Investigating the lexical load of proper names for L2 English readers

Kimberly Klassen

Centre for Language and Communication Research

School of English, Communication and Philosophy

Cardiff University

Thesis submitted for the degree of PhD

Cardiff University

June 2018

Abstract

The central question addressed is whether proper names present a strain for second language (L2) readers. Answering this question helps to establish the soundness of a widely held assumption in L2 vocabulary research, that L2 readers can easily recognise and understand proper names from the form (capitalisation) and function (context) in a text. The investigation is motivated by classroom experience that contradicts this assumption, suggesting a need for reconsideration of how proper names are handled in L2 vocabulary research and language pedagogy. The assumption was approached from three angles using a series of experiments. First, interviews were conducted to investigate how L2 readers perceive proper names and what strategies they use. Another study investigated how L2 readers approach unfamiliar proper names while reading, and found some L2 readers treat proper names as vocabulary to check in a dictionary. The second direction investigated the effect of L2 proper names on higher-level comprehension processes. Two studies compared the effect of culturally familiar and unfamiliar proper names on comprehension, and found no effect for culturally familiar proper names on global comprehension. The third approach considered proper name processing in terms of lower-level reading skills (i.e. word recognition and sub-skills). A study was conducted to determine to what extent L2 readers can identify proper names in context and found that participants were not very successful at using context to identify proper names. Three main claims are based on these results. First, proper names can disrupt reading in that some L2 readers treat them as unknown vocabulary to look up. Second, proper names do not seem to impact global comprehension. Third, L2 readers are not very successful in correctly identifying proper names from context. Based on these results, L2 vocabulary researchers and teachers should consider the potential burden proper names can place on L2 readers.

Declaration

This work has not been submitted in substance for award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Statement 1

This thesis is being submitted in partial fulfilment of the requirements for the degree of PhD.

Statement 2

This thesis is the result of my own independent work/investigation, except where otherwise stated, and the thesis has not been edited by a third party beyond what is permitted by Cardiff University's Policy on the Use of Third Party Editors by Research Degree Students. Other sources are acknowledged by explicit references. The views expressed are my own.

Statement 3

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loan, and for the title and summary to be made available to outside organisations.

Statement 4

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loans after expiry of a bar on access previously approved by the Academic Standards & Quality Committee.

Signed ……..……

Date ……..

Acknowledgements

I would like to thank my supervisors, Dr. Lise Fontaine and Dr. Andreas Buerki, for their guidance and suggestions for bringing this thesis to completion. I would also like to thank my former supervisor Professor Tess Fitzpatrick, who patiently helped me over many years to tease apart the issues. I also thank Professor Alison Wray for her advice and encouragement.

Special thanks goes to Mike for being so supportive and understanding. I also express my gratitude to my parents for their interest and encouragement. Thanks to my sisters: the one who taught me to read, and the other who encouraged me to write.

Lastly, I appreciate the students who generously participated in this research.

Table of Contents

List of tables

Chapter 1: Teacher, what's a Craig?

1.1 Introduction

Reading is a skill that perhaps many people do not fully appreciate for all its complexity. Proficient readers especially might take for granted the sophisticated processes involved when they decode symbols on the page to make meaning in a seemingly effortless manner. Cognitive neuroscientist Stanislas Dehaene is one researcher who advocates for a greater appreciation of reading. He reminds us that we are primates whose brains did not evolve to read, a process he defines as the transforming of texture[1] into speech (Dehaene, 2009). Central to his passion for the feat of reading is his neuronal recycling hypothesis: neurons in the brain, which were originally used for visual recognition, have been repurposed for reading. Since writing developed only about 5,000 years ago, not enough time has passed in evolutionary terms for specialised reading circuits. In sum, areas of the brain have been recycled for the purpose of reading – an amazing feat indeed.

Many people who can read in their first language go on to learn to read the 'texture', or writing system, of a second language (L2). For L2 learners, reading can be a means to learn more about the target language itself as well as its associated culture. For example, through extensive reading (i.e. reading a large amount of material at an appropriate vocabulary level), L2 readers can enhance their existing grammatical and vocabulary knowledge, learn new vocabulary, improve general reading and critical thinking skills, as well as learn about the culture and history of the target language (Day & Bamford, 1998; Nation, 2008). Reading is also important for L2 users who are studying in a formal setting, in that reading will almost certainly be used to assess their language proficiency.

In any assessment of the L2 reader, it is important to view her not as necessarily less proficient than the first language (L1) reader, but as different from the L1 reader. As Cook and Bassetti (2005) point out, users of second language writing systems (L2WS) "read, write, learn and analyse their L2WS *differently* from L1WS users, because they have more than one writing system in their minds" (p. 45). Thus, it is important to view

[1] Dehaene is using 'texture' to mean the appearance of a surface, in this case, the appearance of printed words on the page.

differences in the L2 reader not as incompetence, but as what Cook (2002) refers to as multi-competence: "[when] knowledge of more than one language exists in the same mind" (p. 10). In this regard, L2 reading researchers and teachers need to acknowledge the influence that the reader's L1 has on her L2 reading skills.

Reading skills are generally classified as involving either lower-level processing or higher-level processing. Lower-level skills refer to processing of text-based information including: word recognition, with sub-skills of phonological, semantic and orthographic processing; syntactic parsing; and semantic-proposition encoding (i.e. building meaning from semantic and grammatical information) (Grabe, 2009). In contrast, higher-level processing is generally defined as reader-based: it concerns what the reader brings to the text to build comprehension. This can include, for example, the reader's background knowledge, inferencing skills, and strategy use. Lower- and higher-level processes are thought to be interactive and hierarchical, but not reciprocal (Nassaji, 2014). This means lower-level processing must be efficient and automatic for successful higher-level processing to occur; no amount of higher-level processing can compensate for inefficient lower-level processing.

Traditionally, much L1 reading research has focused on lower-level processing skills, in particular, how L1 children learn that the texture on the page translates into language (Grabe, 2009). Reading models that have driven L1 research are known as 'bottom-up' because they postulate that reading begins with lower-level processes and moves up to higher-level processes. For example, Gough's Model (1972), also known as the Simple View of Reading, suggests that successful reading depends on decoding skills and language comprehension. Another early influential model was Automatic Information-Processing Model (LaBerge & Samuels, 1974), which emphasised the automaticity of decoding skills. Later models predicted that the two levels of processing are interactive (e.g. Rumelhart, 1977). These reading models are examples of cognitive-processing theories that have been influential to L1 reading research since the 1970s (Tracey & Morrow, 2017).

Much L2 reading research, in contrast, has been driven by 'top-down' theories of reading, that is, how the reader directs comprehension in terms of goals and strategies (e.g. Goodman's (1967) Psycholinguistic Guessing Game theory of reading). The reason for this focus on higher-level processing was twofold. First, there has been a belief that reading is universal, no matter what language it is done in: if the reader already understands that texture on the page converts into language, then the reason

for any lag or deficiency is presumed to be the result of deficiencies in higher-level processing. Second, there has been a belief that lower-level processing skills are correlated with L2 proficiency. This means that as the L2 user displays more native-like competency, their lower-level processing skills are assumed to have developed at a comparable rate.

Most contemporary reading models acknowledge the importance of both lower- and higher-level processing (Grabe, 2009; Hudson, 2007; Nassaji, 2014). It is also generally agreed that reading is not universal across languages; that is, readers use different processing strategies for different writing systems (Ehrich, Zhang, Mu, & Ehrich, 2013; Koda, 2005). However, an assumption seems to prevail that L2 readers possess efficient lower-level processing skills as a by-product of their language proficiency, and any deficiencies must therefore lie with higher-level processing. Evidence for this assumption is found in both the research literature and in L2 English reading materials. For example, Eskey (1988) expressed concern that despite the importance that interactive reading models place on both levels of processing, L2 reading literature continued to show a bias toward top-down processes. Writing two decades later, Birch (2007) notes that Eskey's concern is still pertinent, as much L2 reading research remains focused on higher-level processing. Another illustration of the assumption that L2 readers possess efficient decoding skills is seen in L2 English reading textbooks. A quick perusal of some current English as a Foreign Language (EFL) textbooks reveals an almost exclusive emphasis placed on training in higher-level skills, such as predicting content, activating background knowledge, and using context to guess the meaning of words. Conversely, almost no space is devoted to improving automaticity of lower-level skills through word recognition activities, speed reading tasks, or promotion of extensive reading (Grabe, 2009; Grabe & Stoller, 2011).

Thus, there is a prevalent notion that the L2 reader is competent in lower-level decoding skills, and this has led to assumptions about the L2 reader that have been seldom tested. One such conjecture, that will be the focus of this thesis, is an assumption that L2 readers can identify and understand proper names in context. It is common currency among researchers of L2 English vocabulary that the form (initial capital letter) and the function of the name (the context) will signal to the reader that the item is a proper name. On the surface, this seems like a reasonable assumption to make. Indeed, most L2 learners of English possess the declarative knowledge that names of people, places, companies and products require an initial capital letter. However, embedded in this belief is the expectation that L2 readers have the lower-

level decoding skills to consistently and efficiently identify capital letters in continuous text, and furthermore, are skilled at using the context to make inferences about what or who the name refers to. As an EFL teacher for over 20 years, I have much anecdotal evidence that indicates this is not always the case.

I first noticed the difficulty that L2 readers of English can have with proper names when working with Arab learners in the Gulf region. I had implemented an extensive reading program in my classes: the low-intermediate students self-selected and read graded readers (i.e. books in which the vocabulary has been graded to match the reader's vocabulary level). Because many of the students had little or no experience reading for continuous periods of time, even in their L1, extensive reading training took place during class time in the form of weekly Sustained Silent Reading (SSR) sessions (Nation, 2008). During these sessions, students were asked to keep a record of new vocabulary they had come across in their graded readers. This task was set to ensure they were reading at an appropriate level where 98% of the vocabulary was known (Nation, 2006; Schmitt, Jiang, & Grabe, 2011). This level would allow for adequate comprehension (i.e. approximately 70%) (Schmitt et al., 2011), and ensure favourable conditions for uptake of new vocabulary. During one SSR session, a student came to show me her vocabulary list. The first word on her list was:

(1) *n. jack – a tool for lifting heavy objects*

Because this word struck me intuitively as low frequency, I asked to see which book she had been reading: *Jack and the Beanstalk*. The character *Jack* appeared on the cover of the book as well as in almost every picture inside. Furthermore, *Jack* was used in context with verbs like *climbed, said,* and *ran*. The initial capital letter was another clue that the item was a name. According to the assumption in L2 vocabulary literature, all of these clues should have indicated to the reader that *Jack* was a person, not a tool for lifting heavy objects. However, the illustrations, the context, nor the orthography helped this reader: she was surprised to learn that the boy was named *Jack*.

This incident stayed with me because at the time, I was familiar with the work of some L2 vocabulary researchers, like Paul Nation. His view on proper names was that L2 readers should be able to easily recognise and understand them (e.g. Nation, 2006). So, I filed the incident away in my mind, attributing it to a fluke with a low-intermediate learner, one who had little reading experience even in her L1. But the incident made

me aware of the possibility that L2 learners could have difficulty with names. Shortly afterwards, I moved to Japan where I noticed more cases of readers having difficulties with proper names.

The learners I have worked with in Japan have been at an intermediate proficiency level, so slightly higher than the Arab learners. But even at this higher level, they too demonstrate problems with proper names. For instance, in one academic reading skills textbook, there was a narrative text about a woman who undertakes a difficult hike despite having a precarious heart condition. Her husband accompanies her on this hike. In the 800 words or so of text, he is referred to by name six times, as seen in the excerpts (2) to (7):

(2) *. . . my husband Craig and I began to hike the trail . . .*
(3) *Craig, observing the daunting task ahead, gently asked . . .*
(4) *Step for step, Craig stayed directly behind me . . .*
(5) *Craig and I made our way over to the edge.*
(6) *. . . Craig surprised me with a gold charm . . .*
(7) *Craig took a moment to express how proud he was of me . . .*

After the class had read the text along with a CD recording, I asked if there were any vocabulary questions. One student put up her hand and asked, “What is a Craig?” It should be noted that Japanese students are generally very reticent to ask questions in class (King, 2013), so I took her question as sincere. The contextual clues surrounding Craig were rich, as can be seen in examples (2) to (7). Note that his name appeared at the beginning of three sentences in examples (3), (5) and (7), and so the initial capital letter was not overtly helpful in those sentences. The initial capital letter in mid-sentence position in examples (2), (4) and (6) was a clue. When another student in the class whispered to her that Craig was the husband’s name, the student was very embarrassed and never asked another question for the rest of the semester.

The difficulty these learners had with regard to names was not confined to academic texts or graded readers but showed up in other situations. For example, the difficulty with proper names presented itself in an orientation activity I conducted with first-year university students. Even though Japanese students study English for six years before entering university (at the time of writing), many of them have never used a textbook printed by an English publisher. So, one of the orientation activities is a ‘get to know your textbook’ worksheet, designed to get students looking around the book to see how

it is organised and to locate certain information. One question is: *Who are the authors?* The authors for this particular textbook were Margot F. Gramer and Colin S. Ward. I have done this orientation activity for this textbook with three different classes, and each time, about of a third of the students answer the question (*Who are the authors?*) this way:

(8) *Margot F. and Colin S.*

When I asked them to explain why they had written only part of the authors' names, they replied: "Margot F. does the grammar and Colin S. does the words." After hearing this explanation, I realised these particular authors have wonderful names as English textbook writers. Clearly, the students were tuned in to the concepts of grammar and vocabulary in relation to the EFL textbook. However, one can argue that both the context and the form have failed to help them correctly identify the names. The context, how the names were presented on the cover, should have identified these items as names; likewise, the initial capital letter should have alerted the readers. One might suggest that the students' misunderstanding was due to an error in higher-level processing, specifically, a lack of cultural knowledge in how English names are presented. Alternatively, this sort of difficulty may be attributable to inefficient lower-level processing skills, in that the misunderstanding occurred during letter and word analysis. Anecdotal evidence like the above has convinced me, as a language teacher, that proper names are not wholly unproblematic for L2 readers.

1.2 Research aims

The central aim of this thesis is to investigate whether proper names present a burden for L2 readers. This is important because of an assumption in the research literature that L2 readers can easily recognise and understand proper names. Investigating this assumption contributes to addressing an overarching theoretical aim: to determine the extent to which proper names are part of an L2 reader's linguistic system; that is, whether proper names represent lexical or encyclopaedic knowledge. If proper names represent lexical knowledge, then an effect might be seen with lower-level processing level (e.g. word recognition and lexical access). If proper names represent encyclopaedic knowledge, then an effect might be seen with higher-level processing (e.g. the reader's background knowledge).

The investigation approaches the central research question through a series of empirical studies. The findings from the studies are used to advance both the central aim and theoretical aim. Specific research questions are addressed through the set of experiments; each research question connects to a substantive aim. The first substantive aim is to learn how L2 readers approach unfamiliar proper names when reading continuous text. The second aim is to explore the effect of proper names on higher-level comprehension processes; specifically, how cultural background knowledge of proper names affects L2 reading comprehension. This substantive aim also connects to the theoretical aim of the thesis, to determine the extent to which proper names might represent encyclopaedic knowledge for an L2 reader. The third aim is to investigate proper name processing as an aspect of lower-level reading skills; specifically, L2 readers' word recognition skills and their ability identify proper names from context. This aim also relates to the theoretical aim, to determine the extent to which proper names might be considered an aspect of an L2 reader's lexical knowledge. More details on how the set of experiments addresses each of these aims are given below in an overview of the thesis structure (section 1.4). But first, some implications of this research into L2 proper name processing are noted.

1.3 Implications of the research

There are theoretical implications of this research for L2 lexical processing and reading research. For example, in L2 vocabulary and reading research, the difficulty of reading texts can be evaluated in terms of vocabulary coverage. In a lexical analysis of a text, the researcher can treat proper names as encyclopaedic knowledge and therefore, are known items to the L2 reader. Such a treatment of proper names follows the widely accepted philosophical view that proper names refer but do not have meaning (Lyons, 1977). However, if the researcher treats proper names as lexical items, the assumption cannot be made that the reader will be familiar with all the proper names in the text. This conceptualisation of proper names as lexis follows a view held by some linguists that proper names can have categorical meaning, albeit it minimal (J. M. Anderson, 2007; Van Langendonck, 2007). More on this discussion of the theory of proper names follows in Chapter 2.

It is in the area of L2 vocabulary research that an assumption is made regarding L2 readers' understanding of proper names. In the late 1980s and early 1990s, both L1

and L2 reading research began to consider text difficulty with a focus on vocabulary. While prior L1 readability indices had focused on factors of syllable, word and sentence length to determine text difficulty, this later research was less concerned with syntax and more focused on specific words found in a reading text. For example, L1 research (Carver, 1994) considered the ease or difficulty of reading texts in terms of the percentage of vocabulary known to the reader. Around the same time, L2 vocabulary researchers (e.g. Hirsh & Nation, 1992; Hwang & Nation, 1989) began to look at different types of reading texts and how learners might learn new vocabulary from such texts. Later, as computer-based language corpora became more widely available, researchers were able to make frequency counts of the vocabulary in texts. Computer programs like Range (Nation & Heatley, 2002) were developed, whereby a vocabulary profile of a text could be generated. This profile shows which words in the text belong to the first band of 1,000 most common words (1K), which belong to the second band of 1,000 most common thousand words (2K), and so on. However, one problem for vocabulary researchers was what to do with the proper names.

