Second Language Readers' Use of Context to Identify Proper Names

Kimberly Klassen Showa Women's University

ABSTRACT

This study investigated how well second language (L2) readers of English use context to identify proper names as such. It represents a first step in exploring a widely held assumption that L2 readers of English can easily identify proper names by their form and function. The study isolates the issue of function to investigate whether context alone allows identification of proper names. Context may be especially important to the identification of proper names for readers processing a different orthography from their first language (L1), as there is no research indicating that the initial capital letter on English proper names facilitates proper name recognition in L2 readers. Japanese L2 low intermediate readers of English (N = 61) were given sentences in lower-case letters. Target items had a proper name usage and a common word meaning (e.g. Rose, rose). Participants were asked to identify proper names by changing lower-case letters to upper case. Participants correctly identified target proper names in 28% of cases. The findings contradict an assumption in L2 reading and vocabulary research that proper names are easily understood from context. Implications for research and pedagogy are discussed.

INTRODUCTION

Researchers investigating how much vocabulary a second language (L2) reader needs to know in order to read certain texts with adequate comprehension are presented with the problem of proper names: Should proper names be treated as known or unknown to the L2 reader, or should proper names be handled like other lexical items and classified according to frequency? A proper name is a name that refers to a unique person, place or organisation; in English, proper names are capitalised ("Proper noun," n.d.). However, this simple definition belies the complex debate among linguists and grammarians on what items should be regarded as proper names, relating to issues of prototypicality. Orthography is not a reliable guide, as there are irregularities in the English writing system. For example, calendrical items are capitalised in English, though not in other languages like French (e.g. *Monday, lundi)* (Allerton, 1987). Because calendrical items do not refer uniquely, there is disagreement on whether they should be considered proper names. Another example of irregular capitalisation is seen in adjectives of nationality, which are capitalised in English though not in languages like Spanish (e.g. *Japanese, japonesas*). In addition, there are common nouns that "grow capital letters", what Strawson (1950) refers to as "quasi-names" (e.g. *the Great War*) (p. 341).

Adding to the complexity of the debate, there is the issue of the syntactic categorisation of proper names. While proper names have been traditionally regarded as a subclass of nouns in European grammar studies (Lyons, 1977), some linguists including Anderson (2007) argue that proper names do not behave syntactically like nouns, and are more similar to

determinatives (pronouns and determiners) (p. 172). Coates (2006) also discusses the possibility of unlinking proper names from the linguistic noun category, noting the Eurocentric view (p. 373). (To acknowledge this possible alternative syntactic categorisation, the term 'proper name' will be used throughout this paper). As there is very little research into how L2 readers analyse proper names (a notable exception is Kobeleva (2012) who looked at how L2 listeners respond to proper names), one might surmise that L2 readers, drawing on their declarative knowledge about English proper names, might analyse initial capital items in mid-sentence as proper names, without much regard to their prototypicality.

Several early L2 vocabulary studies, beginning in the 1980s, speculated that the L2 reader could easily infer the meaning of proper names. For example, Hwang and Nation (1989) argued that proper names in newspapers can be treated as known vocabulary because these items will have been learnt in the L1 (e.g. *Canada; Margaret Thatcher*) and most names are explained in context (e.g. *Prime Minister Jacques Chirac*) (p. 324). In another study, looking at L1 young adult novels as reading material for L2 learners, Hirsh and Nation (1992) offer "strong reasons" why names do not require prior learning: the form (capitalisation) and function in the story will clearly signal that these are proper names (p. 691). Similarly, Nation and Wang (1999) note that they listed proper names as separate from other lexis: "proper nouns could be easily understood from context and should not be counted as unknown vocabulary" (p. 358).

Because of this supposition, it has become standard practice in text coverage counts to either re-categorise proper names as known (i.e. as 1K band items) or to remove them altogether from the analysis (see Brown, 2010). Since proper names can make up a considerable percentage of vocabulary in any given text (3-6%, depending on the genre) (Nation, 2006), the treatment of proper names in text coverage counts can have a significant impact on vocabulary analyses. In fact, several studies do acknowledge the effect that different treatment of proper names can have on an analysis. Those studies display coverage results to show inclusion of proper names (as known), and exclusion of proper names (as unknown, or off-list) (e.g. Nation, 2006; Webb & Rodgers, 2009). Treating names as known or unknown can sometimes represent the difference in a task that matches a reader's ability and a task that is very difficult for the reader. A more recent trend seen in some studies is to remove proper names from the lexical analysis, making it impossible to see how much of the text is comprised of proper names (e.g. Uden, Schmitt, & Schmitt, 2014).

There is to date little empirical research to guide decisions about how to handle proper names in L2 reading (Klassen, 2021). One study (Kobeleva, 2012) investigated the effect of unfamiliar proper names on L2 listening comprehension. Kobeleva (2012) compared listening comprehension of ESL participants (N = 110) using a short news story in two conditions: Names Known (i.e. pre-taught) and Names Unknown (i.e. unfamiliar). The participants in the Names Unknown condition often mistook proper names for common words. Listening comprehension was significantly higher in the Names Known condition than Names Unknown, though the effect was seen in listening for details, not global comprehension. While the study focused on L2 listening, the findings suggest implications for L2 reading, that proper names might not be always easily inferred from context. However, as noted, there is an absence of empirical research related to proper names and L2 reading.

Conjectures about L2 readers and proper names

Under this assumption that context aids the L2 reader to identify and understand the referent of an unfamiliar proper name, three conjectures are in fact being made: first, that the L2 reader has efficient decoding skills to quickly and easily identify unfamiliar names in connected text; second, that the reader can skilfully use context to infer information related to

the name; and third, that context is always going to be explicit to reveal necessary information about the name. The validity of each of these conjectures is considered in turn, in light of available evidence.

First, the conjecture that L2 readers possess efficient lower-level processing skills to easily identify unfamiliar proper names in context may be unfounded. Information-processing views of reading postulate that efficient lower-level processing, such as letter and word recognition, is critical for successful higher-level comprehension processing, such as using context and inferencing skills, to take place (Grabe, 2009; Perfetti, 2007). However, the few L2 studies that have investigated the role of word recognition skills in L2 adult readers have found that even highly proficient L2 users are lacking in efficient lower-level processing skills (Akamatsu, 2003; Nassaji, 2003a; Shiotsu, 2009). Other research has shown that advanced bilinguals read 30% more slowly in their L2 than L1 readers; furthermore, this research suggests that the slower pace is due to inefficiencies in L2 word recognition, and not with higher-level processes such as text integration or connecting to background knowledge (Segalowitz, Poulsen, & Komoda, 1991). Nassaji (2014) notes that these findings are important not only because they demonstrate the importance of automaticity to L2 reading, but also because they challenge the notion that L2 word recognition skills are developed as a result of language proficiency (p. 8). Rather, word recognition skills might develop as a result of processing experience (Koda, 1996, 2005), which an L2 reader will have less of, in comparison to an L1 reader. Furthermore, as Alderson (2000) suggests, "Since difficulty in processing letters is related to automaticity of word identification, and since speed of word recognition affects speed and efficiency of reading, one might expect that second-language readers processing different orthographies or scripts might experience greater difficulty" (p. 75). Thus, it may be incautious to assume that L2 users, in particular those processing a different orthography than their L1, have efficient word recognition skills to identify unfamiliar proper names when reading continuous text.

