESULTS

To examine whether and how font emphasis affects L2 learners' memory differently from that of native speakers, we compared our L2 group directly to the native speaker group from Fraundorf et al. (2013) for accuracy of recognition memory. We also analyzed initial reading time to test whether L2 learners, like L1 comprehenders, require additional online processing time for words marked as prominent. All models[3] reported below were fit in the R environment for statistical computing using the *lme4* package (Bates, Maechler, & Bolker, 2012).

Initial reading time

We analyzed reading times to examine whether emphasized words were processed differently from nonemphasized words when initially reading the discourse. Reading times were analyzed in a linear mixed-effects model (Baayen, Davidson, & Bates, 2008) with font emphasis (capitals vs. no capitals), region (target word vs. spillover), group (L1 vs. L2), and their interactions as fixed effects. To correct positive skew, the dependent measure was computed as log-transformed reading times on the target and spillover words. The font emphasis and region variables were dummy coded to test their effects relative to the reference level (the nonemphasized critical word). The group variable was coded using mean-centered contrast coding. In all models of log reading times reported below, we included as random effects by-participant random intercepts and slopes for font emphasis, region, and their interaction and by-item random intercepts and slopes for font emphasis and region. A by-item random slope for the interaction between font emphasis and region was excluded because it did not improve the model fit ($p > .05$). Figure 1 displays reading times on the target and spillover words as a function of font emphasis, separately for native English speakers (Fraundorf et al., 2013) and for L1-Korean learners of English. Table 2 presents parameter estimates for the model of log reading times for both L1 and L2 comprehenders.

There was a main effect of group on online reading time ($t = -3.68$), indicating that L2 learners read the target and spillover words more slowly than native

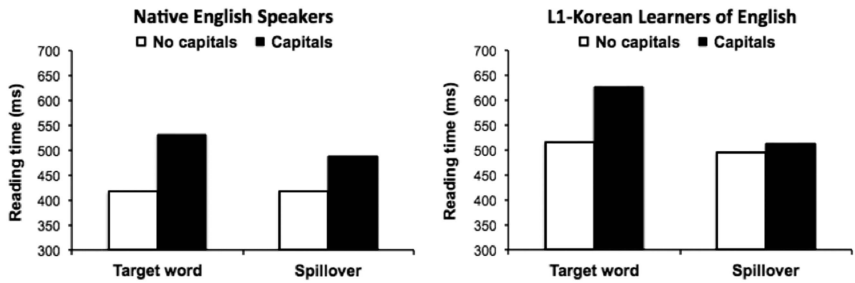

Figure 1. Mean reading time on target words and spillover words as a function of font emphasis for native English speakers (Fraundorf et al., 2013) and L2 learners.

https://doi.org/10.1017/S0142716418000619 Published online by Cambridge University Press

Table 2. *Fixed effect estimates for the model of log reading times for both L1 and L2 comprehenders (N = 13,968, log-likelihood: –6909)*

Fixed effects	β	*SE*	*t*
Nonemphasized critical word (baseline)	6.01	0.03	208.05
Emphasized word	0.15	0.02	9.53
Spillover region	–0.01	0.01	–0.84
Group	–0.21	0.06	–3.68
Emphasized Word × Spillover Region	–0.07	0.02	–3.53
Emphasized Word × Group	–0.01	0.03	–0.24
Spillover Region × Group	0.04	0.02	1.88
Emphasized Word × Spillover Region × Group	0.08	0.04	2.14

Note: SE, standard error. *t*-values greater than 2 are treated as significant (Baayen, 2008).

Table 3. *Fixed effect estimates for the model of log reading times for L2 comprehenders (N = 7,056, log-likelihood: –3482)*

Fixed effects	β	*SE*	*t*
Nonemphasized critical word (baseline)	6.12	0.04	173.73
Emphasized word	0.16	0.02	8.34
Spillover region	–0.03	0.02	–1.35
Emphasized Word × Spillover Region	–0.11	0.02	–4.79

Note: SE, standard error. *t*-values greater than 2 are treated as significant (Baayen, 2008).

English speakers ($t = -3.68$). The effect of font emphasis was significant ($t = 9.53$) such that capitalized critical words ($M = 541$ ms) were read more slowly than those with no capitals ($M = 462$ ms). There was a three-way interaction among font emphasis, region, and group ($t = 2.14$), indicating that the effect of font emphasis on the spillover region varied across groups.

To decompose this three-way interaction, we analyzed just the L2 learners' log reading times in a model with font emphasis (capitals vs. no capitals), region (target word vs. spillover), and their interactions as fixed effects. Table 3 displays parameter estimates for the model of log reading times just for L1-Korean learners of English.

Like native English speakers, L2 learners read the target word more slowly as a function of font emphasis. However, while the effect of font emphasis persisted beyond the target word in native English speakers, the same effect was limited to the target word in L2 learners, as revealed by a significant interaction between font emphasis and region ($t = -4.79$). The presence of font emphasis led L2 learners to stay longer on the target word compared to its absence, and this effect disappeared by the spillover region.

https://doi.org/10.1017/S0142716418000619 Published online by Cambridge University Press

In summary, L2 learners required more time to process emphasized words, despite the emphasized words being the same lexical items in the same syntactic context. This makes it at least *possible* that, in scenarios where L2 learners do not achieve native-like processing of prominent words in spoken language comprehension (e.g., Braun & Tagliapietra, 2011; Lee & Fraundorf, 2017), it is because L2 learners are less well equipped to deal with the time demands of processing prominent words while listening to running speech. If so, we should see more native-like effects on memory in the current self-paced reading task, which allowed L2 learners as much time with the font-emphasized words as needed.