Proper names do not behave like other words. According to the lexicographer Patrick Hanks (2013), “Proper names are special kinds of words, with special rules governing their role as conventional units of a language” (p. 64). Perhaps to the layperson, one obvious difference between proper names and other common words as units of language is that the former are not usually included in dictionaries. Hanks (2013) notes that dictionaries which do not include proper names seem to operate on the assumption that words denote classes, not individuals; on the other hand, dictionaries that do include proper names seem to operate under the assumption that words include all items of culturally shared knowledge (p. 34). As an example, Hanks (2013) remarks that if you do not know who Shakespeare was, then you are not a full-fledged member of the English speaking community (p. 34). Such cultural knowledge of names is frequently exploited, whereby a generic set of individuals is referred to by a proper name. Hanks (2013) offers an example from a newspaper article where the name *Sherlock Holmes* is used to refer to a detective: “A blind Sherlock Holmes: detective fights crime with acute listening” (p. 35). While the subheading of the title aids comprehension, it certainly helps to understand the reference to *Sherlock Holmes.*

When considering how to handle proper names in their vocabulary profiles of texts, researchers may have looked to dictionaries as authoritative sources on lexis for direction on how to treat proper names. Essentially, a decision is taken whether to treat proper names as encyclopaedic or lexical knowledge. If proper names represent world

knowledge, then one can assume proper names are known. For example, some researchers have suggested that proper names will have been learnt in the L1, and therefore can be treated as world knowledge. On the other hand, if proper names represent lexical knowledge, then one cannot assume this information is known to the L2 user. That is, proper names might consist of (minimal) categorical meaning, such as gender or place, information that the L2 reader might not be cognisant of. A more detailed discussion regarding researchers' treatment of proper names as known vocabulary is given in Chapter 4. Here, it is suffice to point out that it has become common practice in L2 vocabulary research to treat proper names as known vocabulary, or even delete proper names from analysis in vocabulary coverage counts (see D. Brown, 2010).

Research into how L2 readers process proper names also has practical implications for the classroom, as has been illustrated above. Teachers and materials developers need to be attuned to any potential burden caused by proper names. For example, not all cultures present proper names in the same order as English speaking cultures, nor do all cultures have concepts of middle names or middle initials.[2] These differences can cause confusion for learners, in particular for, but not limited to, instruction for academic citations. Also, learners may not understand titles that can precede names and the status signified by such titles (e.g. *Lady Mary)*. Furthermore, L2 readers may not be privy to the characteristics implicit in names in literature (e.g. *Goldilocks)*, which L1 children are taught from a young age to identify (Crystal, 2006). Therefore, teachers and textbook writers need to be aware of any potential difficulties that L2 proper names can present.

The implications of research into proper names extend beyond the classroom to include areas of independent learning. As was noted above, there are many benefits of extensive reading, and ideally, learners will undertake extensive reading outside of the classroom (Day & Bamford, 1998). However, Hill (2013) notes that some authors of graded readers often ignore the potential burden of names, introducing many characters in the first few pages of a story. It is difficult to remember names one has never encountered before; it is even more difficult when several names are introduced in a short space, as this creates competition in the reader's short-term memory. Thus,

[2] For example, Japanese, Chinese and Korean names are presented as family name first, followed by the personal. In Arabic cultures, most people have a chain of names that often include the names of the father and grandfather.

materials created for independent learning purposes should consider difficulties that L2 proper names can present.

Another area of concern for how proper names affect L2 reading comprehension is in language proficiency testing. As was noted, reading will almost certainly be part of the assessment for L2 learners in formal education and professional settings. If test writers assume that proper names and their associations are known to the L2 reader, this might disadvantage test-takers who do not possess the relevant knowledge needed to make sense of proper names in the text. For example, important considerations for test writers include the background knowledge inherent in proper names, such as the gender of the person, the location of a place, the type of product related to a brand name, and so on. Especially for high-stakes tests, it should be a concern if such knowledge is assumed, as it can put some test-takers at a disadvantage.

1.4 Overview of thesis structure

As part of this enquiry into how proper names affect L2 reading comprehension, literature from three main fields is reviewed. The first is linguistic theory of proper names, which has been strongly informed by philosophical theories of proper names. Thus, concepts from the philosophical debate surrounding proper names are also briefly referred to. Second, L2 vocabulary research is reviewed as it pertains to proper names and reading comprehension. The third field is L2 reading research. Because L2 reading research has been informed by L1 research, models and theory from L1 reading research are also referred to.

In Chapter 2, the stage is set with a review of literature to address three key aspects of the investigation: proper names, reading processes, and how L2 vocabulary knowledge impacts reading comprehension. I begin with a review of proper name theory, focusing on contemporary linguistic perspectives of proper names. The discussion considers whether proper names are lexical or encyclopaedic knowledge. I move towards a working definition of 'proper name' as it will be used in this research context. Next, a brief overview is given of what is currently known about reading processes. Finally, L2 vocabulary research is considered for implications relevant to L2 reading; specifically, how the percentage of known vocabulary impacts reading comprehension.

Chapter 3 presents a qualitative study consisting of interviews and a read-aloud task with four Japanese L2 English learners. This exploratory study aims to investigate how L2 readers view proper names: how they feel when meeting unfamiliar proper names in texts; what strategies they use for handling unknown names; and any difficulties they perceive with proper names. In the read-aloud task, participants are asked to read aloud an excerpt from level-appropriate material that contains proper names of several different characters and places. The participants are asked to say to what or whom the names refer. In this way, the aim is to identify any difficulties in decoding or understanding of proper name referents. Also, the strategies these readers use when confronted with unknown proper names might be identified. Some of the findings from this initial study are used to motivate the investigation and direct the subsequent studies in Chapters 4 to 6.

Chapter 4 begins with a comprehensive review of the literature concerning an assumption found in L2 vocabulary research: that L2 readers can easily recognise and understand proper names in context, and therefore, proper names can be treated as known vocabulary. I argue why this assumption warrants testing. Following the review, findings are reported from my study that attempts to fill this gap in the literature. The study investigates what Japanese L2 English readers do when they encounter unfamiliar proper names. The study draws on methodology used in L1 reading research (Carver, 1994). Participants are presented with reading texts of varying difficulty and are asked to mark all unknown vocabulary. Then, choosing from marked words, they prioritise the items to check in a dictionary. The texts that participants marked up and their look-up lists are analysed for the presence of proper names. A comparison is made between the proper names identified as unknown to other lexical items marked as unknown.

In Chapter 5, the topic of L2 proper names is examined from a higher-level processing perspective. Specifically, how the role of cultural familiarity with proper names affects reading comprehension is considered. Just as Chapter 4 begins with a focused review, Chapter 5 begins with a detailed review of several studies which investigated the role of cultural knowledge on reading comprehension: two earlier classic studies (Johnson, 1981; Steffensen, Joag-Dev, & Anderson, 1979) and two more recent studies (Alptekin, 2006; Erten & Razi, 2009). Because all of these studies are underpinned by schema theory, this theory is described and discussed as it has been applied to reading research.

Following the review, two empirical studies are presented that look at the effect of cultural familiarity with proper names on reading comprehension. The first is an approximate replication study of Erten and Razi (2009). Because much of the cultural knowledge in their study concerns proper names, their methodology is considered ideal for replication to explore the effect of culturally familiar proper names. The basic premise is to manipulate the proper names and other cultural referents in an American short story to be more culturally familiar to the participant group (i.e. Japanese L2 readers of English). Reading comprehension is compared between groups that read the text with either culturally familiar or unfamiliar proper names and referents. Because post-hoc analysis reveals unreasonable vocabulary demands on the participant group, a second study is run to investigate the effect of proper name familiarity. In the second study, the vocabulary level of the reading material is controlled for, and reading comprehension is compared among three treatments: culturally familiar names (L1 proper names), culturally unfamiliar names (L2 proper names) and no proper names, only common nouns.

In Chapter 6, proper names and L2 reading comprehension are investigated from a lower-level processing perspective. Again, the chapter begins with a focused literature review, which demonstrates the importance of lower-level processing skills to L2 reading comprehension. Factors that can impact L2 lower-level skills are considered: the distance between the reader's L1 and L2, L1 transfer effects, and L2 orthographic processing experience. Following the review, one final empirical study is presented, which investigates the extent to which Japanese L2 English readers can use context to identify proper names. Using context to infer meaning is a skill that draws on the reader's lower-level processing skills to include word recognition, and the sub-skills of orthographic, semantic and syntactic processing. Participants are presented with authentic sentences containing target items that are semantically ambiguous; that is, the items can be used both as proper names and as common words (e.g. *Jack, jack*). Also, the orthographic processing has been disrupted: all initial capital letters have been replaced by lower case letters. The participants' task is to add capital letters where necessary, thus identifying proper names from the sentential context. Correct responses to target proper names are compared to correct responses to target non-names. A multiple regression is run to explore predictors of context richness, frequency ranking of proper names, and parts of speech of non-name target items.

The success of each of these five empirical studies is discussed in Chapter 7. Specifically, the success of each study in terms of meeting its research, theoretical and

methodological aims is considered. The applicability of the findings to a wider context is also discussed, given the important factor of the participants' L1 (i.e. Japanese). Implications of the findings from this investigation are presented for language pedagogy, materials development and testing. The thesis ends with a concluding chapter that articulates the contribution this investigation has made in determining whether proper names present a burden for L2 readers. Several directions are suggested where the research could be taken further.

Chapter summary

This introductory chapter first motivated the research by demonstrating how experiences in the L2 reading classroom do not correspond to an assumption made in the research literature: that L2 readers can recognise and understand proper names from the form and context. The central research aim was stated, which is to investigate whether proper names present a strain to L2 readers. An overview of the thesis was presented, showing how proper names are investigated from three distinctive aspects. The first is from the perspective of L2 readers: how they approach unfamiliar proper names when reading; what strategies they use for unknown names; and what affective factors (e.g. anxiety, apathy) are involved in proper name processing. Then, proper names are looked at in terms of higher-level processing, in particular, as an aspect of cultural background knowledge. Finally, proper names are considered from the perspective of lower-level processing skills, specifically, word recognition and semantic-proposition encoding (i.e. building meaning from semantic and syntactic information). As noted, each of the three perspectives will begin with a narrowly focused review of the literature pertinent to that theme. The literature reviews are followed by empirical studies.

The next chapter presents a short review of terms and concepts that are relevant to each of the three perspectives. The term 'proper name' is defined as well as other terms relevant to reading processes and writing systems. How vocabulary knowledge and background knowledge can impact L2 reading comprehension is also discussed.

Chapter 2: Proper names, L2 vocabulary, and reading processes

2.1 Introduction

The central research aim, as stated in the previous chapter, is to investigate whether proper names constitute a strain for L2 readers. This central aim is furthered through three areas of enquiry: investigating how L2 readers approach proper names; looking at higher-level processing of proper names, specifically, the effect of cultural familiarity on comprehension; and investigating lower-level processing of proper names in terms of word recognition, and use of semantic and syntactic information. While each of these three areas of exploration begins with a focused literature review, there are specific terms and concepts that apply to the overall thesis. Defining these terms and concepts is the focus of this chapter.

I begin with the term 'proper name'. While proper names might seem straightforward, there has been much debate concerning whether proper names belong to the language system, and whether they have meaning. This debate is reviewed to address the important issue in this thesis of whether proper names are part of an L2 reader's vocabulary knowledge or world knowledge. Another important consideration for all three areas of the enquiry is the relationship between an L2 reader's vocabulary knowledge and reading comprehension. Because this investigation is focused on how proper names impact comprehension, it is important to consider key factors that can affect comprehension. To that end, a robust study is reviewed that investigated the interrelationship between the percentage of known vocabulary, background knowledge and reading comprehension. Lastly, there are various terms related to reading processes and writing systems that are used in all subsequent chapters, so they are defined here.

2.2 Proper names

A proper name is a word or groups of words used to refer to an individual person (actual or potential), place, or organisation; in English, proper names are marked by an initial capital letter ("Proper noun," n.d.). This simple dictionary definition belies the complexity surrounding philosophical and linguistic theories of proper names. Because linguistic enquiry into proper names has been strongly informed by philosophical

theorising, this section begins with the fundamental philosophical positions on proper names. The issues are part of ongoing debate, and there is a need to restrict this discussion to what is directly relevant to how proper names might affect L2 reading comprehension. Accordingly, I touch on the philosophical concerns with proper names that are directly relevant to lexical analysis, specifically, whether proper names are part of the language system and whether they have meaning. Then, the contemporary linguistic debate surrounding proper names is reviewed in more detail. Drawing on this review, I move towards providing a working definition of 'proper name' as it is used in this research context. The section ends with some reasons why proper names should be given consideration in L2 research and pedagogy contexts.

2.2.1 Philosophical views on proper names

The philosophical debate on proper names is ongoing and concerns these issues: connotation and denotation; sense and reference; particular and general; truth; and pragmatics (J. M. Anderson, 2007). Only the first two of these concerns will be introduced here, as these are directly pertinent to the discussion of proper names as lexical or encyclopaedic knowledge. The issue of particular and general will be briefly touched on at the end of this section, though only as it relates to L2 reading. Van Langendonck (2007) remarks that philosophical interest in proper names comes from the "conviction that solving the problem of proper names is solving the problem of meaning and reference[3] (and vice versa)" (p. 22). An early and significant philosophical position on proper names comes from John Stuart Mill (1865) who argued that while proper names denote, or refer to[4], an individual, they do not connote; that is, names do not indicate any attributes about their referents[5] (p. 33). Mill (1865) used an analogy for how names are non-connotative: proper names are like the chalk marks put on houses by the robber in the Arabian Nights; while the mark serves a purpose to distinguish the houses, it does not have any meaning or say anything about the house (p. 36). Likewise, a proper name serves to distinguish but does not say anything about its referent. Thus, in the Millian perspective, proper names do not have meaning, as "meaning resides not in what [names given to things] *denote* but in what they *connote*" (Mill, 1865, p. 36).

[3] The term 'reference' here means the relationship between the proper name and the object that is named by it.

[4] The term 'refer to' means to connect to or link to.

[5] The term 'referent' means the person or thing to which a proper name refers.

As Lyons (1977) explains, Mill's use of the term 'connote' is philosophical, and is "intended to suggest that what he calls the signification of the attributes of a subject is something additional to the signification, or denotation, of all the subjects which possess these attributes" (p. 176). Lyons clarifies here that for Mill, connotation indicates the qualities of a thing, and it is in connotation where the meaning of the word resides. In contrast, denotation refers to a thing that has those qualities. Lyons emphasises that Mill's use of connotation is in contrast to denotation. In its non-philosophical usage, connotation has a related meaning (Lyons, 1977): the idea or feeling that a word invokes for a person. To sum up Mill's position, proper names can refer to (denote) entities, but they do not have meaning because they do not signify (connote) any features or qualities of those entities.

Gottlob Frege (1892 [1952]) provided another influential position on proper names. He introduced the idea that in addition to reference, names have *Sinn,* or sense, that is, associative meaning. For Frege, expressions could have the same *Bedeutung* (reference) but not the same *Sinn* (sense)[6] (Lyons, 1977). The example he used to illustrate this distinction was: *The Morning Star is the Evening Star.* These two names have the same reference (i.e. the planet Venus) but different sense. Lyons (1977) explains how some philosophers use the term 'sense', others would say 'meaning'. For Van Langendonck (2007), Frege's *Sinn* is not the same as lexical meaning but could be understood as 'meaningfulness' (p. 27). It is from these two fundamental philosophical positions that the debate on proper names continues. In the Millian view, names refer to entities, but do not signify qualities of those entities. Because names do not have meaning (connotation), they are not part of language (Strawson, 1950). Conversely, under Frege's conception, names do have sense, or associative meaning, and should therefore be considered part of language. Lyons (1977) summarises the widely accepted philosophical view that "proper names may have reference, but no sense, and that they cannot be used predicatively purely as names" (p. 219).

2.2.2 Linguistic views on proper names

The philosophical debate on whether proper names have sense, or meaningfulness, is directly connected to the central issue of proper names in lexical analysis: whether names are an aspect of vocabulary or encyclopaedic knowledge, and the

[6] Frege's use of 'reference' means the object that is indicated by a proper name; 'sense' is what the name expresses. See the example that follows in the text above.

interrelationship between these two types of knowledge (Hanks, 2013). Linguists who take the Millian view of proper names argue that they have no lexical meaning, only encyclopaedic information. For example, Coates (2006) argues that proper names have no sense and are referential in two modes: semantic and onymic. To illustrate these two modes, he gives an example of semantic reference: *I live at the old vicarage*, which means 'I live where a priest used to live'. In contrast, an example of onymic reference is: *I live at The Old Vicarage*, which means 'I live at a place called The Old Vicarage' (Coates, 2006, p. 368). In the two modes of referring, neither reference has sense for Coates.