A second, related conjecture concerns L2 readers' skill at using context to infer word meaning. Using context to guess the meaning of an unknown word is known as 'lexical inferencing', a process that "involves making informed guesses as to the meaning of an utterance in light of all available linguistic cues in combination with the learner's general knowledge of the world, her awareness of context and her relevant linguistic knowledge" (Haastrup, 1991, p. 40). Several studies have investigated L2 readers' strategic use of context to infer word meaning (e.g. Bensoussan & Laufer, 1984; De Bot, Paribakht, & Wesche, 1997; Fraser, 1999; Hu & Nassaji, 2012, 2014; Huckin & Bloch, 1993; Nassaji, 2003b). Findings from several studies suggest that L2 readers are not as efficient at using context for lexical inference as is sometimes assumed. For example, Bensoussan and Laufer (1984) compared their learners' (N = 60) knowledge of 70 lexical items, presented first without context (i.e. in a list) and then in context. They found that context aided guessing of unknown lexical items for only 24% of the words in their text; guessing of the other 76% of items was not aided by context, either because there were no contextual clues or because participants did not make use of the clues. The authors ranked the participants' responses to unknown lexical items and found that the most frequent response was to ignore unknown words (i.e. no attempt to guess was made). This was followed by wrong guesses, which were more frequent than correct or approximately correct answers. Bensoussan and Laufer (1984) also compared their participants' ability to use context to infer word meaning by their proficiency level: they found that higher proficiency participants did not perform any better than those with lower proficiency.

In a more recent study, Nassaji (2003b) also concluded that learners are not very successful when using context for lexical inference. He used think-aloud procedures with 21 adult, intermediate learners of English from five different L1 backgrounds. Analysis of the

learners' responses showed that they were unable to infer word meaning more than half the time (55.8%). Partially successful responses made up 18.6% and successful responses were 25.6%. Nassaji (2003b) notes that for some of the target items, similarity in word form was a source of confusion (e.g. *affluence* mistakenly connected to *influence*), and he suggests that efficiency in decoding skills (word recognition) may be related to the ability to correctly infer meaning (p. 654).

Huckin and Bloch (1993) make a similar suggestion, noting a connection between failure in word recognition and failure to use context for lexical inference. The authors conducted think-alouds with three intermediate Chinese readers of English to determine how they used context to infer word meaning. Out of 44 guessing opportunities, the readers were incorrect 20 times (45%). Of these incorrect inferences, nearly half (9) were what the authors refer to as 'mistaken IDs', that is, words that were mistaken for a similar looking word (e.g. *optimal* mistaken for *optional*). Huckin and Bloch (1993) note that these cases of mistaken IDs are illustrative of how word shape can sometimes override context to reconfirm or check the meaning (p.166).

Though Huckin and Bloch (1993) do not remark on it directly with respect to the mistaken IDs, of import here is that their subjects were L1 logographic (Chinese) readers and as such, may rely more heavily on word shape analysis than non-logographic readers (Akamatsu, 2003; Ehrich, Zhang, Mu, & Ehrich, 2013; Koda, 2005). By the nature of how they read in their L1, logographic readers seem to rely more on decoding skills than higher-level processing skills; these L1 reading skills might then transfer to L2 reading (Koda, 2004). For example, in her comparison of L2 reading strategies of Nigerian and Chinese students, Parry (1996) found the Chinese readers relied almost exclusively on bottom up processing (i.e. the sublexical route). Even in the self-reported use of context to guess an unknown word, it is apparent that a Chinese participant relied instead on morphological analysis to guess the meaning, not the sentential context (Parry, 1996, p. 677). This is illustrative of how L1 logographic readers may rely on decoding skills for L2 reading, as a result of how they read in their L1.

While English proper names are capitalised, it is not known whether the initial capital letter facilitates recognition of proper names for L2 readers, particularly for those whose L1 employs a different writing script or orthography (e.g. in German, all nouns are capitalised). Research has been conducted into L1 proper name processing from an orthographic perspective. Peressotti, Cubelli, and Job (2003) used the lexical decision task paradigm for recognition of L1 Italian proper names and common nouns. (Italian has a similar orthography to English in that proper names are capitalised while common nouns are not). Their findings suggest a "facilitation effect for proper names in the auditory condition, the authors propose the effect occurs at the orthographic level, not semantic. Furthermore, because no effect was found for proper names in the lexical decision task with illegal non-words, the initial capital letter "has a role only when lexical processing is required" (p. 107). However, it remains unknown whether this facilitation effect exists for L2 readers and should not be assumed.

Finally, there is a conjecture that context will always be explicit and provide relevant information about proper names. However, as Nagy and Anderson (1984) found in their survey of L1 school textbooks, some proper names, such as geographical names, are not often explained in context. They argue that lack of knowledge about such names would result in comprehension breakdown, just as unfamiliarity with the meaning of any other word might. They conclude with a "conservative" estimate of 1,000 proper names that should be treated as demands on the L1 reader's vocabulary knowledge, with an increasing number of names that are assumed rather than explained in context (p. 317).

As noted, an important consideration regarding lower-level processing of L2 proper names concerns research that has found L2 readers may be less skilful at using syntactic and semantic information to infer word meaning than has been previously assumed (Bensoussan & Laufer, 1984; Huckin & Bloch, 1993; Nassaji, 2003b). The role of syntactic and semantic information in word recognition processing is a debated issue: word recognition might occur through context, or context might act to disambiguate at a post-lexical level. With respect to proper names, the latter view seems likely given that context determines the referent of a name. Hanks (2013) says, "Context determines the values to be attached to names and how the preliminary probabilistic inferences are to be modified" (p. 37). He gives the example of the name *Peter*. An L1 user might assume this name refers to an English-speaking male. Context, however, might reveal otherwise: *Peter* could well refer to a pet or to Bertrand Russell's wife. Furthermore, as Alderson (2000) correctly notes,

Although context determines the meaning of an unknown word, it may not reveal it: revelation is limited not only by the explicitness of the connection between context and the unknown word, but also by the experience and skill of the reader. (p. 70)

This point is significant with regard to the assumption that proper names are easily inferred from context.