Recognition memory

Analytic strategy. In the memory test, the accurate response is "true" for the correct probes whereas it is "false" for the contrast and merely mentioned probes. Thus, assessing memory performance using proportion accuracy for each probe type confounds participants' *sensitivity* to the truth value of particular probes with their *response bias*, or overall preference (or dispreference) to respond "true." As in prior studies using this paradigm (Fraundorf et al., 2010, 2013; Lee & Fraundorf, 2017), we thus adopted a detection-theoretic analysis, which allows sensitivity to be distinguished from response bias both conceptually and empirically (Green & Swets, 1966; Macmillan & Creelman, 2005; see also Murayama, Sakaki, Yan, & Smith, 2014; Wright, Horry, & Skagerberg, 2008, for applications of detection theory to multilevel models). In this analysis, the proportion of "true" responses (rather than the proportion of accurate responses) is used as the dependent variable. Response bias is then reflected in the baseline rate of "true" responses whereas sensitivity to the truth value of the probes (i.e., veridical memory for the discourse) is reflected by a higher rate of "true" responses to correct probes than to the false probe types. Variables that enhance memory sensitivity are reflected in significant interactions between those variables and the probe type.

The model of "true" responses included font emphasis (capitals vs. no capitals), probe type (correct vs. contrast vs. merely mentioned), group (L1 vs. L2), and their interactions as fixed effects. To examine whether participants' later memory was modulated by online reading time, initial reading time, and its interaction with the other variables were also included in the model as fixed effects. Specifically, we used residual reading time (Ferreira & Clifton, 1986) rather than raw reading time to prevent the effect of font emphasis from being confounded with inter-subject differences in reading speed or inter-item differences in word length. The predicted reading time was computed using a regression model predicting log-transformed reading time based on word length and each participant's baseline reading speed (across all words in the materials). Residual reading time was then computed as a difference between the predicted reading time and the actual reading time. Fraundorf et al. (2013) included residual reading time averaged over the target and spillover words in the model of native speakers' memory because initial reading time on both the target and spillover words was modulated by font emphasis during L1 reading. By contrast, for L2

https://doi.org/10.1017/S0142716418000619 Published online by Cambridge University Press

Table 4. *Fixed effect estimates for the model of "true" responses for both L1 and L2 comprehenders (N = 6,911, log-likelihood: –3740)*

Fixed effects	β	*SE*	Wald *z*	*p*
Baseline rate of "true" responses (response bias)	< 0.01	0.08	0.04	>.1
Emphasis (effect on response bias)	0.01	0.06	0.10	>.1
Reading time (effect on response bias)	< 0.01	0.09	0.05	>.1
Rejections of contrast probes versus baseline (sensitivity)	0.71	0.07	9.70	< .01
Rejections of merely mentioned probes versus baseline (sensitivity)	0.98	0.09	11.40	< .01
Group (effect on response bias)	–0.11	0.08	–1.48	>.1
Emphasis × Reading Time (effect on response bias)	–0.12	0.18	–0.63	< .1
Emphasis × Rejections of Contrast Probes (effect on sensitivity)	0.27	0.08	3.33	< .01
Emphasis × Rejections of Merely Mentioned Probes (effect on sensitivity)	>0.01	0.08	–0.01	>.1
Reading Time × Rejections of Contrast Probes (effect on sensitivity)	–0.04	0.13	–0.33	>.1
Reading Time × Rejections of Merely Mentioned Probes (effect on sensitivity)	0.15	0.13	1.17	>.1
Emphasis × Group (effect on response bias)	0.02	0.12	0.15	>.1
Reading Time × Group (effect on response bias)	–0.14	0.19	–0.74	>.1
Rejections of Contrast Probes × Group (effect on sensitivity)	0.26	0.10	2.63	< .01
Rejections of Merely Mentioned Probes × Group (effect on sensitivity)	0.18	0.11	1.56	>.1
Emphasis × Reading Time × Rejections of Contrast Probes (effect on sensitivity)	–0.44	0.25	–1.77	<.08
Emphasis × Reading Time × Rejections of Merely Mentioned Probes (effect on sensitivity)	0.26	0.25	1.02	>.1
Emphasis × Reading Time × Group (effect on sensitivity)	–0.38	0.37	–1.02	>.1
Emphasis × Rejections of Contrast Probes × Group (effect on sensitivity)	0.09	0.16	0.53	>.1
Emphasis × Rejections of Merely Mentioned Probes × Group (effect on sensitivity)	0.22	0.18	1.30	>.1
Reading Time × Rejections of Contrast Probes × Group (effect on sensitivity)	–0.28	0.25	–1.13	>.1
Reading Time × Rejections of Merely Mentioned Probes × Group (effect on sensitivity)	0.41	0.25	1.59	>.1
Emphasis × Reading Time × Rejections of Contrast Probes × Group (effect on sensitivity)	–0.89	0.50	–1.77	<.08
Emphasis × Reading Time × Rejections of Merely Mentioned Probes × Group (effect on sensitivity)	0.36	0.51	0.72	>.1

Note: SE, standard error.

readers, font emphasis influenced initial reading time just on the target word alone (as discussed above); thus, we included residual reading time on the target word as our reading-time measure in the model of L2 learners' memory. In 1.1%

https://doi.org/10.1017/S0142716418000619 Published online by Cambridge University Press

of the L2 trials, residual reading time for the target word was more than 3 *SD* above or below the participant mean. Those trials were excluded from further analysis.