However, as Van Langendonck (2007) points out, it is not clear why the reference in the semantic mode is a name: it seems to be a phrase because it is possible to insert words (e.g. *the old but beautiful vicarage*); conversely, no insertions are possible for the name *The Old Vicarage* (p. 67). From the perspective of an L2 reader, one might assume that in either referring mode, the word *vicarage/Vicarage* would be unknown and possibly looked up in a reference source. Furthermore, regardless of the presence of an initial capital letter, a dictionary would help an L2 reader understand both the semantic and onymic references. In this way, both *vicarage/Vicarage* can have meaning for an L2 reader.

That names are observed to not be part of the linguistic system might be attributed to their minimal sense and the predominance of encyclopaedic information associated with them (Anderson, 2007, p. 158). Though he disagrees with this conception, Anderson (2007) attempts to describe this view of proper names: the lexical entry of a name, without the phonological and morphological information, consists of a concept of a referent. This concept provides access to encyclopaedic information that is particular to that referent, and the name is simply part of that concept; the concept is not part of the linguistic system (Anderson, 2007, p.158). In other words, proper names can be thought of as "memorized labels" for the entities they refer to (Allerton, 1987, p. 71), or like the robber's chalk marks, serving only to distinguish.

Under the Millian perspective then, when one fails to understand a proper name, it is assumed to be the result of a lack of world knowledge, not linguistic. In this regard, proper names seem to differ from other words: for if one fails to know a common word, this is attributed to a gap in linguistic knowledge, not world knowledge. However, Anderson (2007) gives two examples why this might not always be the case:

> I nevertheless recognize as a language user that in English, for example, *Elise* is a name for women. And, on the other hand, we can also fail to grasp the denotation of common words on the basis of gaps in our knowledge of the world: for instance, I know that 'cantharides' is 'dried Spanish fly', but I would not be able to recognize a sample. Names are simply associated with much less sense, and possibly, in some cases, more encyclopaedic knowledge. (p. 158,159)

In Anderson's example of *Elise*, the gender of the name is linguistic knowledge of the word; encyclopaedic knowledge about *Elise,* on the other hand, would entail other specific information (e.g. her job or her character). Hence, just as understanding of common words can require world knowledge, there can be instances when knowledge of proper names can be of a linguistic nature, even though this may be minimal (i.e. in this example, gender and possibly nationality).

A related issue to whether names have meaning is whether they are translatable. According to Coates (2006), if names have no sense, then they cannot be translated. Lyons (1977) gives the example of *James*, which in France could be translated into *Jacques.* He asks whether in translating the name, the "very Englishness of his name" is lost, which seems essential to the name (p. 222). However, untranslatability is not limited to the class of names, as Anderson (2007) points out: certain loan words are also untranslatable (e.g. *chic*) (p. 158). Moreover, note that some people adopt equivalent names from the target language culture based on etymology (e.g. *John* for Greek name *Yanis*) (Anderson, 2007), or phonology (e.g. *Louis* for the Arabic name *Lowai)*. Anderson (2007) also provides the example of medieval scholars who, when writing in Latin, would Latinise their native language names, sometimes by playing with the etymology of the name. "Can there be any surer signs of belonging to a particular language?" he asks (Anderson, 2007, p.159). In this regard, it would seem that proper names can be treated as part of the linguistic system.

The alternative linguistic view of proper names then is that they do have meaning. For linguists who assign proper names sense, or associative meaning, proper names are part of the linguistic system. Note, however, that Frege's 'sense' is interpreted somewhat differently among linguists. As noted above, for Van Langendonck (2007), 'sense' is meaningfulness, and so "proper names are words just as others" (p. 67). In his view, names have associations or connotations (in the non-philosophical usage; that is, a feeling or idea that a word summons in a person, beyond its primary meaning)

that arise from either the name's referent or from the name's phonological shape (Van Langendonck, 2007, p. 82). Allerton (1987) glosses 'sense' as "language-internal semantic relations" (p. 71). For him, proper names do have meaning; they contribute to the meaning of the sentence in which they occur. However, the meaning is "an isolated, unintegrated one, such that it cannot be related to the meanings of other words in terms of lexical relations" (Allerton, 1987, p. 71). Thus, for Allerton (1987), while names do not lack in connotations, the meaning is "not integrated into the lexical and grammatical system of the language" (p. 81). In this way, proper names seem to exist both inside and outside of the lexicon (Allerton, 1987, p. 62).

Anderson (2007) maintains that naming and calling something by a common word or proper name are both social activities, for which initiation into a linguistic system is required (p. 157). Anderson (2007) says,

> What seems . . . important for an understanding of the use of names in language is that we take 'sense' to involve relations between semantically defined classes of words in the mental lexicon of language users (no matter how inadequate, and variable, are the actual definitions that users can offer). (p. 78)

Anderson goes on to observe that lexical and encyclopaedic knowledge are not necessarily distinct but that 'sense' is affected. He uses the example of *Bertie*, the sense of which is 'human, male'. Other things one might know about *Bertie (Wooster)* is not lexical knowledge but encyclopaedic. Coates (2006), however, takes yet a different view of 'sense'. For Coates (2006), names are senseless but he does not interpret this to mean that names are meaningless: the meaning is the referent (p. 365). That is, the meaning of a proper name is the object or person to which it refers.

Lyons (1977) asserts that names certainly do have connotations (in the non-philosophical sense) and associations. However, he states that "whether names belong to a language or not and whether they have a meaning or not do admit of a simple and universally valid answer" (p. 223). That is, the question of whether names are a part of the linguistic system and whether they have meaning as other words do, is a complex one to which the answer might not be the same in all languages. Van Langendonck (2007) says that to ask whether names have meaning is the wrong question:

> [The right question is] in what way the meanings are construed and function. We could see Frege's *Sinn* as meaning by itself. In common nouns, this meaning crucially determines the denotatum; in proper names, the meaning helps to retrieve an already given denotatum. For practical reasons, we can still call the meaning of common nouns 'conventional meaning', that of proper names 'associative meaning'. (p. 38)

The important issue for Van Langendonck is how proper names function in the text and how the reader makes sense of the proper names within the text; for him, it is clear that proper names have meaning, albeit not in the conventional sense as other common nouns. Lyons (1977) maintains the relationship between a proper name and its bearer is different from a common noun and its denotata; however, he allows that a connection exists, otherwise why would some thinkers confuse the two, just as a layperson refers to words as names for things: "The ordinary speaker of English, reflecting and reporting upon his language, is not similarly bound by the dictates of theoretical or ontological parsimony [as the philosophical semanticist]" (p. 216). This point is important for this research context, given the focus on L2 readers. Researchers and linguistics may analyse proper names in a very different way than the L2 reader. Lyons (1977) allows for the possibility that:

> In the learning of a language the distinction between proper names and common nouns may not always be clear-cut, so that there might be a time when 'chair', for example, is treated as a name which happens to be associated with several otherwise unrelated objects, and conversely, when all the people called 'Horace' are thought of as having one or more other properties by virtue of which the name 'Horace' is peculiarly appropriate. (p. 220)

This is an important possibility to consider in this research context, given its focus on L2 readers' processing of proper names. Thus, an allowance might be made that L2 readers analyse proper names differently than L1 readers and researchers.

If L2 readers do analyse proper names differently, what sort of 'meaning' might they get from proper names that would aid reading comprehension? Van Langendonck (2007) suggests several different types of "presuppositional meaning" that proper names can have: categorical (e.g. man, woman, country, city, month, etc.); associations (about the referent or from the word form) and connotations; emotive (interpreted personally or inherent in the name); and grammatical (gender, number,

definiteness) (p. 86). Neurolinguistic support for the notion that proper names have categorical meaning is found in Van Langendonck's (2007) reporting of German research done by Joseph Bayer on a patient with deep dyslexia (pp. 110-113). The patient could only access written text through the semantic route (i.e. processing via the lexicon), not the phonological route (i.e. written word to speech via phonemic processing). That is, when shown proper names, she could not read them but she could recognise them as names and provide categorical information about them. For example, for personal names, she could usually specify whether the name denoted a man or woman. She identified place names as cities, countries or rivers. Bayer concluded that there must be "a minimal lexical categorical sense belonging to the semantic memory" (Van Langendonck & Van de Velde, 2016, p. 16). The patient could also provide connotative information about names. For example, *Australia* triggered the categorical information of 'country' but also connotations of 'far away' and 'kangaroos' (Van Langendonck, 2007).

Accordingly, if proper names have categorical and associative meaning, they can be treated as lexical items. Hanks (2013) remarks that from "the point of view of corpus linguistics and computational linguists, [names] certainly are [words]" (p. 33). Shcherba (1940 [1995]), in his theory of lexicography, also asserts that names must have meaning because they are used in speech (p. 323). For him, names are words, albeit very different from common nouns, but that is not a reason to exclude them from the dictionary. The question is what meaning to assign names. Meaning is not of an encyclopaedic nature. Rather, Shcherba (1940 [1995]) says,

> The task is to define that necessary minimum without which it would be impossible to operate in a generally understandable way with a proper noun in speech . . . this minimum is the concept which subsumes a given subject, with the general indication that it is not any subject subsumed by a given concept, but a specific one. (p. 323)

Shcherba suggests that there is necessary (minimal) information about a proper name that one must understand in order to distinguish it from other subjects. For example, if the proper name is a country name, then the fact that it is a country and not a person is the necessary information that one would need to understand to make sense of the proper name in a text. This "necessary minimum" connects to Van Langendonck's categorical meaning. Shcherba (1940 [1995]) offers some examples of possible dictionary entries: "*Australia*, 'one of the countries of the world'; *Louis XIV,* 'one of the

French kings'" (p. 324). Furthermore, Shcherba (1940 [1995]) notes that not all names would need to be included in a dictionary, only those commonly known to a linguistic group (p. 324). Certainly, as Lyons (1977) remarks, to say that proper names have been excluded from some dictionaries because they do not have meanings is too simple an argument for not considering them a part of the language system (p. 222).

Whether names refer uniquely is another debated topic in both philosophy and linguistics. I will not go deeply into this debate here, as its importance to L2 reading comprehension seems limited. Some linguists like Coates (2006) argue that names clearly do not refer to unique persons, in that many people may have the name *John*, for example. In response, J. M. Anderson (2007) notes that people behave as if names refer to unique entities, even if two or more referents exist in their lexicon for the same name; and if name has failed to be understood as referring uniquely, then it is the act of speech that has failed (p. 117). There is empirical evidence to support this claim. Hall (1996) conducted studies with L1 preschool children and found that they interpreted a novel word as a proper name when it applied to only one object rather than two (e.g. *This cat is DAXY)*. When syntactic information overrode that interpretation, the children revised their interpretation of the novel word as an adjective. His findings support the notion that children view proper names as referring to unique individuals. Similarly, I suggest that L2 readers understand that a name in a text refers to a unique (real or fictional) individual while the possibility also exists that the same name might be used for another person. That said, context has an important role in differentiating referents with the same name.

2.2.3 Defining proper names

As noted above, a simple dictionary definition of 'proper name' is a name that refers to a unique person, place or organisation, and in English, is capitalised ("Proper noun," n.d.). An important distinction in the role of proper names is that they can be used to 'speak of' a person (as an argument) or to 'speak to' a person (as a vocative) (J. M. Anderson, 2007; Lyons, 1977). However, linguists have differences in opinion as to how to define proper names. Specifically, there is disagreement as to the syntactic categorisation of proper names. Also, linguists differ as to which items should be classified as proper names. Orthography is not helpful here because it is not always a reliable clue to a word's grammatical class; there are various irregularities in the English writing system. Orthography is an important factor in this research context

because the participants are processing a different L2 writing system from their L1. This section discusses various viewpoints related to proper name grammatical categorisation and orthography. Through this discussion, I move towards a working definition of 'proper name' as it will be applied in this L2 reading research context.

Proper names have traditionally been regarded as a subclass of nouns in European grammatical studies (Lyons, 1977; Van Langendonck, 2007). However, Anderson (2007) points out that the Stoics did not view proper names as nouns, but distinguished between names and nouns as distinct classes, translated into Latin, respectively, as *proprium* and *commune* (p. 132 and 172). Anderson (2007) is one notable exception to the European tradition, arguing 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). Perhaps because of the convention of classifying proper names as nouns, proper names are also sometimes referred to as proper nouns.

The grammarian Huddleston makes a distinction between the terms 'proper noun' and 'proper name'. For him, proper nouns are items like *London* and *Queensland*; in contrast, *the University of Queensland* is a proper name because it is composed of a proper noun and a common word element (Huddleston, 1984, pp. 229-230). As Anderson (2007) points out, in this conception, Huddleston allows for complex items composed of common word elements as proper nouns (e.g. *Queensland).* However, he does not allow for compound items (e.g. *the University of Queensland)*, distinguishing such items as proper names. Anderson (2007) sees no need for the distinction – both types are names regardless of what elements they are composed of; whether complex or compound, the items still refer to an entity. Furthermore, proper names based entirely on common word items are still names (e.g. *Long Island*). The only purpose Huddleston's distinction between proper name and proper noun would serve is to demonstrate prototypicality (J. M. Anderson, 2007, p. 191).

In this thesis, the term 'proper name' will be used throughout to include names regardless of whether they are comprised of common word elements or not (*pace* Huddleston). I suggest that an L2 reader would not make distinctions among proper names on this basis. Also, by using the term 'proper name' instead of 'proper noun', the alternate view of the grammatical classification of proper names is acknowledged,

that they might not be accurately categorised as a subclass of nouns. In instances where an original author has used the term 'proper noun', that term shall be retained.

There is divergence among linguists regarding which items should be considered proper names, and it seems this divergence is partly due to orthographic irregularities in English. For example, Allerton (1987) notes that temporal nouns (e.g. *Monday, February*) are not strictly proper names in that they do not refer to unique entities. He notes that some languages, like French, do not capitalise temporal nouns. Allerton (1987) concludes, however, that it is "worth" recognising temporal nouns as proper names (p. 79). Anderson (2007) agrees that calendrical terms are proper names, albeit not prototypical. Van Langendonck (2007) views these terms as appellative (i.e. vocative) or proprial (i.e. proper name related). However, he concludes they could be treated as proper names because they can appear as appositions (e.g. *the month of June).* (Note that proper names can appear as appositions, for example, *my best friend John*). In this thesis, calendrical terms will be treated as proper names, under the assumption that the L2 reader would likely analyse them as names due to orthographic rules of English (i.e. initial capital letter on names).

Related to this issue of how initial capital letters are used in English, Allerton (1987) observes that initial capital letters in English are not reserved only for proper names. Names of languages (e.g. *Japanese*) and the related nationality adjectives (e.g. *Japanese*) are written with a capital letter; in some languages, like French and Spanish, such items are not capitalised. Allerton (1987) suggests this is done in English because these items are derived from a proper name (in this example, the country name *Japan*). To account for such lexical items, Allerton (1987) argues for the creation of the category 'proper name related' to include all "adjectives and common nouns of geographical provenance and common nouns related to social organizations" (p. 77, 78).

Anderson (2007), however, takes a slightly different view. For him, names of languages are indeed proper names, while nouns of nationalities are not: they do not behave syntactically like proper names because they can be predicative (e.g. *Bill is an American*) (p. 129). In this way, the initial capital letter on nationalities is misleading. Van Langendonck (2007) takes yet a different position, viewing names of languages as "a kind of proprio-appellative lemmas that behave primarily as proper names but that are construed as common nouns very frequently" (p. 243). His point is that names of languages can be used as both proper names and as common nouns. In addition,

Strawson (1950) discusses what he calls “quasi-names”, nouns that grow capital letters (e.g. *the Great War*) (p. 341). These are a few examples of how the initial capital letter in English is not a reliable clue for grammatical class.

From the perspective of the L2 reader, it is difficult to say how the orthography would be analysed in such cases. Take, for example, these two sentences:

(1) *He is studying Japanese.*
(2) *He is Japanese.*

Because learners of English are taught that proper names have initial capital letters and are rarely proceeded by the definite article, it is hard to say whether an L2 reader would recognise that sentence (1) contains a proper name while sentence (2) contains an adjective (in Anderson’s (2007) conceptualisation). Likewise, how an L2 reader would analyse quasi-names like *the Great War* is speculative. I would argue that it is unlikely that an L2 reader would examine and differentiate names based on criteria of prototypicality because of their knowledge of English orthographic rules for proper names. Also, recall Lyons’ (1977) acknowledgement that the layperson is not bound by linguistic theory in her analysis of the language. In a similar way, an L2 reader might analyse words with initial capital letters as proper names without much regard to their prototypicality.