Thus, in light of empirical research reviewed here, efficient decoding skills in L2 readers should not be assumed; furthermore, if L2 readers are not proficient at using context to infer meaning, as the research summarised above indicates, then it may be mistaken to assume that context aids proper name identification in connected passages. This study aimed to answer the following research question: To what extent can Japanese intermediate L2 readers of English use sentential context to correctly identify target proper names? In this way, the study is a first step in exploring the assumption that proper names can be easily identified by their form and their function. This study isolates the issue of function and investigates whether it alone allows identification of proper names. Future work will look at whether form enables identification of proper names. The hypothesis is Japanese L2 readers of English will have limited success in using context to identify proper names. As logographic and syllabic L1 readers, they might rely more on decoding skills, such as word recognition, and less on contextual clues. In this regard, they may not be very successful at using context to identify the proper names, despite the assumption found in L2 vocabulary research that proper names can be easily understood from context. Furthermore, it is predicted that a reliance on word recognition skills may lead these readers to misidentify target items as proper names in sentences with the common word usage. This prediction is supported by research that indicates non-alphabetical L1 readers rely heavily on bottom-up processing in L2 reading (Ehrich et al., 2013; Huckin & Bloch, 1993; Parry, 1996).

METHODOLOGY

Participants

There were 61 participants (51 women; 10 men), all Japanese first-year university students who had at least six years of English study before entering university. Almost half (29) of the participants were Business majors from two intact classes at a small private womenonly university. The other participants (32) were Education majors from one intact class at a small private co-ed university. The participants had similar English proficiency levels. The Business majors had combined TOEIC scores ranging from 225 to 400; the Education majors had TOEFL scores ranging from 420 to 470. (Different test scores were used to determine proficiency because in Japan, business students often take the TOEIC test while students in other academic fields take TOEFL or IELTS). These scores place participants on the cusp of A2/B1 levels of the Common European Framework of Reference (CEFR).

Materials

Twenty target names were selected that have both a proper name usage and another common word meaning, as a necessary part of the study design. Table 1 lists the target items, the part of speech of the non-name meaning, and the frequency. These particular target items were chosen for their common word frequency ranking. The first fifteen items in the list are high frequency words; that is, the non-name forms appear in the three sub-lists (A, B, and C) on the Common Core List of shared items from the BNC and COCA corpora (Cobb, n.d.; Gardner, 2013). Because the participants in this study had different exposure to various varieties of English in their learning experiences with foreign teachers in Japan or during study abroad, the Common Core List was used as it has shared items from both British and North American registers. The last five items are off the Common Core List, and as such, might be less familiar to the participants.

| Target item | Part of speech | Frequency | | |
|-------------|----------------------|----------------------|--|--|
| _ | (for non-name usage) | (for non-name usage) | | |
| Rose/rose | verb | A | | |
| Mark/mark | noun | А | | |
| Cook/cook | noun | А | | |
| White/white | adjective | А | | |
| March/march | verb | A | | |
| Wood/wood | noun | A | | |
| Major/major | adjective | А | | |
| Green/green | adjective | A | | |
| Young/young | adjective | A | | |
| Brown/brown | adjective | В | | |
| Frank/frank | adjective | В | | |
| Hill/hill | noun | В | | |
| Grace/grace | noun | С | | |
| Bill/bill | noun | С | | |
| Pat/pat | noun | С | | |
| Nick/nick | noun | Off list | | |
| Mike/mike | noun | Off list | | |
| Bob/bob | verb | Off list | | |
| Jack/jack | noun | Off list | | |
| Cliff/cliff | noun | Off list | | |

Table 1. Target Items, Part of Speech and Frequency on Common Core List

Sentence contexts for each target item were selected from either the BNC (Davies, 2004-) or the COCA (Davies, 2008-) corpora to include both British and North American registers of English. Two sentences were selected for each target item, one that had the proper name usage and the other that had the common word meaning. Pairs of sentences were selected

that were similar in word length; as the sentences were taken from corpora, exact word length matches were not possible.

A task was created with two items sets (named Item Set A and Item Set B) using the selected sentences (see Appendix). Each item set consisted of 20 sentences with a target item appearing only once, either as a proper name or as a common word. Two item sets were used so that the experiment could be repeated; also, this ensured that each target item would appear only once, either as a proper name or a common word, in each item set. Ten sentences in each item set had target items with a proper name meaning and ten sentences with a common word meaning. The sentences were arranged randomly. Punctuation (commas, full stops, etc.) that appeared in the original corpus entry was maintained. However, all initial capital letters were changed to lower case letters; this was done so that participants would need to rely on sentential context to identify proper names. Also, because a lower-case letter was used for the first word in each sentence, this ensured that there was something for participants to do in every sentence, and thus demonstrated whether participants had understood the instructions. Both item sets had the same instructions to the participants, written in English and Japanese (L1). The English instructions read as follows: Read the sentences. Change the small letters to capital letters if necessary. You can use your dictionary if necessary. Look at the example. An example followed the instructions on each sheet, illustrating the instructions to the participants:

I NY

Example: i arrived in new york last night.

Procedure

Participants were asked if they agreed to take part in the experiment. They were given consent forms (in English and Japanese), explaining that participation was not obligatory and data would be kept confidential. All participants agreed to take part.

Data collection took place on two occasions, one week apart. In the first session, the Business majors (n = 29) were given Item Set B to complete, and the Education majors (n = 32) were given Item Set A. In the second session, each group completed the other item set, thereby controlling for any practice effect. Instructions were read aloud, and participants were asked if they had any questions regarding the instructions. The example was used to elicit English capitalisation rules relevant to the sentences in the experiment (i.e. capitalise the first word in each sentence, the pronoun *I*, and any names of people, places, days and months). Participants were told they would have 15 minutes to complete the task and that they could use their dictionaries. Because the sentences were corpus-derived and the participants had an intermediate proficiency level, it was important that they were allowed and encouraged to use dictionaries, and most participants did so. All participants finished the task within the allotted time; however, no papers were collected until 15 minutes passed to ensure no one felt time pressure. In the second session, the instructions, example and capitalisation rules were reviewed. Participants were given 15 minutes to complete the task and again encouraged to use dictionaries.

Participants' responses to both tasks were examined for correct responses. A response was considered correct if the participant had added a capital letter to target items with the proper name meaning (hereafter referred to as 'target names'). A response was also considered correct if the participant had not added a capital letter to target items with the common word meaning ('target non-names'). To focus the analysis on the target items, any errors relating to capitalisation of non-target items were ignored. Data analysis compared participants' responses on Item Sets A and B, so only data from participants who had completed both tasks (n = 54)

was included. (Data from seven participants who had completed only one of the two sets was removed in this initial analysis).