The font emphasis variable was coded using mean-centered contrast coding to yield tests equivalent to the main effects in an analysis of variance. The probe type variable was coded using effect coding to perform two planned comparisons testing participants' sensitivity in rejecting false statements. In the first comparison, the rate of "true" responses to the contrast probes was compared to the overall (mean) rate of "true" responses. The second comparison tested the rate of "true" responses to the merely mentioned probes relative to the overall rate of "true" responses. In both comparisons, the variable was coded such that a positive coefficient indicated more correct rejections. Residual reading time was mean-centered to assess the effect of the other variables at an average reading time for the target word.

The models of "true" responses reported below included by-participant and by-item random intercepts, and by-participant and by-item random slopes for probe type as random effects: random slopes for the other variables and their interactions with probe type did not significantly improve the model fit (all $ps > .05$).

Figure 2 presents the mean rate of "true" responses as a function of font emphasis and probe type, separately for the native speaker group (Fraundorf et al., 2013) and the L2 group. Table 4 presents parameter estimates for the model of "true" responses for both L1 and L2 comprehenders.

When both L1 and L2 comprehenders were considered, the rate of "true" responses to each of the two types of false probes was significantly lower compared to the overall rate of "true" responses (contrast probes: $z = 9.70$, $p < .01$; merely mentioned probes: $z = 11.40$, $p < .01$), indicating that participants could successfully distinguish the false statements from the correct statements. There was also a significant interaction between group and rejections of contrast probes ($t = 2.63$, $p < .01$). L2 learners made fewer correct rejections of contrast probes than native speakers, indicating that L1-Korean learners of English were less successful at rejecting probes referring to salient alternatives. However, there

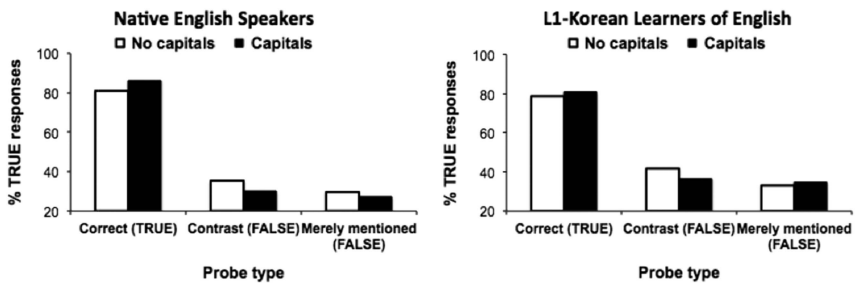

Figure 2. Mean rate of "true" responses as a function of font emphasis and probe type for native English speakers (Fraundorf et al., 2013) and L2 learners.

https://doi.org/10.1017/S0142716418000619 Published online by Cambridge University Press

was no significant interaction between group and rejections of merely mentioned probes, revealing that the groups of participants did not significantly differ in their ability to reject items merely mentioned in the original discourse.

The critical tests here are whether and how rejections of the false probes were affected by font emphasis and whether the effect varied across groups. The results showed that font emphasis facilitated rejecting the contrast probes, collapsed across L1 and L2 groups ($z = 3.33$, $p < .01$). Crucially, the effects of font emphasis on rejections of the contrast probes were similar between L1 and L2 groups, as revealed by the absence of a significant interaction among font emphasis, rejections of alternative probes, and group ($z = 0.53$, $p > .1$).

Moreover, for both L1 and L2 learners, the effect of font emphasis was specific to rejecting the contrast items. There was no effect of font emphasis on rejections of the merely mentioned probes, nor was there a significant interaction with group (all $ps > .1$). There was also no effect of initial reading time nor were there significant interactions (all $ps > .1$).

Further, the same pattern of results (i.e., an effect of font emphasis only in rejecting contrast probes) was observed when the analysis was run on just the L2 group (see Appendix A for parameter estimates), confirming that the L2 group exhibited the same pattern as the L1 comprehenders from Fraundorf et al. (2013).

These data suggest that, unlike contrastive accents, font emphasis led L2 learners to represent a discourse similarly to native speakers, helping them reject exclusively the contrastive alternatives and not the merely mentioned items. The finding that a mnemonic benefit of font emphasis was restricted to rejections of only alternatives that were plausible within the discourse suggests that L2 learners distinguished a salient alternative from an implausible alternative based on a situation model just like native speakers do.

DISCUSSION

Successful language comprehension often requires not only understanding in the moment but also long-term retention of what is heard or read. For native speakers, this process may be enhanced by various cues in the discourse, such as prominence cues that imply the importance of a contrastive alternative (e.g., Braun & Tagliapietra, 2010; Byram Washburn et al., 2011; Fraundorf et al., 2010, 2013; Husband & Ferreira, 2016).