A working definition of ‘proper name’ for L2 reading

For this working definition of ‘proper name’ as the term will be used in the thesis, I draw on Van Langendonck’s (2007) typology to include these subclasses of proper names: prototypical names (personal, animal and place names); appellative proper names (brand names, languages, colours, diseases); and autonyms (common words used as names, e.g. *bank*) (p. 184). Van Langendonck (2007) also includes a fourth category, nouns with restricted proprial function (e.g. *gold* in *the element gold*). This last one will not be included because it is unlikely an L2 reader would analyse such an item as a proper name because of the absence of an initial capital letter. Proper names in this thesis will also include calendrical names (e.g. *Monday, February*) and other non-prototypical, ‘quasi-names’ (e.g. *the Pope*). Because L2 readers get much L2 input from both news sources and works of fiction, proper names shall include names with real world and fictitious referents.

2.2.4 The need for proper name research in L2 contexts

Proper names have been largely neglected in second language acquisition and vocabulary research (Kobeleva, 2012). Yet Hanks (2013) offers a startling statistic why vocabulary researchers should be concerned with proper names: because modern electronic lexicons have no space limits, some lexical databases are demonstrating that over 70% of lexical entries are proper names and this number continues to grow. "In other words, the number of proper names is both larger than all the rest of the lexicon and growing faster than any other area of the lexicon. The nearest rival is multiword expressions" (Hanks, 2013, p. 36). When proper names make up such a large percentage of a language, this has implications for teachers and researchers of L2 reading and vocabulary. This section presents some reasons why proper names deserve more attention from both researchers and teachers.

If L2 vocabulary researchers have overlooked proper names because they are considered encyclopaedic knowledge, there is some research from other fields that suggests L2 users might analyse proper names as lexical items. Chapter 3 considers how L2 reader view proper names; Chapter 4 discusses in detail how proper names have been handled in L2 vocabulary research. Here, I briefly mention two observations from code switching research. For example, Bultena, Dijkstra, and Van Hell (2015) report that based on studies with habitual code switchers, the switches "seemed to co-occur with lexical overlap between languages", such as cognates and proper names (p. 456). They give an example of the L1 Dutch speaker who begins a sentence in Dutch, and after uttering an English proper name, completes the sentence in English. The name is associated with the L2, and triggers L2 use. That a proper name can trigger code switching suggests that names might exist in the bilingual lexicon.

In another code switching study, Park (2006) reports on proper name usage in Korean-Swedish bilingual data. Park notes that proper names have been traditionally excluded from analysis in code switching analysis because they are assumed to be borrowings (i.e. single words borrowed from the other language). Borrowing is distinct from code switching, which involves multiword sequences retaining morphosyntactic structure from the original language. In the data analysis, Park (2006) found that "more than half of the foreign proper names . . . show full morphosyntactic integration into the language of the sentence" (p. 33). For this reason, Park (2006) suggests that in the bilingual

lexicon, proper names might be like other lexical items requiring morphological and syntactic integration. These observations illustrate why more research is needed into proper names and the L2 user.

Likewise, it is important that teachers give adequate attention to proper names. As noted earlier, it is possible L2 readers analyse proper names differently than L1 readers. It is very likely that L2 readers will encounter unfamiliar names by the very nature of what they do: process a second language and its associated culture. Especially in an EFL context, it can be reasonably expected that L2 readers will encounter unfamiliar names. Examples of contexts in which L2 readers might encounter names with opaque referencing include flouting of names, when a name is used as a generic reference. In Hanks' (2013) terms, this is when the language has been exploited, and the usage can sometimes become the norm. For an L2 reader in an EFL context, it is reasonable to assume that flouting of names can be puzzling. Take the example in (3) of a proper name reference that could cause confusion for an EFL reader.

(3) . . . while I'm no Fred Astaire, I'm a damn fine wedding dancer . . . (Pang, 2010, p. 1323)

This sort of referencing can be problematic even for L1 readers when the proper name used as a frame of reference is outdated or esoteric (McPhee, 2015, March 9). As Hanks (2013) reminds us, names are ephemeral. And the exploiting of names as generic references occurs frequently in literature and newspapers, arguably two important sources of input for the intermediate or advanced L2 reader.

Transliteration of names, that is, the conversion of the name's graphemes from one writing system into those of another writing system (Coulmas, 2003), is also an important issue for L2 reading. Languages have different conventions for rendering names from other writing systems, and this could be a potential source of misunderstanding for the L2 reader. To take an example from Hanks (2013), the Russian name *Ельцин* is transcribed as *Eltsine* in French, *Jelzin* in German, and *Yeltsin* in English (p. 36). While context will aid comprehension, it may not always be clear to the L2 reader who the referent is. To take another example from Chinese, in the past, transliteration was done using the Wade-Giles system and now the Pinyin system is used. For readers of a certain age, it may not be obvious that *Peking* is the same place as *Beijing*.

The challenges that natural language computer programs have with proper names are illustrative of difficulties applicable to the L2 reading context. Take example (4) from Hanks (2013):

(4) Birmingham beat Coventry City. (p. 38)

This is an example of metonymy, where the subject is not directly named but referred to by something associated with it. In this example, the Birmingham football team is referred to only by the city name. While most British speakers would understand that what is being reported is the results of a football game, computer language processors are challenged: no mention is made to the sport or to the teams; only two place names are given within the context of the seemingly incongruous verb 'beat'. As Hanks (2013) remarks, the example illustrates the central point about proper names in lexical analysis: "the interrelationship between knowledge of a language and knowledge of the world" (p. 38). For an L2 reader, relevant background knowledge might aid comprehension if one assumes there is an interest in the topic. Unfortunately, for L2 readers, in particular those in a formal educational or testing situation, there is not always a choice or interest in what is read. In that case, background knowledge cannot be assumed to aid comprehension.

To summarise this section on proper names, the debate surrounding proper names was explored, including arguments why proper names might be considered part of the L2 reader's lexical knowledge. This theoretical question of whether proper names represent encyclopaedic or lexical knowledge underlies the empirical studies presented in Chapters 3 to 6. A working definition of proper name was provided; this definition is used as a guide in the experimental work. Several ways that proper names can disrupt L2 reading were briefly mentioned, providing reasons for more research into proper names in L2 contexts.

2.3 L2 vocabulary and reading comprehension

Because the central research aim is to investigate the vocabulary load of proper names on reading comprehension, it is important to clarify what other key factors that may impact comprehension. Two significant factors that can influence comprehension include background knowledge and the percentage of known vocabulary. Several studies have examined the relationship between the percentage of words known and comprehension. For example, Laufer (1989) concluded 95% vocabulary coverage was needed for minimal (55%) reading comprehension. Hu and Nation (2000) suggested 98% coverage was ideal, allowing for better comprehension. Schmitt et al. (2011) addressed some of the limitations in these previous studies and others, including how vocabulary knowledge of participants was measured (more on this below) and the sample size. The Schmitt et al. (2011) study included 661 participants from eight countries with 12 different L1s. In addition to investigating the relationship between known vocabulary and reading comprehension, the authors also examined the effect of background knowledge, including relevant social and cultural knowledge, which has been shown to have a large effect on reading comprehension (Schmitt et al., 2011, p. 30).

Schmitt et al. (2011) identified a relatively linear relationship between vocabulary coverage (between the 90% and 100% levels) and reading comprehension. Based on their findings, the authors suggest that if 60% comprehension is desired, then 95% vocabulary coverage is needed; 70% comprehension can be achieved at 98-99% coverage; and 75% comprehension at 100% coverage. Thus, they recommend that 98% vocabulary coverage is a reasonable target. To test the effect of background knowledge, Schmitt et al. (2011) chose two texts of similar length and difficulty: 'Climate' concerned global warming, a topic that all the participants would be quite familiar with; 'Mice' concerned a scientific study on the effects of mental acuity and exercise done on laboratory mice, a topic that probably few of the participants would have prior knowledge of. At the 90% to 93% vocabulary coverage levels, there was no advantage seen in comprehension with the familiar topic. But at the 94% - 100% levels, comprehension scores were better for the familiar topic, ranging from 7.8 to 13.9 percentage points, suggesting that background knowledge does affect reading comprehension, in addition to vocabulary knowledge. A dip, however, was seen at the 95% level.

As noted above, one of the strengths of the Schmitt et al. (2011) study is how the participants' vocabulary and comprehension was tested. Since the aim was to investigate the relationship between vocabulary coverage and reading comprehension, the participants were directly tested on their knowledge of a high percentage of words from the two texts. To this end, the authors profiled the two texts, identifying words which occurred in both texts and which were unique to one or the other. They sampled more heavily from the 2,000 band and above, assuming these high level participants would know the first 1,000 band quite well. A total of 120 target items were chosen; 30 non-words were added to the list, as the authors had decided to use a yes/no format, given the different L1s of the participants. In this way, the participants were directly tested on their vocabulary knowledge. The authors were also thorough in the design of the reading comprehension test. They opted for multiple-choice questions, avoiding any that directly tested vocabulary items. However, there can be problems using multiple-choice questions to test comprehension; for example, distractors can alter information in the reader's mind. Alderson (2000) argues that distractors present possibilities that may not have occurred to participants, and thus amounts to "a deliberate tricking" of students (p. 211). To address these possible issues, the authors also included a graphic organiser task, which involves greater cognitive processing to recognise text organisation. Graphic organisers are also known as "information transfer tasks". Such tasks may impact cognitive load in that participants need to identify logical relationships between the information provided on the task sheet and the information sought for the blanks. In sum, the vocabulary and comprehension measures used in this study seem to address shortcomings in previous studies. The sample size (N = 661), drawing on participants with 12 different L1s, also added to the study's robustness.

Given these strengths, Schmitt et al.'s (2011) conclusion that beyond 90% coverage, there is no 'threshold' of vocabulary knowledge needed for comprehension but rather that a linear relationship exists between the two, is a convincing one. The authors note that the average 2.3% increase in comprehension per 1% in vocabulary growth they found is similar to findings in Hu and Nation (2000) as well. The conclusion that while vocabulary knowledge is important, it is still only one aspect of comprehension is also borne out by the study's results: after the 94% vocabulary coverage level, the more familiar text topic led to higher comprehension results, suggesting background knowledge also plays an important role. Thus, this robust study clearly demonstrates the importance of vocabulary size and background knowledge to reading comprehension. The study did not mention, however, if either text contained proper

names. Therefore, how the L2 reader's knowledge of proper names, either lexical or encyclopaedic, impacts this equation of vocabulary size, background knowledge and reading comprehension, is explored further in the empirical studies in the following chapters.

2.4 Reading processes and writing systems

To ensure clarity, certain terminology related to reading processes and writing systems that appear in the subsequent chapters are defined here. I begin with reading processes. To convey the complexity involved in reading, Grabe (2009) defines reading as a complex combination of ten processes: rapid, efficient, comprehending, interactive, strategic, flexible, purposeful, evaluative, learning and linguistic processes (p. 14). The overall goal of reading is general comprehension: to obtain a good understanding of the main and supporting ideas, but not necessarily to remember all the specific details, and to be able to relate the main ideas to one's background knowledge as necessary (Grabe & Stoller, 2011). Sadoski and Paivio (2007) describe three aspects or subdivisions of reading: decoding, that is, converting printed text into spoken language; comprehension, that is, making meaningful interpretation of the text at literal, inferential, and critical levels; and reader response, which overlaps with critical interpretation but also can include reader appreciation and application, and can occur during or after reading (p. 341). In the decoding aspect, input is derived mostly from the text. In comprehension, input is balanced between the text and the reader. During response, input derives mostly from the reader's reaction to the text (Sadoski & Paivio, 2007). The three aspects do not operate separately but in an interactive manner.

It is necessary to distinguish between lower-level processes, which are text-based knowledge sources, and higher-level processes, which are reader-based (Nassaji, 2014). Lower-level processes include word and letter recognition, and sub-skills of phonological, orthographic, semantic and syntactic processing (Grabe, 2009). Higher-level processes involve connecting text-based information to the reader's prior knowledge (Nassaji, 2014). These processes include reader-directed attention to the text, such as background knowledge, strategies, goals, inferencing and comprehension monitoring (Grabe, 2009). In most current reading models, researchers view reading as information processing (Nassaji, 2014). In this 'reading as information-processing' perspective, lower- and higher-level processes are integrated and hierarchical, but not

reciprocal. That is, efficient lower-level processing is critical for higher-level processing to occur; however, higher-level processing cannot compensate for inefficient lower-level processing skills (Nassaji, 2014, p. 3-4).

There are many models of reading. As was noted in Chapter 1 (section 1.1), there are metaphorical models of reading that present generalisations about reading comprehension: bottom-up, top-down and interactive theories of reading (Grabe, 2009; Hudson; 2007). Bottom-up theories emphasise the importance of lower-level processing to reading comprehension while top-down models emphasise the role of the reader who samples information from the text to generate meaning. Interactive models postulate that lower- and higher-level processes interact; for example, the reader's background knowledge might serve as support for word recognition. However, as noted above, this view of reciprocal processing in reading is not supported by research (Nassaji, 2014). Rather, restrictive interactive models of reading (e.g. Verbal Efficiency Model, Perfetti, 1985) predict that there is little interaction between processing levels; any interaction that occurs will be within a component skill. For example, word recognition requires interaction between the sub-skills of phonological, orthographic and semantic processing. However, higher-level processes will not interact with lower-level processes for comprehension. In sum, Grabe (2009) notes that metaphorical models of reading are not accurate reflections of current views on reading (pp. 89, 90).

A few influential reading models were briefly noted in Chapter 1 (section 1.1), including: the Simple View of Reading (Gough, 1972), which predicts efficient decoding skills and language comprehension are key to reading comprehension; and the Psycholinguistic Guessing Game model (Goodman, 1967), a top-down model which has largely fallen out of favour as its predictions have been proven incorrect through extensive empirical research. In addition to these, Grabe and Stoller (2011) note three other influential reading models: the Interactive Compensatory Model (Stanovich, 2000); the Word Recognition Model (Seidenberg & McClelland, 1989); and the Dual-Coding Model (Sadoski & Paivio, 2007) (pp. 27, 28). The Interactive Compensatory Model predicts that certain reading processes will interact to compensate for those processes that are not fully automatic. The Word Recognition Model describes how the sub-skills of phonological, semantic and orthographic processing interact for word recognition; the model predicts that word recognition is key to successful reading comprehension. Finally, the Dual-Coding Model emphasises the importance of both visual and verbal input to reading. The importance of lower-level processes to reading comprehension,

with reference to some of these theories of reading, will be taken up further in Chapter 6.

Terms related to writing systems are also important in this investigation of proper names and L2 reading. The 'writing system' of a language refers to the ways in which written symbols connect to a language, as well as the rules for writing in a given language (Cook & Bassetti, 2005). There are three major writing systems: alphabetic, in which symbols represent a phoneme; syllabic, in which symbols represent a syllable; and logographic, in which a character corresponds to the meaning and sound of a morpheme (Koda, 2013). Examples of these three writing systems include: Korean Hangul (alphabet), Japanese Kana (syllabary), and Chinese characters (logography). 'Script' refers to the graphic implementation of the writing system. For example, English has an alphabetic writing system and uses the Roman script; Bulgarian also has an alphabetic writing system but uses the Cyrillic script. 'Orthography' refers to the rules for using a particular script. For example, both English and German use the alphabetic writing system and the Roman script. However, these two languages differ in orthography: in English, proper names are capitalised, whereas in German, all nouns are capitalised.

Koda (2013) points out that L2 reading is a more complicated process than L1 reading because it is cross-linguistic (i.e. two languages are involved at each sub-skill level). Because writing systems differ in the basic unit of orthography (i.e. phoneme in alphabet; syllable in syllabary and morpheme in logography), different processing is assumed for each system; therefore, L1 processing transfer to L2 reading is expected (Koda, 1994). Indeed, there are studies which demonstrate L1 orthographic processing strategies applied to L2 reading (Hudson, 2007). Furthermore, as Alderson (2000) points out, since letter recognition is important for efficient and rapid word identification (lower-level processing), L2 readers of different scripts or orthographies might experience greater processing difficulty (p. 75). Koda (1996) suggests that an L2 reading research framework take into account L2 processing experience, L1-L2 orthographic distance, and the interaction between L1 and L2 orthographic knowledge. Studies[7] that have investigated effects of L1 orthography on L2 reading and L1-L2 orthographic distance are discussed in more detail in Chapter 6. Here, it is sufficient to acknowledge the dual-language role in L2 reading (Koda, 2012), which makes it a more complicated process than L1 reading.

[7] See for example Akamatsu (2003), (Bassetti, 2008; Bassetti & Atkinson, 2015; Chikamatsu, 1996; M. Wang, Koda, & Perfetti, 2003).

Chapter summary

This chapter has explored the term 'proper name' from philosophical and linguistic perspectives, thus illustrating the complex issues surrounding proper names. The most important of these issues for this thesis is whether proper names have meaning and whether they are part of the language system. I suggested that for the L2 reader, proper names do contribute meaning and thus could be analysed as lexical knowledge, as opposed to world knowledge. A working definition of 'proper name' was provided, as it will be used in this thesis: prototypical names (e.g. personal, place); proper name related (e.g. brand names, languages); calendrical terms (e.g. days of the week); quasi-names (e.g. titles); and both real and fictitious names. Then, a robust study was reviewed to show how the L2 reader's vocabulary and background knowledge can affect reading comprehension: generally, 98% vocabulary knowledge of a text will allow for 70% comprehension. Familiarity with a topic also ensures greater comprehension, though the effect of background knowledge will only be seen if at least 95% of vocabulary is known. Finally, terms related to reading processes (e.g. lower-level and higher-level processes) and writing systems (e.g. orthography, logography) were defined, as these are used throughout the thesis.