RESULTS

Descriptive statistics were compiled for correct responses to target names and target non-names in both tasks. The responses in the two tasks were first analysed separately to check for any practice effect. Because the responses were similar across tasks, participants' scores from Item Sets A and B were combined for further analysis. This gave each participant two mean scores: total number of correct target names and total number of correct target non-names (out of 20) (see Table 2).

| Total Correct Responses | | | | | |
|-------------------------|----|------|------|-------|-------|
| | n | Min. | Max. | M | SD |
| Target names | 54 | 0 | 12 | 5.78 | 3.063 |
| Target non-names | 54 | 15 | 20 | 18.37 | 1.233 |

To answer the research question (i.e. to what extent can Japanese low-intermediate L2 readers of English use sentential context to correctly identify proper names), the correct responses to target names and target non-names were compared. It should be note that this comparison is unreliable because no response was required for target non-names, and therefore, it is uncertain whether participants identified the items as non-names by not adding a capital letter. Before submitting the data to paired-samples T-test statistical analysis, a visual inspection of the data was done to check if the assumptions of parametric tests were met. Boxplots indicated that the data was fairly normally distributed for the target name responses, though responses to target names had larger range than target non-names. Boxplots indicated that data for target non-names was not normally distributed; there were no outliers in the data. Histograms confirmed a negative skew for target non-name responses. Tests of normality (Kolmogorov-Smirnov, p < .005) provided further confirmation of non-normal distribution for target non-name responses. As for variance, equal variance is assumed to be true for paired samples T-tests. However, because the data was not normally distributed, a non-parametric test alternative, the Wilcoxon Signed-Ranks Test for paired samples was run to compare responses to target names and target non-names.

The output from the Wilcoxon Signed-Ranks Test indicated that target non-name scores were statistically significantly higher than target name scores, Z = -6.399, p < .001. The effect size was large (r = 0.87). Thus, these results indicate participants were not as successful at using context to identify target names as non-names. In order to determine what factors may have contributed to participants correctly identifying target names, it was decided that participants' responses to each target name be investigated as post-hoc analysis.

Post-hoc analysis

For the post-hoc analysis, the data was re-examined to focus on the total number of correct responses to each target item. Because no comparison was being made between the two tasks, data from all participants (N = 61) was included. Sixty participants responded to ten target names on Item Set A (60 x 10 = 600); 55 participants responded to ten target names on

Item Set B (55 x 10 = 550). Thus, the total data set was 1,150 responses to target names. Likewise, 60 participants responded to ten target non-names on Item Set A ($60 \times 10 = 600$); 55 participants responded to ten target non-names on Item Set B ($55 \times 10 = 550$). The total data set was 1,150 responses to target non-names. Correct responses to each target name and non-name are summarised in Table 3.

The total row shows the correct number of responses to all target names in both item sets as 28.4%; correct responses to all target non-names was 91.4%. Because of the high number of correct responses to target non-names, it was decided to focus the post-hoc analysis on target names; that is, what factors may have led participants to correctly identify target names. Also, this would focus the analysis on the research question; that is, to what extent can L2 readers use sentential context to correctly identify proper names. Because the number of participants who responded to target names in Item Sets A and B was different (60 and 55, respectively), a scaling was done to the number of responses to Item Set B (i.e. dividing by 55 and multiplying by 60) to allow for comparison.

| Target | Total | Correct | | Target | Total | Correct | | |
|--------------|-------------------|-----------|------|--------------|-----------|---------|-----------|--|
| names | number of | responses | | non- | number of | | responses | |
| | responses | n | % | names | responses | n | % | |
| 1. Rose | 60 | 5 | 8.3 | rose | 55 | 54 | 98.2 | |
| 2. Mark | 60 | 31 | 51.7 | mark | 55 | 54 | 98.2 | |
| 3. White | 60 | 19 | 31.7 | white | 55 | 53 | 96.4 | |
| 4. Major | 60 | 7 | 11.7 | major | 55 | 55 | 100 | |
| 5. Bill | 60 | 8 | 13.3 | bill | 55 | 55 | 100 | |
| 6. Green | 60 | 15 | 25 | green | 55 | 55 | 100 | |
| 7. Wood | 60 | 4 | 6.7 | wood | 55 | 55 | 100 | |
| 8. Frank | 60 | 16 | 26.7 | frank | 55 | 52 | 94.5 | |
| 9. March | 60 | 3 | 5 | march | 55 | 48 | 87.3 | |
| 10. Bob | 60 | 40 | 66.7 | bob | 55 | 26 | 47.3 | |
| Subtotal | | 148 | 24.7 | Subtotal | | 507 | 92.2 | |
| (Item Set A) | | | | (Item Set B) | | | | |
| 11. Jack | 55 | 45 | 81.8 | jack | 60 | 33 | 55 | |
| 12. Hill | 55 | 17 | 30.9 | hill | 60 | 57 | 95 | |
| 13. Brown | 55 | 16 | 29.1 | brown | 60 | 57 | 95 | |
| 14. Young | 55 | 3 | 5.5 | young | 60 | 60 | 100 | |
| 15. Grace | 55 | 33 | 60 | grace | 60 | 59 | 98.3 | |
| 16. Nick | 55 | 8 | 14.5 | nick | 60 | 60 | 100 | |
| 17. Cook | 55 | 3 | 5.5 | cook | 60 | 60 | 100 | |
| 18. Mike | 55 | 28 | 51 | mike | 60 | 39 | 65 | |
| 19. Pat | 55 | 2 | 3.6 | pat | 60 | 60 | 100 | |
| 20. Cliff | 55 | 24 | 43.6 | cliff | 60 | 59 | 98.3 | |
| Subtotal | | 179 | 32.5 | Subtotal | | 544 | 90.7 | |
| (Item Set B) | | | | (Item Set A) | | | | |
| Total A+B | <i>n</i> , mean % | 327 | 28.4 | | | 1051 | 91.4 | |

Table 3. Correct Identification of Target Names and Non-names in Item Sets A and B

A multiple regression was carried out to investigate if there is a relationship between correct responses to target names and the explanatory variables of: the richness of sentential context, the frequency of the target name, and the part of speech associated with the corresponding target non-name. Data for this post-hoc analysis was gathered in three parts.

The first step in data collection was to generate a context rating for each sentence with a target name. Six L1 adult users of English were given the same sentences as the participants, though the target names were replaced by a blank. The L1 users were asked to fill in each blank with one suitable word. A context rating for each sentence was generated from the number of L1 users who entered a specific name or pronoun in the blank, thus indicating the likelihood that the target item referred to a person. This created a scale from 6 to 0, where if most L1 users entered a name or pronoun, the sentence was considered rich in context, strongly indicating the target item was a name. Conversely, if none of L1 users entered a name or pronoun, then the sentence was considered to have no context to indicate the target item was a name. Table 4 shows the L1 users' response types and tallies listed by the target name (which had been replaced by a blank).