Can L2 learners make similar use of these cues to go beyond what is explicitly stated in a discourse and to retain in memory what is inferred? Several past studies have suggested that, highly proficient L2 learners evoke salient alternatives in response to prominence cues in L2 discourse (e.g., Braun & Tagliapietra, 2011; Lee & Fraundorf, 2017; Reichle & Birdsong, 2014). However, the way they process L2 contrastive pitch accents does not converge with that of native speakers (e.g., Akker & Cutler, 2003; Baker, 2010; Braun & Tagliapietra, 2011). For example, L2 learners fail to integrate salient alternatives into memory as fully as native speakers do and to later rule them out when remembering the content of the discourse (Lee & Fraundorf, 2017). In the present

https://doi.org/10.1017/S0142716418000619 Published online by Cambridge University Press

study, we tested whether such non-native behavior in encoding salient alternatives is intrinsic to L2 discourse processing or whether it may stem from cognitive resource limitations. To do so, we tested whether L2 learners, specifically, L1-Korean speakers with moderate or high proficiency in English, would be able to use a prominence cue to enhance memory in a more native-like way in a task in which online processing demands are reduced compared to listening; namely, self-paced reading, which allows as much time as needed to process each word.

We found two sources of evidence for a role of processing constraints in comprehending prominence in L2 discourse. First, L2 learners, like L1 readers, read words with font emphasis (specifically, capitals) more slowly than nonemphasized words. This means that it is possible that online processing demands account for the non-native-like prominence processing of L2 learners in some circumstances, demands that should be reduced by the present self-paced reading task. Second, more direct evidence came from a recognition memory test in which memory for the discourse was assessed using probes that referred to the correct item, to the salient contrast item, or to a merely mentioned item that was not a plausible alternative. Korean-speaking L2 learners of English were overall less successful at correctly rejecting contrast probes than native speakers, indicating that L2 learners were more confused about which of the members of an alternative set was the correct item. Crucially, however, font emphasis had the same effects on L2 learners' discourse memory as for native speakers: font emphasis helped L2 learners reject the contrast probes, but not the merely mentioned probes, indicating that L2 learners were native-like in using font emphasis to facilitate memory for discourse.

Previous studies on L2 learners' processing of prominence cues have focused primarily either on whether they know the mapping between those cues and information structure (Akker & Cutler, 2003; Baker, 2010; Donaldson, 2012; Reichle, 2010b) or on *whether* they encode an alternative set in response to contrastive focus (e.g., Braun & Tagliapietra, 2011; Reichle & Birdsong, 2014). Taking a step further, the present study also provided an answer to *how* L2 learners constrain an alternative set, if they encode it. The distinct effects of font emphasis on different types of false probes, both of which referred to items mentioned in the original discourse, suggest that an alternative set was constrained by discourse plausibility rather than mere mention in discourse. Taken together with the results above, these results suggest that L2 learners, at least in comprehension of written discourse, are similar to native speakers in how they restrict a set of salient alternatives and represent them in memory.

Native-like and non-native-like processing of prominence

Previous work reported that L2 learners are not always similar to native speakers in encoding and representing salient alternatives in a discourse; for instance, contrastive accents fail to lead L1-Korean learners of English to encode contrastive information fully enough to successfully reject false statements about salient alternatives on later memory (Lee & Fraundorf, 2017).

Our present results with written discourse provide empirical evidence that L2 discourse processing may not be intrinsically disadvantaged. Rather, the non-native-like behavior observed in past studies could have resulted from the cognitive demands of processing a prosodic cue in addition to the segmental content of running speech, an assumption underlying the shallow representation account. We found that prominent words require more reading time than non-prominent words for L2 learners. In addition, once these online processing demands were managed by allowing L2 learners as much time as needed (i.e., using a self-paced reading task), moderate- or high-proficiency L2 learners encode salient alternatives in a native-like way in response to a prominence cue in a written text. We have attributed this discrepancy to the difference in whether the task allows them to have sufficient cognitive resources to process the contrastive information.

The fact that L1-Korean learners of English perform better with a font emphasis cue than a prosodic cue is not attributable to the availability of each prominence cue in L1-Korean. Neither pitch accenting nor capitalization is used to mark prominence in Korean; Seoul Korean lacks pitch accents, and the Korean writing system does not have a distinction between uppercase and lowercase letters. However, one might nevertheless argue that our present results reflect Korean-speaking L2 learners' distinct sensitivity to different prominence cues in English; that is, perhaps Korean-speaking L2 learners simply have less knowledge about contrastive pitch accents than font emphasis. Some data suggest against this possibility: high-proficiency L1-Korean learners of English were sensitive to the contrastive meaning of an L + H* pitch accent even when they failed to fully integrate the contrast relation into memory (Lee & Fraundorf, 2017; see also Braun & Tagliapietra, 2011, for further evidence of L2 learners' sensitivity to L2 contrastive pitch accents). This suggests that the differences in L2 learners' comprehension of font emphasis and pitch accenting do not stem from differences in their knowledge about the mapping between each prominence cue and contrast. Still, it is possible that L2 learners' non-native encoding of salient alternatives in response to a contrastive pitch accent in speech may be due at least in part to the inherent difficulty of learning to process the contrastive pitch accent itself. This possibility can only be fully tested in the future by manipulating the prominence cue while holding time pressure constant, or vice versa.