The next chapter presents an exploratory study that served as a springboard for the empirical studies that follow in Chapters 4 to 6. The exploratory study consists of interviews conducted with a small sample of L2 readers. The aim is learn more about: the affective factors involved in proper name processing; what strategies are used for unfamiliar proper names; and what sort of processing difficulties L2 readers might have with proper names.

Chapter 3: L2 readers' perspectives on proper names

3.1 Introduction

Anecdotal evidence was given in the introductory chapter from my own classroom experience, observations that run contrary to an assumption that L2 readers can easily recognise and understand proper names from the form (i.e. initial capital letter) and the context. The literature surrounding this assumption is reviewed in detail in the next chapter. Before looking at how L2 vocabulary researchers have approached proper names, this chapter explores L2 readers' perspectives on proper names. Contrasting views on proper names from philosophers and linguists were discussed in Chapter 2. A pertinent aspect of that debate for this thesis is whether proper names represent lexical or encyclopaedic knowledge. I suggest that from the perspective of the L2 reader, proper names might be analysed as lexical items. It may be that L2 readers peruse proper names just as they do other L2 lexis to make meaning of a text. As processors of a second language, L2 readers will almost certainly encounter unfamiliar proper names. It is also likely that not all referents of proper names are inferable from context.

That being the case, an initial consideration for this investigation concerns how L2 readers regard proper names. Specifically, this chapter aims to determine what affective factors (i.e. emotional factors that can influence learning) are involved in processing unfamiliar proper names. For example, if L2 readers are indifferent to unknown names, then it might be reasonable to presume that names present a low reading burden. On the other hand, if unfamiliar names cause anxiety for L2 readers, then this finding would warrant further investigation into what makes certain names problematic, for example. Also of interest is whether L2 readers use any strategies for handling unknown names. If it is the case that L2 readers can effectively deal with unfamiliar proper names with specific strategies, such as inferring from context, then it might be reasonable to assume proper names cause little strain for L2 readers. Lastly, it is worth looking at whether L2 readers have any particular difficulties related to decoding of proper names, difficulties that might be distinct from processing of other lexical items.

3.2 Study: L2 readers' perspectives on proper names

A qualitative study is presented here that explores L2 readers' perspectives of proper names. The aim of the study is to learn more about: affective factors related to L2 proper name processing; strategy use for handling unknown proper names; and any particular processing difficulties specific to proper names. In these aims, the study is exploratory in nature, and can be seen as a springboard to further the central research aim of determining whether proper names are a processing burden for L2 readers. To this end, interviews are conducted with Japanese L2 English participants. Interviewing is a "known communication routine" which makes it so useful as a research instrument, and it is the method used most often in qualitative research (Dornyei, 2007, p. 134). A semi-structured interview format is used: a combination of prepared questions targeting specific aspects of proper name processing, and opportunities for interviewees to raise and elaborate on factors relevant to their own experience which the researcher may not have anticipated. This format should maximise the potential to learn more about how L2 readers perceive proper names.

In addition to interviews, a read-aloud task is used in order to investigate how well L2 readers are able to decode and infer information about unfamiliar proper names. The read-aloud task is an adaptation of think-aloud protocol. In a think-aloud task, participants are asked to focus on a task and articulate, without explanation, any thoughts that occur. In the read-aloud task used here, participants are asked to read a short passage aloud, and when they encounter a proper name, to verbalise what that name refers to. This read-aloud task differs from think-aloud protocol in that in the read-aloud task, participants are not asked to verbalise all thoughts that occur to them while reading the text, but rather, to identify proper names and say who or what they refer to. In this way, insight might be gained as to participants' ability to identify, decode and understand proper name referents.

Theoretical motivation for the study was drawn from the philosophical and linguistic debate on proper names discussed in Chapter 2. That is, proper names might be analysed by the L2 reader as lexical items with some minimal, categorical meaning; in that case, how does the L2 reader approach unknown names in a text? Alternatively, it may be that proper names are analysed as referring expressions representative of world knowledge. In that case, how does the L2 reader handle or feel about unfamiliar names? The research questions are:

1. How do Japanese L2 readers of English feel when meeting unfamiliar proper names in texts?
2. What strategies do they use when encountering unfamiliar names?
3. How well are they able to decode and comprehend unfamiliar names?

3.2.1 Participants

I recruited[8] four students from two English classes I was teaching at a private university in Japan. The four students (two women and two men) were all in their first year of university (18 or 19 years old). They had studied English for about six years, though one participant had studied for ten years. Their English language proficiency was intermediate (A2/B1 on the Common European Framework of Reference, or CEFR). They were English majors, studying in a program that included eight 90-minute classes per week of academic English with L1 English teachers. A non-obligatory goal of the program is that the students will spend their third year abroad in an English-speaking country at one of the university's partner schools. Because of this goal, the students were typically motivated to do well in their courses.

My sampling strategy in recruiting participants was purposive: the sampling was driven by theoretical objectives, typical of qualitative research (Dornyei, 2007, p. 126). Two criteria were established. Of primary interest were those students who had previously remarked on having difficulties with proper names while reading. That is, during the academic year, there were occasions when students had commented on, either in tutorials or in writing, difficulties with proper names. I had kept a record of any comments made by students related to proper names (in the interest of research). When recruiting for this interview study, I checked this record for current students who had noted problems with proper names while reading. A secondary criterion for recruitment was reasonable confidence in speaking English. Japanese university students are typically hesitant to speak English (King, 2013). For purposes of an interview, interviewees' hesitancy to speak could have a negative effect on the quantity or quality of data collected. Therefore, reasonable confidence in speaking English was another consideration in the sampling strategy. Below I will describe each of the participants recruited for the interview and how well they met the selection criteria I had

[8] The University's ethical procedures were followed and ethical approval was obtained. Further details are provided in section 3.2.3.

set out. In accordance with identity protection regulations at the Japanese university, I will not refer to the students by their names, only by the first initial.

Student T had come to me on one occasion for tutoring with academic citations: he asked how one could decide which names were surnames or personal names, and in which order they appear in a citation. In the classroom, T was unlike the majority of his peers: he was always very eager to communicate in English and was not afraid of asking questions or offering comments. Due to his confidence in communicating in English and his questions about English names in academic citations, he was considered a good interview participant.

Student K had alluded to some difficulties with identifying and understanding proper names during independent extensive reading. He noted on his vocabulary-to-learn lists several proper names as words that he needed to check. For example, when reading *The Secret Garden*, he had written “n. robin – person’s name?” These sorts of notes indicated that he had some difficulty in deciding which lexical items were names, for example, when the context had not provided enough information. K was also an unusual student in that he was very keen to speak English and did not seem embarrassed when he made mistakes. When he was misunderstood, he would always try to rephrase to clarify. In this way, he met the criteria set out for the interview.

Student W made reference to difficulties with English names in a written assignment that was meant to probe students’ attitudes towards to reading. She wrote in her reflective essay: “I also can’t memorise the name of foreign person so when I read the book which has many characters, I often can’t find who did something or whom story is written or who did the conversation.” W did not meet the second criterion: she was a very reserved, quiet student in the classroom. I thought she might become more communicative in an interview setting because her peers would not be in the room. Because she met the first criterion of having remarked on some specific difficulties with proper names, I decided to invite her for the interview.

Student A made written comments related to struggles with names in a timed (fluency) writing task, for which the topic was “Learning to read in English”. In her response, A wrote that she had problems when a story had many different characters because it was hard to remember and follow all the various names. A was also a talkative, active participant in the classroom. In these two ways, she met the criteria outlined for the recruitment of interviewees.

To conclude on the sampling of these four participants, they were selected from intact classes, which were nearly homogeneous in terms of language proficiency. In this regard, the interviewees were considered typical of their classmates in their reading experience and ability. As noted above, three of the interviewees were less typical in that they were self-assured speakers of English. This was important, given the interview format. While I assumed that more students than just these four were having difficulties with proper names while reading, I wanted to lessen my bias as the researcher and not project this assumption onto potential participants. Therefore, rather than selecting students who were suspected of having difficulties with proper names, only those students who had independently noted the difficulty were chosen. It was predicted that with these four interviewees, the data would be insightful and informative with respect to the research questions. Table 3.1 summarises information about interviewees.

Table 3.1

Interview participant details

Participant	Gender	Years of English study	Proper name difficulty	Oral English confidence
T	male	6	citations	high
K	male	6	extensive reading	high
W	female	10	extensive reading	low
A	female	6	extensive reading	high

I asked for their participation to interview via email. In receiving the invite to interview in an email, students would have time to formulate a negative response if they were not inclined. Also, by not asking the students in the classroom, they were not singled out in front of their classmates. (See Appendix 1.1 for a copy of the invitation to interview). All four students responded favourably to the interview request, and we scheduled individual appointments for the interviews.

3.2.2 Instruments

Interview schedule

In order to direct the interviews and ensure that the interviews were consistent across participants, an interview schedule was prepared. As noted above, a semi-structured format was chosen: the questions on the interview schedule were meant to serve as a guide for the researcher to elicit information about the learners' attitudes about and strategies for handling proper names in reading texts. Below, each question is listed, followed by an explanation of the purpose of the question.

What is your first language? How long have you been studying English?

These first two questions were meant to be easy-to-answer questions to put the interviewees at ease (Dornyei, 2007, p. 137). How long the participants had been studying English was also important to establish as an indication of their L2 orthographic processing experience (Koda, 1996).

When you read something in Japanese, like a news article, does it ever happen to you that you see a name in Kanji that you don't know? What do you do when that happens?

This set of questions was meant to enquire into participants' strategies for handling unknown L1 proper names, in order to get them thinking about proper names in a familiar context. Japanese names are most often written in Kanji (Chinese characters). It is possible that a Japanese reader may come across an unknown L1 proper name, and not know how to pronounce it or know who or what it refers to. This is a common problem with proper names for readers of meaning-based orthographies like Japanese (Cook & Bassetti, 2005, p. 15). Also, it is useful to explore how participants process unfamiliar names in their L1 in order to identify whether these behaviours transfer to L2 processing (Koda, 2005, p. 315).

When you read something in English, do you ever see names that you don't know? What do you do when that happens?

Following on the set of questions related to unfamiliar L1 names, participants are asked whether a similar situation occurs when reading in their L2 (English), and if so,

what strategies they use for handling unknown names. The answers to these questions are relevant to Research Question 2 (what strategies Japanese L2 readers use when meeting unfamiliar names in a text).

How do you feel when you are reading a text with lots of English names? Or a text that has names you aren't sure about?

This set of questions was meant to investigate affective factors related to proper names and reading, in order to answer Research Question 1 (how Japanese learners of English feel about unfamiliar proper names while they are reading in English). Asking participants how they feel when reading texts heavy with proper names is an attempt to gauge if these readers have a general feeling of apathy toward unfamiliar names. Alternatively, readers may have a sense of anxiety when they see unfamiliar names. For example, Kobeleva (2012) found in her study with L2 listeners that unfamiliar proper names caused anxiety for the participants.

How would you describe your level of knowledge of English names? Would you say you know a lot of names, enough names, or a few names? Do you think it's important to know about English names, for example, which are family names and which are first names?

Participants are asked to self-report on their level of knowledge of English names, and whether this knowledge is sufficient or lacking for reading tasks. The purpose of these questions is to investigate the importance these learners place on proper names as special kinds of words, and whether they feel their knowledge of this type of lexis is sufficient for their academic studies. In this respect, the questions were also targeting Research Question 1 (affective factors).

Do you recognise names easily/quickly while reading? How do you recognise a name in English (what clues do you look for)? Do you think it's easier to recognise names in Japanese or English?

These questions are meant to explore participants' awareness of their decoding strategies for identifying English proper names. This set of questions targeted both Research Questions 2 (strategies) and 3 (how well they are able to decode and comprehend unfamiliar names while reading).

Is there anything you'd like to add? What should I have asked you that I didn't think to ask?

The closing question was included to allow the interviewee the final say (Dornyei, 2007), and to elicit any further comments from the interviewees on the topic of proper names. In summary, it should be noted that the interview questions were phrased neutrally to avoid researcher bias.

Read-aloud task

A read-aloud task was given to each participant after the interview questions. The purpose of the read-aloud task was to gain insight into whether proper names are always recognised in context, and what information can be inferred about each name (to address Research Question 3). Participants were asked to read aloud a short excerpt (323 words) and say who or what each proper name in the text refers to. The excerpt was taken from a graded reader, *My Family and Other Animals* by George Durrell, published by Penguin Readers (see Appendix 1.2 for the text excerpt). Graded readers are condensed books for L2 readers in which the vocabulary has been graded or matched to certain vocabulary levels. The reader was designated Level 3, the same level that these participants were reading in their extensive reading program. The excerpt has ten proper names: names of four people and a pet (*Larry, Leslie, Margo, George* and *Roger*) and five place names (*England, Greece, Spain, Italy* and *Corfu*). I predicted that most of the names would be familiar to the participants, while three would be unfamiliar: *Leslie, Margo,* and *Corfu*. This mix of familiar and unfamiliar names was considered appropriate, as participants would have confidence with most of the names; however, a few unfamiliar names would present a challenge so they could demonstrate their decoding and inferencing skills.

3.2.3 Procedure

At the beginning of each interview session, before recording began, I reiterated the purpose of the interview and gave the participant a consent form to sign (see Appendix 1.3). The consent form, in English and Japanese (L1), gave assurances that the student's grade would not be affected and that the data collected would be kept confidential. The form also outlined what participation involved (i.e. answering questions about reading attitudes and strategies, and reading a short text aloud). Lastly, students were assured that they could withdraw from the research at any time

without penalty. After the student had signed the consent form, I explained that I wanted to record the interview and asked for their permission. All agreed. The interviews were recorded on a MacBook Pro using Audacity software version 2.0.6.

The interviews lasted between 18 min 13 s and 23 min 15 s, including the read-aloud task. The average length of interview was 20 min 25 s. The interviewer asked questions and waited for the participants to answer. If the interviewees did not understand the question, it was rephrased. Where interviewees gave very short answers, they were encouraged to give an example or to explain further.

The read-aloud task was conducted after all the interview questions had been asked. Instructions given to students (orally) were to read aloud and stop reading every time they came to a name of a person or place, and to say who or what that name referred to. I modelled the read-aloud task for each student by reading a short paragraph from a different graded reader (*Strangers on a Train* or *The No. 1 Ladies' Detective Agency)*. One participant (K) asked if he should explain his reason for the name reference given during the read-aloud; he was told it was not necessary. When another participant (T) did not stop reading after several proper names had been mentioned in an extended portion of the text, he was interrupted and asked if there were any names in that portion of the text; he was also reminded to say what or who the names referred to.

3.2.4 Data analysis

Transcription of interviews and read-aloud tasks

The interviews were transcribed by the researcher/interviewer. The interview transcripts are provided in full in Appendix 1.4. Guidance on transcribing speech orthographically was taken from Wray and Bloomer (2012). The main interest of this study was the content of *what* the participants said as opposed to *how* they said it. Therefore, it was felt appropriate that the transcription focused for the most part on the information provided by the interviewees (Wray, personal communication). As a result, only the following detailed features of speech were transcribed: any non-verbal communication, such as laughter; unusual pronunciation of lexical items; and long pauses. Any comments or explanations that may aid the reader of the scripts are included in brackets (e.g. explanation of Japanese utterance). The main features of speech that were transcribed are shown in the key in Table 3.2.

Table 3.2

Key: Main features of transcription

Symbol	Feature
I:	interviewer's speech
A/K/T/W:	participant's speech (initial letters of interviewees' names)
[ ]	non-verbal communication, e.g. laugh
/ /	non-standard pronunciation in phonetic script
(.)	long pause
( )	comment or explanation

Punctuation symbols are not normally included in orthographic speech transcriptions. However, to facilitate reading of these scripts, the following regular punctuation conventions were included: question marks to signal questions, periods to mark the end of statements, and commas to indicate short pauses. The participants' speech was transcribed as it was heard on the audio recording by the transcriber; that is, no grammatical errors in their speech were corrected, and no vocabulary choices were altered.

For the read-aloud task, it was felt that non-standard pronunciation of names or other vocabulary items that were not names should be noted in the transcription. Since uncommon pronunciation of names might be indicative of the participants' unfamiliarity with the items, this feature was important for analysis. Unusual pronunciations were noted using phonetic script. For example, this unusual pronunciation of *Corfu* is taken from T's interview:

(1) T: George says /kɔːf/ is wonderful. Why don't we go there?

Data coding

After the interviews were transcribed, the next stage of analysis involved data coding. Typological analysis was used to analyse the data. In typological analysis, the data is divided into categories based on pre-determined typologies, which can come from theory, common sense or the research objectives (Hatch, 2002). For this study, the

typologies were derived from the research questions. Hatch (2002) notes that typological analysis works best with structured interview schedules in that the researcher can be confident that particular points will be addressed in the interview. Semi-structured schedules were used in this study; perhaps because of their intermediate proficiency level, the interviewees did not deviate from the interview schedule. Therefore, it was felt a typological approach to analyse this data would be appropriate.