| Target name ^a | Response types and tally | Context rating ^b |
|--------------------------|---------------------------------------|------------------------------------|
| Rose | her (4), Rose, some | 5 |
| Mark | he (6) | 6 |
| Jack | she (6) | 6 |
| (the) White (House) | top, White (2), owners, lower, main | 3 |
| Hill | Shaw (3), Green, Smith, | 5 |
| | unsurprisingly | |
| Brown | he (5), Walter | 6 |
| (John) Major | Smith, had, Major (2), diligently, | 3 |
| | confidently | |
| Bill | myself (3), her, him, Sr. Lopez | 6 |
| (Sean) Young | Connery (4), Penn (2) | 6 |
| Green | she (2), Maggie, Elizabeth, Lord | 6 |
| | Byron, I | |
| (Martin) Grace | Shaw (2), Grace, Scorsese, Sheen, | 6 |
| | Short | |
| (Brian) Wood | Robson, Gibson, Jones, Khan, Orser, | 6 |
| | More | |
| Frank | Peter, he (3), David Suzuki, Dick | 6 |
| (the) March (hearing) | committee, court (2), arraignment, | 0 |
| | public, preliminary | |
| Nick | Stephen, him (4), the DA, | 6 |
| Cook | Smith, then, inadvertently, | 3 |
| | enthusiastically, O'Toole, Pan | |
| Mike | Peter, him (2), Albert, herself, Brad | 6 |
| Bob | he (3), she (2), Trump, | 6 |
| Pat | you (5), Julia | 6 |

| Table 4. | Context | Rater Res | sponses by | 7 Target Name | (N = 6 adu | lt L1 English) |
|----------|---------|-----------|------------|---------------|------------|----------------|
| | | | | | | |

^a Target names were replaced by a blank. ^b 6 = rich context; 0 = no context.

Most (16) sentences with target names had rich context ratings of 5 or 6. Three names (*Major, White*, and *Cook*) had sentential context ratings of 3 (moderate context). One target name (*March*) had no context (0 rating) to indicate the item was a name. Test item #14 from Item Set A is shown in excerpt (1):

(1) 14. during the march hearing, the sheriff, city officials, firemen and other witnesses testified about the death.

One can see that the target item *March* is not being used as a name in this sentence but as an adjective. Indeed, all the L1 users entered common nouns or adjectives in this blank. Therefore, there were no contextual clues to indicate the item is a name.

Next, in order to investigate the relationship between the frequency of the target name and the participants' ability to correctly identify target names, the frequency ranking of each target name was determined from a search using an unpublished proper name frequency list based on the BNC (Parent, 2016). Table 5 shows how frequent each proper name is in English.

| Target name | Frequency ranking |
|-------------|-------------------|
| 1. March | 9 |
| 2. White | 203 |
| 3. Jack | 234 |
| 4. Mark | 310 |
| 5. Brown | 336 |
| 6. Frank | 385 |
| 7. Mike | 452 |
| 8. Bob | 453 |
| 9. Major | 458 |
| 10. Nick | 471 |
| 11. Young | 579 |
| 12. Wood | 652 |
| 13. Pat | 995 |
| 14. Cook | 1011 |
| 15. Grace | 1212 |
| 16. Bill | 1222 |
| 17. Green | 2018 |
| 18. Hill | 2080 |
| 19. Rose | 2349 |
| 20. Cliff | 23897 |

Table 5. Frequency Ranking of Target Proper Names in the BNC

From Table 5, one can see the majority (16) of target names were quite frequent, appearing in first 2,000 words. One notable exception was *Cliff*, which is relatively infrequent as a proper name, with a ranking of 23,897.

Lastly, to investigate the effect of the parts of speech of the corresponding target nonname items, an online dictionary, Longman's Dictionary of Contemporary English (LDOCE) (https://www.ldoceonline.com/dictionary) was used to determine which part of speech was most frequent. The LDOCE lists meanings in order of frequency, so that learners can see which meanings are most common (1995, p. xv). One might predict that the part of speech in which the target item occurs most frequently would be the form that the L2 reader would be most familiar with (Hoey, 2007). It might follow then, that for the non-name items that appear most frequently as nouns, the corresponding proper names might be more recognisable to the L2 reader in that grammatical position. Table 6 shows the parts of speech of the target non-name items in order of frequency. From Table 6, it can be seen that nearly half (9) of the non-name items appear most frequently, or only, as nouns: *rose, bill, wood, jack, hill, grace, nick, mike* and *cliff*. The items that appear most frequently as verbs are: *mark, march, bob, cook,* and *pat*. The items that appear most frequently as adjectives are: *white, major, green, frank, brown* and *young*. Regarding the part of speech that was used in the 20 sentences for non-name items, this is indicated in Table 6 with an asterisk.

| Target non-names | Parts of speech by frequency |
|------------------|--------------------------------|
| rose | noun, verb*, adjective |
| mark | verb, noun* |
| white | adjective*, noun, verb |
| major | adjective*, noun, verb |
| bill | noun*, verb |
| green | adjective*, noun, verb |
| wood | noun* |
| frank | adjective*, verb, noun |
| march | verb*, noun |
| bob | verb*, noun |
| jack | noun*, verb |
| hill | noun* |
| brown | adjective*, noun, verb |
| young | adjective*, noun |
| grace | noun*, verb |
| nick | noun*, verb |
| cook | verb, noun* |
| mike | noun*, verb |
| pat | verb, noun*, adjective, adverb |
| cliff | noun* |

Table 6. Frequency of Parts of Speech for Target Non-name Items

Note. Parts of speech marked with an asterisk (*) denote the part of speech used in the item sets.

The multiple regression was carried out using the total number of correct responses to each target name as the response variable, and the context rating, frequency ranking and most frequent part of speech as explanatory variables.

Post-hoc analysis findings

The standard multiple regression analysis indicated that when correctly identified names were predicted, it was found that none of the predictors were statistically significant: context rating ($\beta = .361$, n.s.) and frequency ranking ($\beta = .038$, n.s.). As for the predictor of the most common part of speech of the corresponding non-name target items, comparing those items that occur most frequently as adjectives to those items that occur most frequently as nouns was not predicted ($\beta = .169$, n.s.). Comparing items that occur most frequently as verbs to items that occur most frequently as nouns was also not predicted ($\beta = .008$, n.s.). The overall model fit was R² = .173.

Squared semi-partial correlations indicated that context rating had the largest contribution to the model ($sr^2 = .330$) while frequency ranking of the proper names had a much lower contribution ($sr^2 = .037$). The squared semi-partial correlation for those items that are

most common as adjectives was negative ($sr^2 = -.148$). This is an indication that if the most common part of speech of the corresponding non-name item was an adjective, then participants were less likely to correctly identify the target item as a proper name. The square semi-partial correlation for those items most frequent as verbs had the lowest contribution to the model ($sr^2 = .007$). However, as noted, none of predictors were significant.

Multiple regression has several assumptions and these were tested. Initial examination of scatterplots indicated linearity. There was not much multicollinearlity between the explanatory variables. P-P plots of standardized residuals indicated normal distribution. A scatterplot of studentized residuals and predicted value of standardized residuals indicated the assumption of homogeneity of variances was not completely met. However, the regression is thought to be robust enough for this not to have had a decisive influence.