While the current data showed that even moderate-proficiency L1-Korean learners of English evoked a contrast set in response to font emphasis in written discourse, the ability to process contrastive accents in L2 speech appears modulated by L2 proficiency (e.g., Baker, 2010; Braun & Tagliapietra, 2011; Lee & Fraundorf, 2017). For instance, only high-proficiency L2 learners showed memory benefits of an L + H* pitch accent in L2 speech (Lee & Fraundorf, 2017). This suggests that learning to process contrastive pitch accents requires more experience with L2 input compared to capitalization. Given that the L2 learners tested in the current study and Lee and Fraundorf (2017) all have spent less than a year in an English-speaking country, it is possible that L2 learners with more exposure to L2 speech and higher proficiency may reveal native-like encoding of salient alternatives in response to a prominence cue in L2 speech as well. At least

in some cases, L2 learners' behavior may converge with that of native speakers as a function of increased L2 proficiency or exposure (e.g., Frenck-Mestre, 2002; Hopp, 2006, 2009; 2010; Rossi, Gugler, Freiderici, & Hahne, 2006). Conversely, it is also possible that learners with very low L2 proficiency, whom we did not examine, may face challenges beyond resource limitations and may exhibit representational deficits in their L2 discourse processing.

Reading emphasized words in L2

Using a self-paced reading task, we measured word-by-word reading time while L2 learners were reading the original discourse. The data showed that L2 learners did spend extra time on words emphasized with capitals compared to those with no capitals, indicating that they were sensitive to a font manipulation while reading. Unlike native speakers, L2 learners' slowdown was restricted to the critical word, revealing no spillover effects. L2 learners' mean reading time was also overall longer than native English speakers. Thus, one possibility is that the effect of capitalization may not have been carried over to a subsequent region in L2 reading because L2 learners may have spent as much time as required to process capitalized words before moving on to another word.

Consistent with the L1 study using the same stimuli (Fraundorf et al., 2013, Experiment 3), initial reading time had no direct effect on L2 learners' performance on the later memory test. The absence of an effect of online reading time on later memory performance may simply reflect a lack of statistical power. It may also be due to a lack of variability in online reading time. We would observe a memory difference between words read slowly versus quickly only if readers slowed down in *some* cases but not others; instead, we observed that participants consistently read emphasized words more slowly. The third possibility is that longer reading time failed to predict better performance on the later memory test because the slowdown stemmed from decoding capital letters rather than from computing the alternative set; this possibility could be tested in future research by manipulating font emphasis for words inside and outside a contrast set. Furthermore, there have been mixed findings across studies as to whether native English speakers show better performance in offline comprehension as a function of longer reading times (Caplan, DeDe, Waters, Michaud, & Tripodis, 2011; Christianson & Luke, 2011; Daneman, Lennertz, & Hannon, 2007; Fraundorf et al., 2013; Reder & Kusbit, 1991; Ward & Sturt, 2007). More research is required to draw a firm conclusion about the relation between online processing and offline comprehension in L1 and L2 processing.

We do note that the particular self-paced reading task we chose may also alter reading behavior from natural reading insofar as readers could not regress or return to previous words; this may increase the burden placed on readers to remember the previous words of the text. However, this may be advantageous for the goals of our present study, in which we sought to compare results from self-paced reading to prior results from listening. As it is also not possible to return to previous words when listening to spoken speech, the use of noncumulative self-paced reading makes the memory demands more comparable between tasks.

Sensitivity to discourse-level cues

Previous studies of L2 processing at the discourse level have shown that L1-L2 differences exist, in at least some aspects of discourse processing, such that L2 learners are actually more sensitive to information given explicitly in the discourse than native speakers. For example, L2 learners can use referential discourse context in resolving syntactic ambiguity when native speakers fail to do so (Pan & Felser, 2011; Pan et al., 2015). L1-L2 differences are also seen in anaphor resolution. L2 learners tend to rely more on discourse-level cues than syntactic cues in interpreting anaphoric expressions even when the resulting interpretation is syntactically inappropriate (Felser & Cunnings, 2012; Felser et al., 2009; Patterson et al., 2014). We did not find that L2 learners were *more* sensitive to discourse-level cues as these studies have suggested. Instead, we found that L2 learners can use prominence cues at least in written discourse in a native-like manner to infer a contrast between one element and its alternatives, and to later rule out the alternatives when remembering the content of a discourse.

Resource limitations in discourse processing

A growing body of research suggests that L2 learners differ from native speakers in retrieving lexical information (Ardila, 2003) and in using and integrating multiple sources of information online (Hopp, 2006, 2009, 2010; Kilborn, 1992; Sorace & Serratrice, 2009; for review, see White, 2011) in part because of cognitive resource limitations rather than because of a fundamental difference in processing mechanism and representation. For example, Sorace and Serratrice (2009) have proposed that L2 learners' processing limitations may be a source of non-nativeness in integrating discourse/pragmatic information with syntax. The present study extends this finding to the domain of L2 discourse representation by showing that how salient alternatives are represented in memory for L2 discourse may also be constrained by cognitive resources. Further, Fraundorf et al. (2012) showed that limited cognitive resources cause difficulty allocating attention to information other than the contrastive information even in L1 discourse. The current study, combined with that of Lee and Fraundorf (2017), takes this finding a step further by showing that limited cognitive resources may also modulate the depth of encoding the contrastive information itself in memory, at least for L2 learners.

Conclusion

It is not possible to attend to and remember everything in a discourse. The present study demonstrates that prominence cues, such as font emphasis, may help L2 comprehenders identify what to attend to and remember from a discourse. Font emphasis facilitates memory for L2 discourse by enhancing a distinction between what happened and salient alternatives that did not happen, just as it does in L1 processing. Further, there is also an L1-L2 convergence in how an alternative set is restricted. These findings suggest that L2 comprehenders may construct a

similar kind of discourse representation as that of native comprehenders when sufficient resources are available.