Hatch (2002) outlines nine steps in typological analysis (p. 153), which were followed here. First, typologies are identified from the research questions, and then the data is read and coded to the typologies. Next, a summary sheet by typology is made of the main ideas that emerged from the initial analysis. This summary sheet is examined for patterns (regularities), relationships (links) and themes (integrating concepts) in the typologies. Then, the data is read again and coded for patterns. Examples are looked for that do not fit the patterns. Links in the patterns are identified; these patterns are written as one-sentence generalisations. Lastly, data excerpts are selected that best support the generalisations.

The typologies identified from the research questions were as follows. First, with regard to participants' affective responses to proper names (Research Question 1), the code 'difficulties' was given to passages where interviewees mentioned challenges with processing proper names; it was also given to entries in the read-aloud task where participants demonstrated problems with decoding proper names (Research Question 3). The code 'helpful' was used for entries where interviewees mentioned how proper names aided comprehension (Research Question 1). Lastly, the code 'strategies' was given to entries concerning approaches to handling unfamiliar proper names (Research Question 2).

A summary sheet was made based on the three typologies and examined for patterns. Then, a secondary analysis of data was done to code for the patterns identified in the summary. Under the 'difficulties' typology, three patterns were identified in the interviews and read-aloud task: distinguishing between first and family names; understanding proper name referents; and phonology of names. Under the 'helpful' typology, no further patterns emerged. Under the 'strategies' typology, seven specific approaches to handling unfamiliar proper names were identified across the data: ignoring; guessing from context; using pronoun references; using name charts; looking

for orthographic clues; using grammatical knowledge; and doing an online or dictionary search.

Finally, the data was checked for examples that contradicted the patterns identified, and links were identified between the interviews and the read-aloud task. Generalisations were formulated as themes that emerged from the data. These generalisations are presented as sub-headings in the Results and Discussion section (3.3), and data excerpts are presented to support the patterns identified under each theme.

3.3 Results and Discussion

In this section, three major themes that emerged from the data will be presented: difficulties that L2 readers have with proper names; ways that proper names can aid comprehension; and strategies that L2 readers use with unfamiliar names. Under each theme, patterns identified in the data will be explained and illustrated using data excerpts that best exemplify those regularities.

3.3.1 Difficulties in understanding or pronouncing proper names

From the challenges with proper names that participants mentioned in the interviews or demonstrated in the read-aloud task, three patterns emerged: problems in distinguishing between personal and family names; difficulties in understanding or identifying proper name referents; and challenges related to the phonology of names. Excerpts from the data are given to illustrate each of these patterns. Findings are interpreted in reference to relevant literature. The key for transcription symbols was listed above in Table 3.2. Ellipses (. . .) within the quotes denote that some speech has been omitted.

Distinguishing between personal and family names

It may be unsurprising that Japanese readers of English would have difficulty distinguishing between personal and family names: Japanese names are usually written with the family name first, followed by the personal name. Korean and Chinese names are also written this way. This is, of course, the reverse order of how names are usually written in English, with the notable exception of academic citations. Name order

can be a source of confusion for Japanese students. When looking at an author's name on a book cover, for example, they are uncertain which is the family name; if they draw on L1 knowledge, they might infer that the name that appears first is the family name. Three of the interviewees mentioned the difficulty in differentiating between family and personal names. W raises the problem with regard to citations, a new academic skill she is learning:

(2) W: I can't decide the first name or the last name. . . . In Japanese, I can find this is the last name or the first name. But in English, I can't imagine that so, for example, in citation, I can't which is one should put down the statement.

This difficulty can result in citations listed by personal names, not family names. There are other situations where this confusion over name order results in inappropriate usage, for example, addressing people with a title and personal name (e.g. *Professor Mike; Ms. Elizabeth)*.

Identifying proper name referents

A prevalent assumption in L2 vocabulary research is that L2 readers can easily understand proper names in context. (This assumption is reviewed in detail in Chapter 4, section 4.2). However, the participants in these interviews discussed various struggles they had identifying proper name referents in context. For example, two participants mentioned problems related to the gender of personal names; this seems unsurprising in that they are processing names from a different culture. Also, two participants referred to mix-ups with unknown place names; this is to be expected as place names are often not explained in context (Nagy & Anderson, 1984). Another difficulty that participant A remarked on is when characters in stories are referred to by more than one name. When I asked her if she had anything further to say about proper names, A remembered a story she had been reading recently. She describes the confusion that resulted when a character in the story was called *brother* by his sister, and referred to by his personal name by his friend:

(3) A: Last month I read the book. . . . There are three persons, the main character was a girl, and second is her brother, older brother. And he is the friend of the brother. And she called her brother, brother. But he called . . . his name, and so sometimes . . . friend is talking about her brother but I confused who is he.

> Because usually . . . the main character call her brother, brother. . . in her conversation but sometimes, suddenly appeared the name.

Just as one character can be called by different names, different characters in a story can have the same name, perhaps creating a burden for the working memory. As the American author Elif Batuman (2010) observes of the characters in Tolstoy's *Anna Karenina*:

> Anna's lover and her husband had the same first name (Alexei). Anna's maid and daughter were both called Anna, and Anna's son and Levin's half brother were both Sergei. The repetition of names struck me as remarkable, surprising, and true to life. (p. 8)

This is a good illustration of how difficulties with proper names are not limited to L2 readers. It also demonstrates how context, descriptions and nicknames can be necessary for proper name differentiation (J. M. Anderson, 2007, p. 117). However, as A has shown in example (3), L2 readers might not always be successful using context to understand who is being referred to and by what name.

Nicknames also present a challenge for L2 readers. While nicknames and the link to the original name might seem obvious to L1 users, this connection might not be so clear to L2 users. When I asked A if it was important to learn about English names, she raised this point about nicknames. She explained how her lack of knowledge regarding nicknames resulted in confusion about her American friend:

(4) A: Yeah, because I have exchange student, name Tom. Actually his name is Thomas. Often the American or other foreign people say Thomas, Tom and I didn't know such things before I come to here (to university). So sometimes when I'm talking with other people I'm not sure who is he. So it's important.

Although example (4) did not come from reading but from her campus life, one can easily think of examples in literature where characters are referred to by nicknames. While nicknames or diminutive forms might seem so obvious as to not warrant explanation (e.g. *Tom, Thomas; Mike, Michael*), that is not the case for L2 users, as A has conveyed in example (4).

Some examples of misunderstanding proper name referents were observed in the read-aloud task. Recall that participants were asked to read aloud a short excerpt and to stop at each proper name to say what or who the name referred to. For two participants, proper names without overt referents in the text were misunderstood. In the text excerpt, the oldest brother Larry is trying to convince his mother to relocate the family and escape the rainy English weather. He mentions someone named George, who has told him that Corfu is a nice place. No information is given to the reader as to who George is or where Corfu might be. This lack of specificity creates difficulties for two participants. In example (5), A guesses that both names, *George* and *Corfu*, might be members of the family. The text is presented first as it appears in the book; this is followed by A's reading of that text.

(5) Text: 'George says Corfu's wonderful. Why don't we go there?'

A: George, George is brother, hmm (.) George? George, says Corfu's wonderful (.) Corfu. George and Corfu, heh (.) [laughs] Why don't we go there? Maybe they are family, members of the family.

Later in the excerpt, the family is sailing towards Corfu, and this is the second mention of Corfu in the text. In example (6), A remembers her earlier guess as to the meaning of Corfu (i.e. a family member), and retains this inference.

(6) Text: We slept when the boat left and then, very early the next morning, we watched for Corfu.

A: We slept when the boat left and then, very early the next morning, we catched (*sic*) for Corfu, the family.

A seems to remember her guess about the meaning of Corfu from the first mention (i.e. that it refers to a family). She does not check whether the guess is correct in the context of the second mention, though at the second mention, it could feasibly be a person. Huckin and Bloch (1993) noticed similar behaviour in their L2 readers. They found that when their participants had made an inference about a word, they did not use context to ascertain the plausibility of the guess; rather, they would retain the incorrect guess, even when contextual clues refuted that inference.

Pronouncing unfamiliar proper names

Two participants mentioned difficulties related to the phonology of proper names. When I asked A about the strategies she used when meeting new names in graded readers, A talked about her struggles with pronunciation of unknown names.

(7) A: Sometimes, for example, the last book I read, there are two persons, the names begin with A. I very confused. So in such situation I really concentrate on the name, very look carefully but usually I don't sure how pronounce especially English name or Indian name, it's very difficult. But I can understand . . . who is he if I can't pronounce so I not so care.

A's strategy use in (7) is an example of what Koda (1995) noted in her study with Japanese readers: when meeting new L2 words that they are not sure how to pronounce, they treat them as Kanji (Chinese script), and try to remember them visually. Also in excerpt (7), A seems to suggest that knowing how to pronounce the name is not important; later in the interview, however, she changes her mind. When I asked her about recognising names in Japanese and in English, she reconsiders the importance of pronunciation to comprehension. She tries to explain the difference between reading Japanese names, which are written in Kanji, and English names in excerpt (8):

(8) A: Of course I think pronounce is very important. Because if I know how to pronounce the name, it's easy to memorise because I think I can understand from eyes and ears. . . . Japanese is Kanji so Kanji usually can read only one or two ways. So one letter has only one or two ways to read. . . . Like Kanji is if I wrote like this (gestures on hand) it's meaning 'one'. But English is o, n, e and English makes the word use many letters so it's difficult.

Here, A has articulated the importance of having both an orthographic representation ("understand from eyes") and a phonological one ("and ears") for proper names. As Hulstijn (2001) notes, processing new lexis more elaborately, that is, through more than one dimension, will lead to greater retention. Thus, having both representations available to her is what makes processing Japanese names easier, and why processing English names is more difficult for her (i.e. the phonological representation is absent). When we discussed unfamiliar L2 proper names, A agreed that knowing how to pronounce the name can aid comprehension.

(9) A: Yes, yes. Because the book like Romeo and Juliet, it's very famous, so it's easy to understand ah, this is Romeo's phrase and if I could pronounce the name so I think it help me to read.

What A seems to say in excerpt (9) is that being familiar with the name is useful because she can pronounce it. And because she can pronounce it, this helps with comprehension.

It has been suggested that there may be more plausible phonological sequences for names than for other words, thus making names more difficult to recall or learn (Brennen, 1993; James & Fogler, 2007). In his plausible phonology hypothesis, Brennen (1993) suggests that learning of new phonology is done more often for names than it is for other words. He does not suggest that there are novel phonemes but rather new sequences or syllables. Brennen (1993) argues that when one encounters a new proper name with a novel phonological sequence, the name essentially represents a new word to be learnt; this can make recall difficult. He gives an example of such a novel name he came across: a researcher's name, *Intriligator*. It is important to note that Brennen (1993) presents his plausible phonology hypothesis in the context of L1 users. Whether the same is true for L2 users, that the learning of new phonology is done more often for names than for other types of words, is difficult to say; L2 users arguably learn a lot of new vocabulary. However, it is interesting to consider this hypothesis in light of the difficulties that the interviewees mention related to proper name phonology.

For instance, K also mentions struggles with proper name pronunciation. When asked what he finds difficult about names, K tries to explain why names might be harder to pronounce than other types of words. He uses the name of one of his teachers to illustrate:

(10) K: Because the pronunciation of name is quite different from the spelling. I learned a lot of words from junior high school and there are many words and almost pronunciation is same as spelling. So I can pronounce the word I don't know with the spelling. But name is quite different from spelling. Name pronunciation is different so I can't guess how to pronounce it. . . . Ah, like Miss Zoë Jenkins. Her name Zoë is difficult for me. Because the z, o and e (gestures a dieresis). And also Jenkins is difficult for me, maybe for Japanese to

pronounce. I only could know its pronounce Jenkins after she said, I'm Zoë Jenkins.

K's explanation in example (10) for why names are more difficult to pronounce seems to lend some support for the plausible phonology hypothesis for L2 users. K feels that he can work out the pronunciation for words, but when it comes to proper names, the orthography seems more distant from the phonology. It may be the case that he finds proper names difficult to pronounce in that they are low frequency items. As novel items, with potentially more plausible phonologies, proper names might seem more difficult to pronounce than other words.

From the read-aloud task, there are two instances noted in which participants misread a proper name. While it is very difficult to say what has occurred during a miscue in oral reading (Goodman, 1969), these two examples are presented here as possible demonstrations of difficulties related to proper name phonology and orthography. Recall that in the read-aloud task, the participant stops at each name and says to whom or what it refers. In the text excerpt, Larry and Leslie are brothers; the narrator has distinguished between the two by noting that Larry is the older brother, and Leslie has problems with his ears. At the second mention of *Leslie*, both A and W misread the name *Leslie* as *Larry*.

Let us look first at A's misreading. In the example (11), A misreads *Leslie* as *Larry*. Then, because she wants to say who this refers to, she looks back in the text, searching for the name *Leslie* to check if she has correctly remembered which brother it is (i.e. the brother with the problem with his ears). She finds the first mention of *Leslie* where reference is made to his problem with his ears. But when she returns to the current point in the passage, she does not alter her pronunciation of the name, again pronouncing the name as *Larry.* The original text is presented first in example (11), followed by A's reading of it and her explanations of the proper name referents.

(11) Text: We travelled by train with our clothes and our most important belongings: Mother's cookbooks, Leslie's guns, something for Margo's spots, Larry's books, my favourite insects and Roger, my dog.

A: We travelled by train with our clothes and our most important belongings: Mother's cookbooks, Larry's, Larry, Larry is brother who has illness . . . (looks

back in text). Ah yeah, with ears. Larry's guns, something for Margo's spots, Margo is sister, Larry's books, my favourite insects and [rɒgə] my dog, dog.

Notice that A does not say to whom *Larry* refers in the phrase *Larry's books*, perhaps because she feels she has just given the referent. She might think Larry has both guns and books to take on the trip. This is unlikely, however, as she seems aware that there are two brothers, one of whom has a problem with his ears. She does not seem aware that she has misread the name. I suggest that a possible explanation of her misreading of *Leslie* as *Larry* could be that she has not carefully analysed the individual letters in the name, and has focused on the initial letter only. If so, this could be an example of how L1 logographic readers do not analyse intraword components (i.e. individual letters) in English words as carefully as alphabetic readers do (Akamatsu, 2003; Koda, 1996).

Considering this possible explanation, it is interesting that W also misreads the same name at the same point in the text. For W, *Larry* seems to be an unfamiliar name because at the first mention of his name, she hazards a guess that *Larry* is a woman. Later, in example (12), she says that *Larry* refers to a person, and then, a man. Just as with A's misreading, W does not seem to notice reading Larry's name twice in this list of family members and their possessions. In example (12), the original text is presented first, followed by W's reading of the passage and explanation of proper name referents.

(12) Text: We travelled by train with our clothes and our most important belongings: Mother's cook books, Leslie's guns, something for Margo's spots, Larry's books, my favourite insects and Roger, my dog.

W: We travelled by train with our clothes and our most important belongings: Mother's cook books, Larry's guns. Larry means people, something for Margo's spots, Margo maybe person, Larry's books, Larry the man, my favourite insects and Roger, Roger is the dog, my dog.

Unlike A, W does not demonstrate whether she understands how the characters are related (i.e. by referring to them as sister or brother). Like A in example (11), W does not notice that she misreads *Leslie* as *Larry*, and thus repeats the name twice in the list. It may well be that both participants were fully aware of the two brothers and their names, and had simply misread the names due to anxiety of reading aloud, for

example. As noted above, it is very difficult to conjecture what has happened in an oral miscue (Birch, 2007; Goodman, 1969). It is of course possible that the participants had difficulty pronouncing the phonemes in the name *Larry;* it is difficult for Japanese speakers of English to make distinctions between the /l/ and /r/ phonemes. For that reason, the participants may have been overly focused on or worried about the name *Larry*. However, since two participants misread the same name at the same spot in the text, I offer it here as a possible illustration of how L1 logographic readers might not carefully analyse the individual letters in words, and instead focus on the initial capital letter (see Chapter 6 for more on L1 transfer and L2 decoding skills).

3.3.2 Proper names aid comprehension

Proper names might not cause only disruptions for L2 readers; indeed, names might aid comprehension by serving as an anchor to the events in a text. For example, when characters are introduced by proper names rather than their role names, they are usually more prominent in a text; also, named characters are usually more accessible to the reader for anaphoric reference (Sanford, Moar, & Garrod, 1988). Thus, while recruitment for the interviews was purposive in that students were selected who had mentioned difficulties with proper names, I did not want to focus the interviews on negative aspects of proper names. When asked if understanding proper names can aid reading comprehension, two participants gave examples of how they felt proper names were helpful. K suggests that knowing place names can help with reading.

(13) K: It's important for us Japanese to know the geographic because if the person will study about geographic, person will understand many countries name. Then it will be hint for reading. But I'm a little poor at geographic so sometimes it will be weak point.