DISCUSSION

The findings suggest that Japanese L2 low intermediate readers of English are not very successful at using sentential context to identify proper names. The participants in this study were able to use context to correctly identify names, on average, in 28.4% of the cases. In comparison, they correctly identified target non-names on average 91.4%. The difference between correctly identified names and non-names was statistically significant (p < .001), and the effect size was large (r = 0.87). Thus, sentential context was not very helpful for these participants to identify proper names. This is an important finding because of the assumption that exists in L2 vocabulary research that L2 readers can use context to recognise and understand proper names. The results from this study seem to contradict that assumption.

The post-hoc analysis was run to identify why some proper names were correctly identified while most (71.6%) were not. The multiple regression accounted for 17.3% of the variance, though none of the predictors were statistically significant. Richness of context made the largest contribution to the model. This suggests that context has a more important effect on participants' ability to identify names than the frequency of the name itself. However, the fact that frequency had little effect on the model may have come about because most (16) of the proper names in the study were quite frequent, appearing in the first 2,000 words of the BNC. The target items were chosen for this study because they were common words that the participants would be familiar with.

As for the predictor of the part of speech of the corresponding non-names, the multiple regression indicated that if the non-name item occurred most frequently, or only, as a noun, there was greater chance the participants were able to identify the proper name item. If the most common part of speech of the non-name item was a verb or adjective, participants were less likely to correct identify the target name. However, none of these predictors were statistically significant, making it impossible to draw any generalisable conclusions. The data set is an important consideration for the non-statistically significant findings: the post-hoc analysis examined the data in terms of target names, which resulted in a very small data set of 20 items. It is not likely to find a statistically significant result with such low statistical power.

There were some indications in the data that these participants, as L1 logographic readers, may have been relying on word recognition skills over contextual clues, as seen in some other studies (Ehrich et al., 2013; Huckin & Bloch, 1993; Parry, 1996). Recall that for the target non-name items, no action was required from the participant (i.e. the correct response was no addition of an initial capital letter). That no action was required may have contributed to the high scores for target non-name items. However, some participants wrongly identified some target non-names as proper names in 8.6% of cases (see Table 2), which might suggest a reliance on word recognition skills. There were three non-name items that were misidentified

as proper names by a significant number of participants: *bob* (by 52.7% of participants), *jack* (45%), and *mike* (35%). In the case of *bob* (shown in excerpt (2) from test item #18), the context in which it appeared might be seen as a classic example of leading readers down the garden path, with the phrase *it seemed to* preceding the item:

(2) 18. i kept my eyes on it, but it seemed to bob in front of us, keeping its distance like a mirage.

Participants may have stopped reading after the target item, interpreting the phrase as *it seemed to (someone)*. While the prepositional phrase *in front of* clarifies the part of speech is a verb, it is easy to see how participants could mistake the item for a name. An alternative possibility is that participants were familiar with *Bob* as a name, and relying on word recognition skills, did not check the context to confirm.

Looking at the two other examples of misidentified non-name items, many participants (n = 27) misidentified *jack* as a name. As shown in excerpt (3), the target item appeared in test item #3:

(3) 3. furthermore, this jack can deliver two different pickup signals or can be adapted.

Two contextual clues may have alerted the L2 reader that the item was not a name. First, the determiner *this* does not normally precede English proper names, unless one is distinguishing between different people called *Jack*. Another clue was the passive verb *can be adapted*. It would be unusual to refer to a person in this way. However, the participants may have stopped reading after the first verb (*can deliver*) and interpreted this as a clue that the item was a name. Alternatively, another explanation is that these L1 logographic readers relied on word recognition skills. This seems particularly plausible in the case of *mike*, as seen in example (4), for which the contextual clues seem to strongly favour a non-name response:

(4) 17. she stood at the mike and looked out at the white and hispanic faces of the congregation, and remarked, "welcome, all you pilgrims!"

The definite article precedes the item; also, the preposition at would not likely occur here (i.e. *stand at* + person). Thus, it seems likely that participants recognised the name and did not use contextual clues to confirm.

These examples point to a limitation of the study, that the inferencing strategies used by the participants remain unknown. Several of the studies reviewed above attempted to identify what strategies the readers used for lexical inference. For example, think-aloud protocol was used by Huckin and Bloch (1993) and Nassaji (2003b). Those studies had smaller sample sizes (3 and 21, respectively), making that approach more feasible. The sample size in this study (N = 61) was better in terms of quantitative analysis; however, think-aloud procedures with a smaller sample would have allowed for more insight into the inferencing strategies used by participants.

Because the results from the multiple regression were not statistically significant, conclusions cannot be drawn as to why some proper names were correctly identified and most were not. The results from the Wilcoxon Signed-Ranks test were statistically significant, with a large effect size, indicating that participants are not very successful in using context to identify proper names. It bears repeating that the aim of the study was to isolate the aspect of function, to determine whether this alone could aid in identification of proper names; in that regard, the study design did not represent an authentic reading task in that the initial capital letter was absent, which would normally provide a clue to the reader about proper names.

Therefore, the next step would be to isolate the aspect of form (i.e. the initial capital letter on English proper names). Indeed, it might be incautious to assume L2 readers have orthographic skills to process upper and lower cases efficiently; this is particularly true for those L2 readers processing a different writing system (Alderson, 2000). For example, L2 readers with a different L1 orthography might not be able to efficiently process letters that look similar in upper and lower cases (e.g. C, c; K, k; O, o). Thus, the assumption that the initial capital letter on proper names serves as an orthographic cue for L2 readers requires empirical support.

Another way in which the findings might not be applicable to authentic reading is that the participants could use the context of only one sentence to determine which items were names. While this is a valid point, it bears repeating that the participants were given ample time (15 minutes) to analyse twenty sentences. They were also encouraged to use their dictionaries to check the meaning of any unknown words. All participants finished the task in less than the allotted time, so it seems this was sufficient time to analyse the twenty sentences.

CONCLUSIONS

The results from the study suggest that L2 readers are not very successful at using context to identify proper names. The participants were able to use context to identify proper names in 28.4% of cases. While it is difficult to make direct comparisons, these results are very similar to those found in other studies investigating L2 readers' ability to use context to infer word meaning. For example, in Bensoussan and Laufer (1984), participants made correct inferences from contextual clues for 24% of the words. Likewise, in Nassaji (2003b), correct inferences from context made up 25.6% of responses. One might view such results in a positive light (i.e. when a reader can correctly infer the meaning of one in four words, this is quite good). However, in the case of proper names, the findings suggest that it should not be assumed L2 readers can easily identify and understand all proper names from context.

The post-hoc analysis did not indicate why some proper names might be easier to identify using context than others. Further research into L2 processing of proper names could employ qualitative approaches, such as think-aloud protocol, which might reveal more about how L2 readers make sense of the proper names they encounter. Also, by using slighter longer texts, researchers could look at how readers build meaning of proper names as they progress through a text.