NOTES

1. This term refers to an account of prosodic and discourse processing in which L2 learners evoke but do not fully process an alternative set. We do not intend for this hypothesis to be equated with the *shallow structure hypothesis* (Clahsen & Felser, 2006a, 2006b) of L2 *syntactic* processing, nor do we intend for our present investigation of discourse processing to be a test of the shallow structure hypothesis.
2. Although the word *font* is often taken to refer to differences between *faces* such as Times New Roman and Arial, it more properly refers to variations within a particular face, such as italicization ("Font," 2011).
3. For both reading times and recognition memory, the model including proficiency as a factor did not yield significant effects of proficiency, nor did proficiency interact with the other variables in the model. Thus, we present the data collapsed over moderate- and high-proficiency groups.

REFERENCES

Akker, E., & Cutler, A. (2003). Prosodic cues to semantic structure in native and nonnative listening. *Bilingualism: Language and Cognition*, *6*, 81–96.

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

Archibald, J. (1997). The acquisition of second language phrasal stress: a pilot study. In S. J. Hannahs, & M. Young-Scholten (Eds.), Focus on phonological acquisition (pp. 263–289). Amsterdam: John Benjamins.

Ardila, A. (2003). Language representation and working memory with bilinguals. *Journal of Communication Disorders*, *36*, 233–240.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics*. Cambridge: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412.

Baker, R. E. (2010). *Non-native perception of native English prominence*. Paper presented at the 5th International Conference on Speech Prosody, Chicago, IL, May 11–14, 2010.

Bates, D., Maechler, M., & Bolker, B. (2012). lme4: Linear mixed-effects models using S4 classes. R package version 0.999999-0. Retrieved from http://CRAN.R-project.org/package=lme4

Beckman, M. E., & Ayers, G. M. (1997). *Guidlines for ToBI labelling, version 3.0. Columbus*, OH: Ohio State University.

Blok, P. I., & Eberle, K. (1999). What is the alternative? The computation of focus alternatives from lexical and sortal information. In P. Bosch & R. van der Sandt (Eds.), *Focus: Linguistic, cognitive, and computational perspectives* (pp. 105–119). Cambridge: Cambridge University Press.

Braun, B., & Tagliapietra, L. (2010). The role of contrastive intonation contours in the retrieval of contextual alternatives. *Language and Cognitive Processes*, *25*, 1024–1043.

Braun, B., & Tagliapietra, L. (2011). On-line interpretation of intonational meaning in L2. *Language and Cognitive Processes*, *26*, 224–235.

Byram Washburn, M., Kaiser, E., & Zubizarreta, M. L. (2011). Focus facilitation and non-associative sets. In R. Artstein, M. Core, D. DeVault, K. Georgila, E. Kaiser, & A. Stent (Eds.),

SemDial 2011: *Proceedings of the 15th workshop on the semantics and pragmatics of dialogue* (pp. 94–102). Los Angeles, CA: SemiDial.

Calhoun, S. (2009). What makes a word contrastive? Prosodic, semantic, and pragmatic perspectives. In D. Barth-Weingarten, N. Dehé, & A. Wichmann (Eds.), *Studies in pragmatics: Where prosody meets pragmatics: Research at the interface* (pp. 53–77). Bingley, UK: Emerald.

Caplan, D., DeDe, G., Waters, G., Michaud, J., & Tripodis, Y. (2011). Effects of age, speed of processing, and working memory on comprehension of sentences with relative clauses. *Psychology and Aging*, *26*, 439–450.

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

Christianson, K., & Luke, S. G. (2011). Context strengthens initial misinterpretations of text. *Scientific Studies of Reading*, *15*, 136–166.

Clahsen, H., & Felser, C. (2006a). Continuity and shallow structures in language processing. *Applied Psycholinguistics*, *27*, 107–126.

Clahsen, H., & Felser, C. (2006b). How native-like is non-native language processing? *Trends in Cognitive Sciences*, *10*, 564–570.

Cowles, H. W. (2003). *Processing information structure: Evidence from comprehension and production* (Unpublished doctoral dissertation, University of California, San Diego).

Daneman, M., Lennertz, T., & Hannon, B. (2007). Shallow semantic processing of text: Evidence from eye movements. *Language and Cognitive Processes*, *22*, 83–105.

Dekydtspotter, L., Donaldson, B., Edmonds, A. C., Fultz, A. L., & Petrusch, R. A. (2008). Syntactic and prosodic computations in the resolution of relative clause attachment ambiguity by English-French learners. *Studies in Second Language Acquisition*, *30*, 453–480.

Dekydtspotter, L., Edmonds, A. C., Fultz, A. L., & Renaud, C. (2010). Modularity of L2 sentence processing: Prosody, context, and morphology in relative clause ambiguity in English-French interlanguage. In M. Iverson, I. Ivanov, T. Judy, J. Rothman, R. Slabakova, & M. Tryzna (Eds.), *Proceedings of the 2009 Mind/Context Divide Workshop* (pp. 13–27). Somerville, MA: Cascadilla Press.

Donaldson, B. (2012). Syntax and discourse in near-native French: Clefts and focus. *Language Learning*, *62*, 902–930.

Felser, C., & Cunnings, I. (2012). Processing reflexives in English as a second language: The role of structural and discourse-level constraints. *Applied Psycholinguistics*, *33*, 571–603.