K recognises the usefulness in knowing place names for text comprehension. T agrees that understanding proper name referents can help comprehension. He explains:

(14) T: Because the more we get to know the name, the more we can get easily remember and remember the story too. So yes it's important I think. . . . Because in the story, so for example, in the test, in the story, in the novel, we have to remember or we have to imagine for ourselves what the character looks like. And that's of course connected to name. And under the name, we make

> character for our own. But if we don't know name, that's we often forget what this character doing or not.

He seems to be saying that the more information one can infer about a name, the more helpful it can be for making connections in a text. This points to ways that teachers can support L2 readers, for example, by alerting them to characteristics that might be inferable from literary proper names in particular. Crystal (2006) notes that L1 children are taught from a young age to recognise characteristics behind names (e.g. *Goldilocks*). Such training would benefit L2 readers, for whom these connections are most likely not obvious. Training in looking for traits behind names might help L2 readers imagine characters and retain related information.

Nevertheless, there might be a limit to the connections readers can make from proper names. Certainly not all literary characters have names that indicate aspects of their appearance or personality; furthermore, not all characters have names. As Batuman (2010) notes in her discussion of names in Russian fiction:

> Chekhov's characters, many . . . didn't have names at all. In "Lady with Lapdog," Gurov's wife, Anna's husband, Gurov's crony at the club, even the lapdog, are all nameless. *No contemporary American short-story writer would have had the stamina not to name that lapdog.* They were too caught up in trying to bootstrap from a proper name to a meaningful individual essence. (p. 20)

Batuman's point here is that the scope for proper names to develop characterisation is limited. In this respect, proper names as a tool for L2 readers to create effective connections might be finite.

3.3.3 L2 readers' strategies for unfamiliar proper names

The final theme to be explored concerns the various strategies that these participants reported using when meeting unknown proper names. Several strategies were identified in the data: ignoring; guessing from context; searching in a dictionary or online; using pronoun references; using grammatical knowledge; using orthographic clues; and using name charts. Each of these strategies will be looked at in turn, drawing on excerpts from the interviews for support.

Ignoring

As was noted above, one problem for Japanese readers is that if they are not familiar with a particular L1 proper name, the pronunciation and referent might not be inferable from the Kanji (Cook & Bassetti, 2005). The reader will need to either research the name or ignore it. It was of interest how these participants handled L1 proper names because of possible transfer effects to L2 reading. When asked how they approach unfamiliar L1 proper names, two participants reported that they ignore the name and continue reading. As A suggests in (15), knowing how to pronounce the name is not important to understanding the main idea.

(15) A: I don't care. Because I want to know the . . . contents of article. So I don't care if I couldn't read the person's name correctly.

T agrees, saying that when he sees an unknown L1 proper name over and over again, it becomes more familiar and does not disrupt his reading.

Guessing from context

The other two participants reported that when meeting unfamiliar L1 proper names, they try to guess from context. As K explains, he first tries to guess from context, and if he is still not sure, he does some research to find out about the name.

(16) K: I guess the meaning from the context. . . . I often asked my mother or grandmother how to read. They will know about it.

As for using context to infer L2 proper name referents, none of the participants mentioned doing so.

Doing online searches or using a dictionary

Three of the participants mention searching in a dictionary or doing an online search to find more information about unfamiliar L2 proper names. The two comments in examples (17) and (18) are also illustrative of difficulties related to insufficient contextual clues and pronunciation.

(17) A: I thought the person's name is the place name. I misunderstood it is a place name. So I searched in dictionary. But it was person's, man's name. So dictionary said it's a man's name (.) It has very popular names.

(18) K: I will check it on the Internet and because spell is very complex and I really don't know how to even pronounce it. So I copy the spell, on Internet, it's easy for me. In Internet, I can know in Japanese because there are many facts on Internet dictionaries. . . . Like Wikipedia or Google maps.

The interest that participants show in knowing more about new proper names they meet is important: it indicates they consider proper names important enough to comprehension of the text to take time to research them. Perhaps they expect to come across those names again. In any event, L2 readers tend to look up lexical items that they consider relevant to understanding the text, and ignore those that are not considered relevant (Hulstijn, 1993). For that reason, it is important to note that these participants report taking time to research unfamiliar proper names.

Using pronoun references

Participants mentioned using pronoun references to keep track of different characters in stories. The gender of proper names is not always clear to L2 readers. A explains in excerpt (19) how she used pronouns to distinguish between two characters whose names both started with the letter *A*.

(19) A: The last book I read, there are two persons, the names begin with A. I very confused. . . . They were related, one of them is a girl, and the other is a man, an uncle of her so often they appear at the same time. The main character called the girl she or her, but uncle is he or him so I could organise the two person.

Excerpt (19) also points to the difficulty L2 readers can have when they do not have phonological representations of names. In this particular case, because both names start with the same letter, the participant was not able to distinguish between them from the initial letters (this strategy is discussed further below). So instead, she relies on pronouns to make distinctions.

Using syntactic knowledge

One participant mentioned using his grammatical knowledge to differentiate between unknown proper names and other unfamiliar lexis. T explains this strategy:

(20) T: Ah, look the words like the or a. We don't say a before the name. So first that's a strategy.

While one can think of exceptions (e.g. *the United Kingdom*), T has demonstrated that how his knowledge of English grammar helps him to recognise unfamiliar proper names.

Orthographic clues in English

Because proper names are capitalised in English, one might consider this a reliable clue for L2 readers. (However, recall that proper name related items are also capitalised; see Chapter 2, section 2.2.3). Three of the participants noted using the initial capital letter clue to identify proper names. Relying on the initial capital letter only to distinguish between characters is not always practical, especially when both names start with the same letter, as was noted in excerpt (19). This same difficulty arises in the read-aloud task for W. After W finished the read-aloud task, I asked her what strategies she would use to remember so many different characters, like the ones in this text.

(21) I: Ok, for example, this story has two people with the letter L, Larry and Leslie, so how could you distinguish between those two because almost same length, Larry and Leslie. How could you distinguish between those two?

W: Hmm (.) the atmosphere.

I: Ok, from the character?

W: Yes.

I: Ok, from what you read, what do you know about Larry?

W: He is the man.

I: Yes. What do you know about Leslie?

W: Leslie is the sister.

I: Yeah?

W: Hmm (looks back at text). Oh, Leslie is the brother.

I: Yes, he is also brother. So they are both men and they both start with L. So what strategy would you use then?

W: [Laughs]

W is a very reserved, quiet student, and in pointing out her referencing error about Leslie, I may have embarrassed her: she offered no further comment, only laughed. This strategy of relying on the initial capital letter to differentiate between characters illustrates how a lack of phonological representation can cause processing difficulties for L2 readers. Also, as previously noted, the /l/ and /r/ phonemes in the name *Larry* may have been causing processing difficulties for this Japanese participant.

Name charts and family trees

One way that publishers of graded readers offer support for proper names is by providing names charts or family trees to show the relationship between characters in the story. These kinds of charts are found in L1 novels as well, especially epic works spanning generations. In graded readers, the characters' names sometimes appear under head sketches, which match the action sketches throughout the book. In this way, the reader is supported with both pictures and a name chart to keep the various characters organised. K says that he sometimes refers to these charts if there are many characters in the story.

(22) K: But when names often appear, commonly the book has the page to list the names. . . . Hmm, the characters. At first or at last. So when I confused, really confused, I check it, return back to the page and check it, who it is.

When there is no chart to refer to, both K and A report making their own notes on characters and their roles in the story. This illustrates how useful such name charts are for readers: not only do L2 readers seem to make use of such charts, but the charts also provide examples to the students of how to keep characters organised in longer pieces of fiction.

3.3.4 Summary of results

In answer to Research Question 1 (i.e. what affective factors are involved in proper name processing), these interviewees reported feeling confused by unfamiliar names. For example, it is not always clear to them which are family or personal names (excerpt (2)), or which are male or female names (excerpt (21)). They also expressed frustration at not knowing how to pronounce new names because having a serviceable pronunciation can help to remember and follow the characters in a story (see excerpts (8) and (10)). On the other hand, some participants noted that proper names can aid comprehension; for example, knowing places names, and using names to build images of characters can help comprehension (excerpts (13) and (14)). Regarding Research Question 2 (i.e. strategy use), the participants discussed several strategies they use to approach unfamiliar names. These strategies included: ignoring; guessing from context (for L1 names); checking dictionaries or doing online searches; using pronoun references; drawing on grammatical knowledge; using orthographic clues; and checking name charts or making notes (see excerpts (15) to (22)). As for Research Question 3 (i.e. difficulties in decoding and understanding proper names), a few examples from the interviews illustrated problems in identifying proper name referents (e.g. nicknames and place names, in excerpts (4) and (17)). From the read-aloud task, some examples were observed of how similar looking names (i.e. *Leslie* and *Larry*) caused decoding and inferencing difficulties for two participants (see excerpts (11) and (12)).

3.4 Conclusion

In the interviews, participants reported having difficulties identifying proper name referents. Reasons for these difficulties included their unfamiliarity with family and personal names, the gender of personal names, and nicknames. This finding suggests that it might be incautious to assume proper names do not present a burden for L2 readers and can be treated as known vocabulary. Other challenges that the

interviewees remarked on concerned the phonology of unknown names. Some interviewees agreed that it would help comprehension to be able to pronounce the names. Because proper names are low-frequency items, it seems intuitively correct that L2 learners would have difficulty pronouncing them. In support of this conjecture, the plausible phonology hypothesis (Brennen, 1993) was considered from the perspective of L2 readers: it may be that they need to learn new phonological patterns more often for proper names than they do other lexical items.

The interviewees also provided examples of various strategies they use when meeting new names in reading texts. These strategies ranged from drawing on syntactic knowledge and orthographical clues, to doing online or dictionary searches and making notes on proper names. That the participants reported using these various strategies is important: strategy use suggests that proper names might present a processing difficulty for L2 readers.

In the read-aloud task, it was observed that these participants made wrong inferences about names either because contextual clues were lacking or they did not check the inference against the context. For example, one participant wrongly inferred that *Corfu* was a family member; the contextual support[9] was not rich to indicate it is a place. This finding supports other research that has shown L2 readers are not very successful at using context to infer meaning, either because contextual clues are lacking or readers do not always confirm their guess against the context (Bensoussan & Laufer, 1984; Nassaji, 2003b). In this respect, it would have been good to have used a control for this study: the text used in the study may not have been rich enough in context for the participants to make inferences. Grabe (2009), in his discussion of background knowledge and how it affects comprehension, notes that weak readers can draw on the wrong information and make wrong inferences (p. 74). The findings from this study suggest this can happen with proper names as well.

There were a few limitations to note concerning this study. Although several interesting themes emerged from the interviews, the L2 oral proficiency levels of the participants constrained the amount of data collected. Two of the participants were very quiet in the interviews. In the profile of participants, I noted that W was a reserved student in the classroom. I thought she might open up in the interview because her peers were not there; this did not happen. She gave very short answers and did not elaborate when

[9] Contextual support here refers to the other words in the sentence that may have helped the participant understand what *Corfu* referred to.

invited to do so. T was also very quiet in the interview, which was surprising because in the classroom he was very talkative. He also gave very brief answers. The other two participants (K and A) were quite animated and forthcoming in the interviews. This is perhaps evident in the Results and Discussion section (3.3), as many of the examples were drawn from their interviews. K and A offered fuller answers in terms of providing concrete examples and explanations.

There are two possible solutions to conducting interviews with participants of lower proficiency levels. The first is to conduct the interviews in the L1 so that participants are able to express their answers more fully and completely. Another possibility is to give the students the questions in advance of the interview. That way, they could have time to consider the questions being asked, and to think of examples to support their answers. Indeed, it was often at the end of the interviews, when I asked if they had anything further to add, that participants were able to offer specific examples of difficulties they had, or strategies they used for L2 proper names. Thus, if I had given them a few days to think about the questions and prepare some answers, this might have reduced their anxiety and in turn, produced more insightful comments.

Another limitation to the study concerns the read-aloud task. The aim of the task was to identify decoding difficulties, and to learn more about how information is inferred about proper names. However, as has been noted, it is very difficult to know what has occurred during an oral reading miscue (Goodman, 1969). Reading aloud requires much processing and attention (Birch, 2007). Thus, the data from the read-aloud task needs to be interpreted with caution. The participants were probably anxious about the task (i.e. reading aloud while being recorded). One can imagine they did not want to make any errors. This anxiety would have taken away from their attention for the reading task. In these respects, the task was not ideal to capture decoding and inferencing difficulties related to proper name processing.

Due to the exploratory nature of this study, more questions may have been raised than answered about L2 proper names. However, these questions concerning how L2 proper names are processed can be used to move this investigation forward. In particular, three points that emerged from this study are identified as deserving further investigation. First, there is the issue that as a qualitative study, the sample size was small. Also, recall that the selection of these participants was purposive: prior to the study, they had self-reported having some sort of difficulty related to proper names. In this respect, these participants might not be representative of most Japanese L2

readers of English. For that reason, it merits conducting a study with a random, larger sample. Also, it is worth looking at how L2 readers deal with proper names in reading tasks without drawing explicit attention to the issue of proper names, as was done in this qualitative study. This is the focus of Chapter 4: how L2 readers approach proper names in reading texts.

Another issue raised from this study is that many of the difficulties that the participants reported with proper names stemmed from their unfamiliarity with these items. For example, participants mentioned being unfamiliar with: the gender of some proper names; nicknames; places; and surnames vs. personal names. Therefore, it seems worthwhile to investigate the effect of familiarity with proper names on reading comprehension. It may be that whether an L2 reader is acquainted with the proper names in a text or not has little or no effect on comprehension. Conversely, an L2 reader's familiarity with proper names in a text might aid comprehension. For example, there may be an effect for pre-teaching of proper names in texts. The effect of proper name familiarity is the focus of Chapter 5. The third topic identified as deserving further investigation concerns the read aloud task: while this task was not ideal for capturing the challenges of proper name processing, some of the data suggests that context might not always be reliable for proper name inferencing. Thus, it warrants investigating the extent to which L2 readers can use context to recognise and understand proper names; this is the focus of Chapter 6.

Chapter summary

The aim of this chapter was to explore L2 readers' perspectives on proper names. This aim was achieved by gathering self-reported data from L2 readers on: how they feel when encountering unknown proper names; what strategies they use for new names; and any difficulties they experience in proper name processing. It was found that main source of confusion for the interviewees stemmed from their inexperience with proper names; that is, uncertainty about which are family names and personal names; the gender of names; nicknames; and proper name phonology. Participants also reported various strategies they use when encountering new proper names.

In the next chapter, a comprehensive review is given of how proper names have been handled in L2 vocabulary research. Rationale is provided for why an assumption that L2 readers can easily understand proper names should be empirically investigated. A

study that investigates this assumption is presented. The study explores how L2 readers approach proper names with text difficulty as a variable.

Chapter 4: How L2 readers approach proper names

4.1 Introduction

It was established in the previous chapter that for some Japanese L2 readers of English, proper names can be problematic. The interview participants reported various processing difficulties that stemmed from their inexperience with certain proper names. Their inexperience resulted in challenges related to identifying proper name referents, distinguishing between personal and family names, and pronouncing novel names, for example. It was noted that the sampling for the interviews had been purposive, which may have impacted the findings. Also, as the study was exploratory in nature, participants' attention had been drawn to issue of proper names in reading. Given that these participants reported difficulties with proper names, the next consideration is whether the findings are applicable to a larger sample from the population, using a reading task in which explicit attention is not drawn to proper names. In other words, how do L2 readers from a random sample approach proper names in a reading text?

Some of the implications of knowing how L2 readers approach proper names while reading were considered in Chapter 2. For example, researchers can make informed decisions as how to categorise proper names in lexical analyses of reading texts. L2 reading teachers can decide how much, or how little, attention to give to proper names in a given text. Similarly, test writers and material developers can consider the processing load of the proper names that appear in their materials. For these reasons, it would be informative to know more about what L2 readers do when they meet unfamiliar proper names in reading texts. For instance, they might recognise that the unfamiliar item is a proper name, try to guess its referent and then continue reading. Alternatively, they might consider knowing the referent of the proper name important enough to look it up in a learner dictionary or from an online source.

The purpose of this chapter is to investigate how L2 readers handle proper names while reading. The chapter begins with a literature review, focusing on how L2 vocabulary researchers have treated proper names in lexical analyses of texts. From this review, it emerges that by and large, researchers have worked on an assumption that L2 readers can easily recognise and understand proper names in context. The review is followed by an empirical study that aims to investigate this assumption. The

study looks at how L2 readers handle proper names with text difficulty as a dependent variable. Text difficulty is determined by the percentage of known vocabulary, as was discussed in Chapter 2 (section 2.3). The overarching aim is to determine what L2 readers do when they encounter unfamiliar proper names and thus, explore how sound the assumption is that proper names are known vocabulary.