If L2 readers cannot reliably recognise and identify most proper names in context, then this has implications for L2 vocabulary and reading research. Recall that because of an assumption L2 readers can use context to understand proper names, it has become standard practice in vocabulary analyses of text coverage counts to re-categorise proper names as known (i.e. placing these items in the 1K band, regardless of their actual frequency). In other studies, the proper names are removed from the vocabulary analyses altogether (e.g. Uden et al., 2014), making it impossible to ascertain what percentage of the text is proper names. The findings here, however, indicate that the assumption is imprecise, and therefore, a re-examination of how proper names are handled in vocabulary analyses is warranted.

The findings from the study also have pedagogical implications. Because proper names contribute to the meaning of a text (Allerton, 1987), it is important that adequate attention be given to proper names in classroom texts. As Kobeleva (2012) found in her study with L2 English listening, students who were pre-taught proper names perceived the comprehension tasks as easier to do, and self-reported higher comprehension than the group with unfamiliar proper names. There are various ways that teachers can draw attention to and check comprehension of proper names in texts. For example, students can be tasked with scanning for and circling proper names. Then, students can share what they know about each referent,

either from the text or from an Internet search. Comprehension should be checked of any titles that appear with proper names (e.g. *Lady Mary, Chief Birch*), as students might not understand the usage of such titles. In literary texts, teachers can draw attention to nicknames, which may seem obvious to L1 users but might not be easily inferred by L2 users (e.g. *Nicholas, Nickie*). The nuances behind the usage of diminutive forms should also be discussed. Teachers can support students by making sure they have a phonological representation of unfamiliar proper names because being able to pronounce names will lessen the burden on the short-term memory important for text comprehension. Finally, teachers can alert students to the patterns in L2 proper names (e.g. compare family names *Edwards* and *Adams* with personal names *Edward* and *Adam*). Being able to distinguish between personal and family names is important for academic citations. These are just a few strategies teachers can use to help their students become more familiar with L2 proper names.

ACKNOWLEDGEMENTS

I would like to thank Dale Brown for his comments on an earlier version of this paper.

REFERENCES

- Akamatsu, N. (2003). The effects of first language orthographic features on second language reading in text. *Language Learning*, *53*, 207-231. doi:10.1111/1467-9922.00216
- Alderson, J. C. (2000). Assessing reading. Cambridge: Cambridge University Press.
- Allerton, D. J. (1987). The linguistic and sociolinguistic status of proper names: What are they, and who do they belong to? *Journal of Pragmatics*, *11*, 61-92. doi:10.1016/0378-2166(87)90153-6
- Anderson, J. M. (2007). The grammar of names. Oxford: Oxford University Press.
- Bensoussan, M., & Laufer, B. (1984). Lexical guessing in context in EFL comprehension. *Journal of Research in Reading*, 7(1), 15-32. doi:10.1111/j.1467-9817.1984.tb00252.x
- Brown, D. (2010). An improper assumption? The treatment of proper nouns in text coverage counts. *Reading in a Foreign Language*, 22(2), 355 361.
- Coates, R. A. (2006). Properhood. Language, 82(2), 356-382. doi:10.1353/lan.2006.0084
- Cobb, T. (n.d.). VP-Compleat [computer program] (Version 2). Retrieved from https://www.lextutor.ca/vp/comp/
- Davies, M. (2004-). BYU-BNC. Retrieved December 2, 2017 http://corpus.byu.edu/bnc/
- Davies, M. (2008-). The Corpus of Contemporary American English (COCA): 520 million words, 1990-present. Retrieved December 2, 2017 http://corpus.byu.edu/coca/
- De Bot, K., Paribakht, T., & Wesche, M. (1997). Toward a lexical processing model for the study of second language vocabulary acquisition: Evidence from ESL reading. *Studies in Second Language Acquisition*, *19*, 309-329. doi:10.1017/S0272263197003021
- Ehrich, J. F., Zhang, L. J., Mu, J. C., & Ehrich, L. C. (2013). Are alphabetic language derived models of L2 reading relevant to L1 logographic background readers? *Language Awareness*, 22(1), 39-55. doi:10.1080/09658416.2011.644796
- Fraser, C. A. (1999). Lexical processing strategy use and vocabulary learning through reading. *Studies in Second Language Acquisition, 21*, 225-241.
- Gardner, D. (2013). Exploring vocabulary: Language in action. New York: Routledge.

- Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Cambridge: Cambridge University Press.
- Haastrup, K. (1991). Lexical inferencing procedures or talking about words: Receptive procedures in foreign language learning with special reference to English. Tubingen, Germany: Gunter Narr.
- Hanks, P. (2013). Lexical analysis: Norms and exploitations. Cambridge: MIT Press.
- Hirsh, D., & Nation, P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language*, 8(2), 689-696.
- Hoey, M. (2007). Grammatical creativity: a corpus perspective. In M. Hoey, M. Mahlberg,M. Stubbs, & W. Tuebert (Eds.), *Text, discourse and corpora*. London: Continuum.
- Hu, M. H., & Nassaji, H. (2012). Ease of inferencing, learner inferential strategies, and their relationship with the retention of word meanings inferred from context. *The Canadian Modern Language Review*, 68(1), 54-77. doi:10.3138/cmlr.68.1.054
- Hu, M. H., & Nassaji, H. (2014). Lexical inferencing strategies: The case of successful versus less successful inferencers. *System*, 45, 27-38. doi:10.1016/j.system.2014.04.004
- Huckin, T., & Bloch, J. (1993). Strategies for inferring word meaning in context: A cognitive model. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning*. Norwood, N.J.: Ablex Publishing Corp.
- Hwang, K., & Nation, I. S. P. (1989). Reducing the vocabulary load and encouraging vocabulary learning through reading newspapers. *Reading in a Foreign Language*, 6(1), 323 335.
- Klassen, K. (2021). Proper name theory and implications for second language reading. *Language Teaching*, 1-7. doi:10.1017/S026144482100015X
- Kobeleva, P. (2012). Second language listening and unfamiliar proper names: Comprehension barrier? *RELC Journal*, *43*(1), 83-98. doi:10.1177/0033688212440637
- Koda, K. (1996). L2 word recognition research: A critical review. *Modern Language Journal*, *83*, 51-64. doi:10.1111/j.1540-4781.1996.tb05465.x
- Koda, K. (2004). *Insights into second language reading: A cross-linguistic approach*. Cambridge: Cambridge University Press.
- Koda, K. (2005). Learning to read across writing systems: Transfer, metalinguistic awareness, and second-language reading development. In V. Cook & B. Benedetta (Eds.), Second language writing systems. Clevedon: Multilingual Matters Ltd.
- Lyons, J. (1977). Semantics (Vol. 1). Cambridge: Cambridge University Press.
- Nagy, W. E., & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly*, 19(3), 304-330. doi:10.2307/747823
- Nassaji, H. (2003a). Higher-level and lower-level text processing skills in advanced ESL reading comprehension. *The Modern Language Journal*, 87(261-276). doi:10.1111/1540-4781.00189
- Nassaji, H. (2003b). L2 vocabulary learning from context: Strategies, knowledge sources, and their relationship with success in L2 lexical inferencing. *TESOL Quarterly*, *37*, 645-670. doi:10.2307/3588216
- Nassaji, H. (2014). The role and importance of lower-level processes in second language reading. *Language Teaching*, 47(01), 1-37. doi:10.1017/s0261444813000396
- Nation, I. S. P. (2006). How large a vocabulary size is needed for reading and listening? *The Canadian Modern Language Review*, 63, 59-82. doi:10.3138/cmlr.63.1.59
- Nation, I. S. P., & Wang, K. M. (1999). Graded readers and vocabulary. *Reading in a Foreign Language*, *12*(2), 355 380.
- Parent, K. (2016). [BNC proper noun frequency list]. Unpublished raw data.