Felser, C., & Roberts, L. (2007). Processing wh-dependencies in a second language: A cross-modal priming study. *Second Language Research*, *23*, 9–36.

Felser, C., Roberts, L., & Marinis, T. (2003). The processing of ambiguous sentences by first and second language learners of English. *Applied Psycholinguistics*, *24*, 453–489.

Felser, C., Sato, M., & Bertenshaw, N. (2009). The on-line application of binding principle A in English as a second language. *Bilingualism: Language and Cognition*, *12*, 485–502.

Ferreira, F., & Clifton, C. Jr. (1986). The independence of syntactic processing. *Journal of Memory and Language*, *25*, 348–368.

Font (2011). *Oxford English Dictionary.* Retrieved November 8, 2011, from http://www.oed.com.proxy2.library.illinois.edu/view/Entry/73956

Fraundorf, S. H., Benjamin, A. S., & Watson, D. G. (2013). What happened (and what did not): Discourse constraints on encoding of plausible alternatives. *Journal of Memory and Language*, *69*, 196–227.

Fraundorf, S. H., Watson, D. G., & Benjamin, A. S. (2010). Recognition memory reveals just how CONTRASTIVE contrastive accenting really is. *Journal of Memory and Language*, *63*, 367–386.

Fraundorf, S. H., Watson, D. G., & Benjamin, A. S. (2012). The effects of age on the strategic use of pitch accents in memory for discourse: A processing-resource account. *Psychology and Aging*, *27*, 88–98.

Frenck-Mestre, C. (2002). An on-line look at sentence processing in the second language. In R. Heredia & J. Altarriba (Eds.), *Bilingual sentence processing* (pp. 217– 236). New York: Elsevier.

Frenck-Mestre, C., & Pynte, J. (1997). Syntactic ambiguity resolution while reading in second and native languages. *Quarterly Journal of Experimental Psychology*, *50A*, 119–148.

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

Harrington, M. (1992). Working memory capacity as a constraint on L2 development. In R. J. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 123–136). Amsterdam: North-Holland.

Hopp, H. (2006). Syntactic features and reanalysis in near-native processing. *Second Language Research*, *22*, 369–397.

Hopp, H. (2009). The syntax-discourse interface in near-native L2 acquisition: Off-line and on-line. *Bilingualism: Language and Cognition*, *12*, 463–483.

Hopp, H. (2010). Ultimate attainment in L2 inflection: Performance similarities between non-native and native speakers. *Lingua*, *120*, 901–931.

Husband, E. M., & Ferreira, F. (2016). The role of selection in the comprehension of focus alternatives. *Language, Cognition, and Neuroscience*, *31*, 217–235.

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

Jackson, C. N., & Bobb, S. C. (2009). The processing and comprehension of *wh*-questions among second language speakers of German. *Applied Psycholinguistics*, *30*, 603–636.

Juffs, A., & Harrington, M. (1996). Garden path sentences and error data in second language processing. *Language Learning*, *46*, 283–323.

Jun, S.-A. (1993). *The phonetics and phonology of Korean prosody* (Unpublished doctoral dissertation, Ohio State University).

Jun, S.-A. (2005). Prosodic typology. In S.-A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing*. Oxford: Oxford University Press.

Kilborn, K. (1992). On-line integration of grammatical information in a second language. In R. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 337–350). Amsterdam: Elsevier.

Kluender, R., & Kutas, M. (1993). Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience*, *5*, 196–214.

Ladd, D. R. (2008). *Intonational phonology*. Cambridge: Cambridge University Press.

Lee, E.-K., & Fraundorf, S. H. (2017). Effects of contrastive accents in memory for L2 discourse. *Bilingualism: Language and Cognition*, *20*, 1063–1079.

Lee, E.-K., & Garnsey, S. M. (2015). An ERP study of plural attraction in attachment ambiguity resolution: Evidence for retrieval interference. *Journal of Neurolinguistics*, *36*, 1–16.

Lee, E.-K., Lu, D. H., & Garnsey, S. M. (2013). L1 word order and sensitivity to verb bias in L2 processing. *Bilingualism: Language and Cognition*, *16*, 761–775.

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

Marinis, T., Roberts, L., Felser, C., & Clahsen, H. (2005). Gaps in second language processing. *Studies in Second Language Acquisition*, *27*, 53–78.

McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, *55*, 381–401.

Murayama, K., Sakaki, M., Yan, V. X., & Smith, G. M. (2014). Type I error inflation in the traditional by-participant analysis to metamemory accuracy: A generalized mixed effects model perspective. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 1287–1306.

Pan, H.-Y., & Felser, C. (2011). Referential context effects in L2 ambiguity resolution: Evidence from self-paced reading. *Lingua*, *121*, 221–236.

Pan, H.-Y., Schimke, S., & Felser, C. (2015). Referential context effects in non-native relative clause ambiguity resolution. *International Journal of Bilingualism*, *19*, 298–313.

Papadopoulou, D. (2005). Reading-time studies of second language ambiguity resolution. *Second Language Research*, *21*, 98–120.

Papadopoulou, D., & Clahsen, H. (2003). Parsing strategies in L1 and L2 sentence processing: A study of relative clause attachment in Greek. *Studies in Second Language Acquisition*, *24*, 501–528.