- Parry, K. (1996). Culture, literacy and L2 reading. *TESOL Quarterly*, *30*(4), 665-689. doi:10.2307/3587929
- Peressotti, F., Cubelli, R., & Job, R. (2003). On recognizing proper names: The orthographic cue hypothesis. *Cognitive Psychology*, 47, 87-116. doi:10.1016/S0010-0285(03)00004-5
- Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, 11(4), 357-383. doi:10.1080/10888430701530730
- Proper noun. (n.d.). *English Oxford Living Dictionaries*. Retrieved from https://en.oxforddictionaries.com/definition/proper_noun
- Quirk, R. (Ed.) (1995) Longman dictionary of contemporary English (3rd ed.). Harlow, England: Longman.
- Segalowitz, N., Poulsen, C., & Komoda, M. (1991). Lower level components of reading skill in higher level bilinguals: Implications for reading instruction. In J. H. Hulstijn & J. F. Matter (Eds.), *Reading in two languages* (Vol. 8, pp. 15-30): AILA Review.
- Shiotsu, T. (2009). Reading ability and components of word recognition speed: The case of L1-Japanese EFL learners. In Z. Han & N. J. Anderson (Eds.), Second language reading research and instruction: Crossing the boundaries (pp. 15-39). Ann Arbor, MI: University of Michigan Press.
- Strawson, P. F. (1950). On referring. Mind, 59(235), 320-344.
- Uden, J., Schmitt, D., & Schmitt, N. (2014). Jumping from the highest graded readers to ungraded novels: Four case studies. *Reading in a Foreign Language*, 26(1), 1-28.
- Webb, S., & Rodgers, M. P. H. (2009). Vocabulary demands of television programs. Language Learning, 59, 335-366. doi:10.1111/j.1467-9922.2009.00509.x

Kimberly Klassen completed her PhD in applied linguistics (lexical studies) at Cardiff University, Wales. She is currently an instructor at the International Christian University, Tokyo, Japan. Her essay on proper name theory and implications for second language reading was awarded the Christopher Brumfit Essay Prize 2020, by Language Teaching and Cambridge University Press. Other publications have appeared in Reading Psychology and The Curriculum Journal. She has taught English as a foreign language in Korea, Taiwan, the UAE and Japan.

Email: <u>kimi.klassen@gmail.com</u>

APPENDIX

ITEM SET A

Instructions: Read the sentences. Change the small letters to capital letters if necessary. You can use your dictionary if necessary. Look at the example.

I N Y Example: i arrived in new york last night.

1. the doctor offered rose slimming tablets to help her lose even more weight.

- 2. he suspects babur is a secret smoker because sometimes mark can smell the tobacco on him.
- 3. furthermore, this jack can deliver two different pickup signals or can be adapted.
- 4. it was agreed that the statement should come from the white house, with immediate confirmation in london.
- 5. it is very difficult to work out owing to the tremendous variation in hill ground types.
- 6. james pointed to the brown envelope christina still clutched in her hand.
- 7. after winning, john major patrolled the commons tea-room, soliciting opinions on the next debate.
- 8. next morning i continued to make excuses for bill, but as i now knew i was making them, they sounded false.
- 9. most adults do it almost unthinkingly, but for young children it's a painstakingly complicated business.
- 10. in magisterial style, green dealt peremptorily with the committee's inflexible attitude.
- 11. in his eyes their saving grace was something he could only define as that truly human feeling.
- 12. brian wood was the strong, solid, ever-dependable central defender.
- 13. the most likely cause of the dark spots, frank reasoned, was water, a common molecule that absorbed at the wavelengths detected by his camera.
- 14. during the march hearing, the sheriff, city officials, firemen and other witnesses testified about the death.
- 15. if this were a movie, you'd think that jeremy and his mother were escaping in the nick of time.
- 16. can the typical cook finish three of ray's recipes in 30 minutes?
- 17. she stood at the mike and looked out at the white and hispanic faces of the congregation, and remarked, "welcome, all you pilgrims!"
- 18. i don't get the feeling bob is under a lot of pressure.
- 19. i gave him a loving pat as i went by.
- 20. he shook himself, scrambled up the cliff and disappeared into a crevice.

ITEM SET B

Instructions: Read the sentences. Change the small letters to capital letters if necessary. You can use your dictionary if necessary. Look at the example.

I N Y Example: i arrived in new york last night.

- 1. the unemployment rates in all advanced industrial societies rose substantially during this period.
- 2. only one of the landscape planners and artists left their mark in this biggest and best known park of berlin.
- 3. she led her out of the nursery and then julie picked her up when jack came out of the school.
- 4. he looked at young john, now white and shaken with the shock.
- 5. bernard hill played a maverick detective in last night's new drama telltale.
- 6. chesham secretary tony greeham said brown resigned for personal reasons, and that there was no animosity.
- 7. all russia's major rivers are estimated to have between 10 and 100 times the safe limit.
- 8. a town in scotland managed to cut 6,000 from its annual water bill when an employee pointed out an unnecessarily wasteful use of water.
- 9. i asked warren beatty not to cast sean young in the role for dick tracy.
- 10. we took the road to harar, through mountains that were beautiful and green after recent rain.
- 11. martin grace had to run along the top of a train doubling for roger moore.
- 12. when she went downstairs henry was feeding wood into the stove in the living-room.
- 13. if you can't have frank conversation in these meetings, then you can not trust each other.
- 14. salazar left after an hour to march in the saturday-morning parade celebrating the opening of the state fair.
- 15. the fact that he could not drive the image from his mind, struck nick as a proof that yvette's beauty had been her curse.

- 16. peter cook revealed another deeply held secret that shocked and embarrassed everyone in the court.
- 17. sometimes she pictured mike and thought about a future in which they had a child.
- 18. i kept my eyes on it, but it seemed to bob in front of us, keeping its distance like a mirage.
- 19. take those cucumbers pat admired so much.
- 20. what if cliff suddenly appeared and caught me cuddling with his wife?