Patterson, C., Trompelt, H., & Felser, C. (2014). The online application of binding condition B in native and non-native pronoun resolution. *Frontiers in Psychology*, *5*, 147. doi:10.3389/fpsyg.2014.00147

Pierrehumbert, J., & Hirschberg, J. (1990). The meaning of intonational contours in the interpretation of discourse. In P. Coehn, J. Morgan, & M. Pollack (Eds.), *Intentions in communication*. Cambridge, MA: MIT Press.

Prince, E. F. (1981). Toward a taxonomy of given/new information. In P. Cole (Ed.), *Radical pragmatics* (pp. 223–255). New York: Academic Press.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372–422.

Reder, L. M., & Kusbit, G. W. (1991). Locus of the Moses illusion: Imperfect encoding, retrieval, or match? *Journal of Memory and Language*, *30*, 385–406.

Reichle, R. V. (2010a). Near-nativelike processing of contrastive focus in L2 French. In B. Van Patten & J. Jegerski (Eds.), *Research on second language processing and parsing* (pp. 321–344). Amsterdam: Benjamins.

Reichle, R. V. (2010b). Judgments of information structure in L2 French: Nativelike performance and the critical period hypothesis. *International Review of Applied Linguistics in Language Teaching*, *48*, 53–85.

Reichle, R. V., & Birdsong, D. (2014). Processing focus structure in L1 and L2 French: L2 proficiency effects on ERPs. *Studies in Second Language Acquisition*, *36*, 535–564.

Reichle, E. D., Warren, T., & McConnell, K. (2009). Using E-Z Reader to model the effects of higher level language processing on eye movements during reading. *Psychonomic Bulletin & Review*, *16*, 1–21.

Reves, T., & Levine, A. (1988). The foreign language receptive skills: Same or different? *System*, *16*, 327–336.

Roberts, L., & Felser, C. (2011). Plausibility and recovery from garden-paths in second-language sentence processing. *Applied Psycholinguistics*, *32*, 299–331.

Roberts, L., Gullberg, M., & Indefrey, P. (2008). On-line pronoun resolution in L2 discourse: L1 influence and general learner effects. *Studies in Second Language Acquisition*, *30*, 333–357.

Rossi, S., Gugler, M. F., Freiderici, A. D., & Hahne, A. (2006). The impact of proficiency on syntactic second-language processing of German and Italian: Evidence from event-related potentials. *Journal of Cognitive Neuroscience*, *18*, 2030–2048.

Segalowitz, N. (2003). Automaticity and second languages. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 382–388). Oxford: Blackwell.

Sorace, A., & Serratrice, L. (2009). Internal and external interfaces in bilingual language development: Beyond structural overlap. *International Journal of Bilingualism*, *13*, 1–16.

Van Berkum, J. J. A., Brown, C. M., & Hagoort, P. (1999). Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory and Language*, *41*, 147–182.

Ward, P., & Sturt, P. (2007). Linguistic focus and memory: An eye movement study. *Memory & Cognition*, *35*, 73–86.

Watson, D. G., Tanenhaus, M. K., & Gunlogson, C. A. (2008). Interpreting pitch accents in online comprehension: H* vs. L+H*. *Cognitive Science*, *32*, 1232–1244.

Weber, A., Braun, B., & Crocker, M. W. (2006). Finding referents in time: Eye-tracking evidence for the role of contrastive accents. *Language and Speech*, *49*, 367–392.
White, L. (2011). Second language acquisition at the interfaces. *Lingua*, *121*, 577–590.
Williams, J. (2006). Incremental interpretation in second language sentence processing. *Bilingualism: Language and Cognition*, *9*, 71–88.
Wright, D. B., Horry, R., & Skagerberg, E. M. (2008). Functions for traditional and multilevel approaches to signal detection theory. *Behavior Research Methods*, *41*, 257–267.
Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*, 162–185.

APPENDIX A

Fixed effect estimates for the model of "true" responses for just L2 comprehenders (N = 3,489, log-likelihood: –1981); the model also included by-participant and by-item random intercepts, and by-participant and by-item random slopes for probe type as random effects

Fixed effects	β	*SE*	Wald *z*	*p*
Baseline rate of "true" responses (response bias)	0.06	0.09	0.67	>.1
Emphasis (effect on response bias)	< 0.01	0.08	0.05	>.1
Reading time (effect on response bias)	0.05	0.11	0.43	>.1
Rejections of contrast probes versus baseline (sensitivity)	0.59	0.08	7.36	< .01
Rejections of merely mentioned probes versus baseline (sensitivity)	0.91	0.11	8.66	< .01
Emphasis × Reading Time (effect on response bias)	0.09	0.21	0.41	>.1
Emphasis × Rejections of Contrast Probes (effect on sensitivity)	0.23	0.11	2.04	< .05
Emphasis × Rejections of Merely Mentioned Probes (effect on sensitivity)	–0.11	0.11	–0.94	>.1
Reading Time × Rejections of Contrast Probes (effect on sensitivity)	0.11	0.14	0.73	>.1
Reading Time × Rejections of Merely Mentioned Probes (effect on sensitivity)	<0.01	0.14	0.00	>.1
Emphasis × Reading Time × Rejections of Contrast Probes (effect on sensitivity)	–0.06	0.29	–0.22	>.1
Emphasis × Reading Time × Rejections of Merely Mentioned Probes (effect on sensitivity)	0.11	0.28	0.40	>.1

Note: SE, standard error.

https://doi.org/10.1017/S0142716418000619 Published online by Cambridge University Press