4.2 Treatment of proper names in reading and vocabulary research

Researchers investigating how much vocabulary an L2 reader needs to know in order to read certain kinds of texts with adequate comprehension are presented with the problem of proper names. For example, should proper names be treated as known or unknown to the L2 reader? Should proper names be handled like other lexical items and classified according to frequency? Several early studies beginning in the 1980s speculated that the L2 reader could easily infer the meaning of proper names. For example, Hwang and Nation (1989), in their investigation of vocabulary learning through newspaper reading, argue 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" that 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), in their examination of graded readers as a means for vocabulary learning, 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). In a study that investigated learners' productive vocabulary knowledge, Laufer and Paribakht (1998) take a different view of proper names. The authors note that when examining students' compositions, they "deleted proper nouns because [the authors] did not consider them part of learners' vocabulary knowledge" (p. 375).

L1 reading researchers were also investigating the demands of vocabulary knowledge on reading comprehension, at around the same time as L2 reading researchers. Carver (1994) investigated L1 reading text difficulty in terms of the percentage of unknown vocabulary. The participants were elementary school and university students. Participants were given different reading passages, from easy to difficult levels. Instructions to participants were to underline any words that they did not know the

meaning of or had not seen before. To encourage honest responses, participants were told that texts were of varying difficulty, so that while some students might underline many words, others might underline only a few. In his analysis, Carver (1994) created a list of Unknown Basic Words and excluded any words from that list which would not be taught in a vocabulary or spelling lesson: proper names, numbers, foreign words, abbreviations, and hyphenated words (p. 419). Having excluded those items from his analysis, Carver concluded that reading texts were easy if no vocabulary was unknown; texts matched the reader's ability if only 1% of vocabulary was unknown; and texts were difficult if 2% or more of the vocabulary was unknown. This is the same conclusion that subsequent L2 reading studies came to. Now it is generally agreed that 98% vocabulary coverage is needed for adequate comprehension of a text (Laufer & Ravenhorst-Kalovski, 2010; Nation, 2006; Schmitt et al., 2011). (The Schmitt et al. (2011) study was reviewed in detail in Chapter 2, section 2.3).

What is noteworthy for discussion here is Carver's (1994) justification for removal of proper names from his analysis: he argues that like the other eliminated items, proper names would not be taught in an L1 classroom as vocabulary to learn. In this respect, Carver may be treating knowledge of proper names as encyclopaedic knowledge and not lexical, though he does not explicitly state this. This is in contrast to Nagy and Anderson (1984), who do acknowledge the "complex issue" surrounding proper names, that is, whether to treat comprehension of these items as world knowledge or vocabulary knowledge (p. 361). In their vocabulary analysis of L1 school textbooks, they argue that some proper names are often not explained in context, such as geographical names, and that lack of knowledge of such names would result in comprehension failure just as unfamiliarity with the meaning of any other word might. Nagy and Anderson (1984) 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 higher school grades (p. 317).

Returning to L2 reading studies, Hu and Nation (2000) analysed vocabulary in fiction texts to gauge the lexical load for L2 readers. The authors conclude that to reach 98-99% coverage in a fiction text, a vocabulary size of 5,000 words would be needed, assuming proper names are known (Hu & Nation, 2000, p. 407). The authors note that proper names account for 4-5% of the running words in fiction, citing Francis and Kucera (1982). Thus, their conclusion rests on the assumption that the L2 reader

understands those 4-5% of proper names in a fiction text, an assumption that does not have empirical backing.

By looking at the Hu and Nation (2000) study in more detail, one can see how the authors handled proper names in their investigation. The experiment involved the manipulation of a short fiction text (673 words) to test the effect of different levels of vocabulary coverage (80%, 90%, 95%, and 100%) on comprehension by replacing low frequency words with nonsense words. The authors state that low frequency words were chosen by referencing Francis and Kucera (1982) and cross-checking with West's (1953) General Service List (GSL). They report that there were 14 proper name tokens, which were "left unchanged in the text and were counted as familiar words" (p. 411). However, one can see from their Appendices that in fact there was manipulation of two of the 14 proper name tokens: *Governor* and *Northfields Hospital*. In the original version (i.e. the 100% vocabulary coverage version), *Governor* appeared in this sentence as a proper name:

(1) *George Hitchcock?. . . But that's the name of the Governor of the hospital.*

Although the Governor's name is not repeated after his title, this is an example of what Allerton (1987) refers to as a mixed proper name, which has a descriptive common noun element (e.g. *President Obama).* In the manipulated version (i.e. 80% vocabulary coverage), a pseudo-word replaces *Governor*, even though *governor* does appear in the GSL 2,000 band, and therefore is not a low frequency item. Moreover, the pseudo-word loses the capital letter in the manipulated version:

(2) *But that's the name of the phantropist of the slivian home.*

By removing the initial capital letter, the pseudo-word becomes the equivalent of a common noun, and is no longer the title of a particular person. Of course, it is difficult to say if the omission of the initial capital letter had a direct impact on readers' comprehension of the 80% version. One could argue that an initial capital might have provided the reader with a clue that the item referred to a person, in addition to the sentential context and the *–ist* suffix ending. That the authors first counted *Governor* as a proper name token, and then replaced it as a low frequency item in the manipulated version demonstrates a lack of consistency in their handling of proper names.

Turning to the other manipulated proper name in Hu and Nation's (2000) text, the common noun *hospital* was a part of the name *Northfields Hospital* in the original version; it was changed to *Northfields Slivian Home* in the 80% version. Again, it is not clear why they changed this item because *hospital* does appear in the GSL 2,000 band and is not a low frequency item. However, more important for our discussion than the issue of frequency is that *Slivian Home* first appears in the text as part of a proper name with initial capital letters, and on second mention, it appears in lower case (i.e. *slivian home*). By changing the case of the letters, the authors assume that the L2 reader can make a connection between the first and second mention:

(3) *This is the Northfields Slivian Home*.

(4) *But that's the name of the phantropist of the slivian home.*

If one assumes that the initial capital letter on a proper name will help the reader to recognise it as such, then to alter the case in this way has processing implications for the reader.

There is also a lack of consistency noted in how proper names are handled in Nation's (2006) study, a frequently cited paper. Nation (2006) reports on his creation of 14 lists of 1,000 words each, based on the British National Corpus (BNC). Nation (2006) describes how he placed proper names in a separate list, word list 15[10]. To demonstrate the use of the word lists, he provides text extracts in which words beyond the first 1,000 band (1K) are marked with the number of the word list to which they belong. One can see from these extracts that his treatment of proper names is not consistent. For example, *Paris* belongs to list 7, and *Rome* to list 2, while *Florence* appears in list 15 (Nation, 2006, p. 68). One assumes then that *Paris* and *Rome* are more frequent in the BNC than *Florence* but Nation does not explain the discrepancy in treatment of these names, only stating, as noted above, that he placed proper names in a separate list.

Later in the same paper, Nation provides a lexical analysis of *Lady Chatterley's Lover*, showing text coverage percentages with and without proper names. He points out that 95% text coverage is achieved at the 4,000-word level, "assuming that proper nouns are easily understood" (Nation, 2006, p. 70). The table provided shows that if proper

[10] Word list 15 is a list only for proper names.

names are not assumed known, then 95% text coverage is achieved only at the 7,000-word level. Nation concludes that if "proper nouns can be counted as having a minimal learning burden" then 98% text coverage (the desired minimum) is obtained at the 9,000-word level; that is, a reader with a vocabulary size of 9,000 could read this novel without meeting too many unfamiliar items (2006, p. 70). However, when proper names are not treated as known, then 98% coverage is not achieved even at the 14,000-word level. Therefore, treating proper names as known or unknown often represents the difference between a reading goal that is attainable for an L2 reader and one that is much more demanding. Of course, it should be noted that by providing this comparison of text coverage with proper names known and unknown, Nation (2006) acknowledges that the assumption regarding proper names is debatable.

For his text analysis, Nation used the Range program (Nation & Heatley, 2002), which creates a vocabulary profile of a text based on frequency. The program allows for proper names to be excluded from analysis by ignoring mid-sentence words with an initial capital letter. Cobb (2010), discussing VP-Compleat, his online version of the Range program, says that users of the program need to understand that proper names are not lexical items (p. 187). He offers a simple example: *Pierre lives in Beaurepaire.* Cobb (2010) suggests that this sentence is comprehensible enough for the L2 reader without knowing more than that *Pierre* refers to a person and *Beaurepaire* to a place. He adds that other information (e.g. that these words are French, and that *Pierre* refers to a male) is not necessary to process the sentence. However, in continuous text, there may be several characters in different places, and retaining this information accurately may be important to comprehension of the text. Therefore, it may be helpful to know that *Pierre* refers to a man, and *Beaurepaire* is a French name. In addition, having phonological representations of the names can aid the working memory; as was seen in Chapter 3, it should not be assumed that L2 readers have the recoding skills to pronounce unfamiliar names. Moreover, having efficient word recognition skills to distinguish *Pierre* from *Perrin,* for example, and *Beaurepaire* from *Beauclaire,* is also important. In short, one should not assume L2 readers always have either the requisite phonological decoding or word recognition skills to effectively understand proper names in a text. The importance of lower-level processing skills to L2 reading comprehension will be discussed further in Chapter 6.

Moving forward in the vocabulary coverage research to more recent studies, one can see that the problem of how to handle proper names persists. In addition, the assumption that the L2 reader can easily understand the meaning of proper names in

context prevails. Authors of these later studies begin to cite the assumptions made in the earlier studies as justification for their treatment of names. For example, in a study investigating the vocabulary demands of television, Webb and Rodgers (2009) follow Nation's (2006) approach to proper names, first placing them in word list 15 for analysis, and then in their discussion, showing text coverage percentages with proper names listed separately and with proper names added to the first 1,000 band (i.e. treated as known vocabulary) (p. 345, 346). In doing so, the authors seem to acknowledge the difficulty in what to do with proper names. For again, this makes a significant difference in text coverage, with 98% coverage achieved at the 7,000-level if proper names "have a minimal learning burden"; 98% coverage is not achieved at all in these lists when proper names are listed separately (Webb & Rodgers, 2009, p. 345, 346). The authors do note that proper names made up 2.96% of tokens and 11.91% of types in their analysis of TV programs. Indeed, proper names were the largest lexical category after the 1,000 and 2,000 lists. The authors acknowledge that this large percentage illustrates the importance of "being able to recognize proper nouns when watching TV" (Webb & Rodgers, 2009 p. 345). Webb and Rodgers (2009) suggest that teachers help learners by pre-teaching low frequency vocabulary but they make no mention of pre-teaching proper names (p. 357). This is surprising given that proper names account for nearly 12% of the word types in their lexical analysis.

Matsuoka and Hirsh (2010) also cite previously made assumptions and contribute their own beliefs about learners' knowledge of proper names. Their investigation looked at learning opportunities for academic vocabulary in English Language Teaching (ELT) course books. The authors cite Hwang and Nation (1989) and Hirsh and Nation (1992) to account for their position on proper names, that is, names have either been previously learnt in the L1 or can be easily inferred from the context (Matsuoka & Hirsh, 2010, p. 58). For their lexical analysis of ELT course books, the authors created baseword lists, which "allow for the classification of words in a text according to word families" (p. 60). They explain the creation of additional baseword lists beyond the first 2,000 and the Academic Word List (Coxhead, 2000): proper names, technical words, textual words and technology. These additional baseword lists were "considered to be known or easily understood by learners at the upper-intermediate level . . . by one of the authorial team who has taught EFL extensively over the past 12 years" (Matsuoka & Hirsh, 2010, p. 61). From the analysis, one can see that proper names make up the third largest baseword list after the 1,000 and 2,000 word lists. At 2.8% (of types, 504 tokens), there are more proper names than academic words (2.1%), which is the target vocabulary the authors wish to investigate the learning opportunities for. It is

noteworthy here that one of the authors has postulated, based on his prior teaching experience, that learners will know or easily understand all 504 proper names found in their vocabulary analysis. However, there is no empirical support for this assumption.

Webb and Macalister (2013) suggest that learners can recognise and understand proper names due to the high number of names they will encounter while reading. Their study compared vocabulary coverage in texts written for children, language learners and adults. Citing Nation (2006), they first explain the inclusion of proper names in the calculation of coverage because "they may have a lower learning burden and be more easily learned" (Webb & Macalister, 2013, p. 309). The authors state that learners who engage in extensive reading should have learned to recognise names because of the initial capital letter used to indicate names. In their results, the authors point out that graded readers (i.e. simplified readers for language learners) have the highest percentage of proper names (4.02%) among the different text types in their study. Because of this high percentage of names in these texts, the authors suggest that, "language learners may quickly develop the skills to recognise and understand proper nouns to some degree. This is likely to allow readers to process the text more easily when they encounter unknown proper nouns rather than other unknown words" (p. 311). They note, however, that learners might not acquire proper names in the same way as other words because proper names tend to be context specific and therefore, the rate at which these items are encountered might be very different than other words.

While Webb and Macalister (2013) are careful to hedge their claims about L2 readers and proper names, there are a few concerns with their assertions. First, consumers of graded readers are typically beginner to intermediate learners. It is not cogent that language learners at this proficiency level will quickly develop word recognition skills. Nassaji (2014) points out how researchers and teachers often erroneously equate increased L2 proficiency with efficient L2 lower-level processing skills. He notes that even high proficiency L2 users tend to read more slowly than L1 readers, most likely due to deficiencies in word recognition skills (Nassaji, 2014). So to assume that readers at a lower proficiency level will quickly develop word recognition skills is unfounded. Secondly, as noted in Chapter 2, L2 reading is a more complicated process than L1 reading because of the dual-language involvement. If the orthography of L2 reader's first language differs with respect to how proper names are marked, this might have processing implications with respect to English names. Finally, it is not clear why meeting unfamiliar names allows for easier text processing than meeting other

unknown lexis. As noted above, names are not always explained in context, and an unfamiliar name might cause a breakdown in comprehension, just any other unknown vocabulary item might (Nagy & Anderson, 1984).

While the assumption in the literature is that proper names are not problematic for L2 readers, it is not clear why this might be the case. For example, Horst (2013) presumes that proper names are transparent for L2 readers. In her paper, she argues for a more focused pedagogical approach to ensure that learners acquire the first 2,000 most common words. She illustrates the advantages of knowing these items by providing a text excerpt (approximately 200 words, titled "Exploring the Arctic") from a language test that her Arabic students needed to pass. She shows that knowledge of the first 2,000 words, plus some "presumably transparent" place names like *the Arctic* and *Scotland,* would render this text comprehensible, leaving only six unknown words (Horst, 2013, p. 174). (There are three other proper names as well, the name of a ship, *The North Star*, an island, *Bear Island,* and another country, *Norway).* To take one of these names as an example, one can see that understanding *the Arctic* is important to comprehension of the passage. But it is not apparent why Horst presumes that *the Arctic* is transparent to Arabic readers. In Arabic, *the Arctic* is القطب الشمالي, pronounced as /alqutb alshamali/, and translates literally as the *North Pole* ("the Arctic," n.d.). Thus, there is no phonological clue for Arabic readers as to the meaning of *Arctic*, nor would the literal translation of the Arabic name help them in this text as no mention is made of the North Pole.

Yet these assumptions continue to be cited and used by other researchers as supposedly sound rationale for treating proper names as known. In fact, so ingrained is the notion that names are known to L2 readers that in the most recent studies, one can see a trend toward obscuring or removing names completely from the lexical analysis. For example, Douglas (2015) justifies his classification of proper names in the first 1,000 frequency band (i.e. known vocabulary to the L2 reader) by quoting Horst's assertion that names are "usually transparent" (2013, p. 176). He goes on to state that proper names are "easily acquired and understood by learners", citing Nation, 2006; Schmitt, 2008; and Webb and Rodgers, 2009 (Douglas, 2015, p. 52). Unlike the earlier studies discussed above, in Douglas' (2015) lexical analysis, the percentage of proper names in the text is not apparent because they are hidden within the 1,000 band.

An example of a study in which all proper names have been completely removed from the analysis is seen in Uden, Schmitt, and Schmitt (2014). In their vocabulary analysis

of two graded readers and two unsimplified novels, the authors report that most of the proper names were people's first names or places names, and thus "despite being low frequency, would be easily understood from context," citing Nation and Wang (1999) (Uden et al., 2014, pp. 6, 7). As mentioned previously, however, place names in particular are often not explained in context (Nagy & Anderson, 1984). Because the proper names have been removed from Uden et al.'s (2014) lexical analysis, the percentage of names in the texts is not observable. The participants in the study were Lithuanian and Polish; both of those languages employ an initial capital letter on names, so it is quite likely these participants are sensitive to the initial capital letter on names. However, unlike the participants in the Uden et al.'s (2014) study, not all L2 readers have an L1 with the same orthography or script as English. Because L1 transfer effects are expected to impact L2 processing, this transfer would limit the generalisations that can be drawn from Uden et al.'s (2014) conclusion, that moving from graded readers to unsimplified novels is quite easy for the L2 reader. Effects of L1 transfer to L2 processing will be discussed in further detail in Chapter 6.

A few publications appear in the 2010s that begin to question the assumptions made concerning L2 learners' knowledge of proper names. It should be noted that in the papers reviewed above, the researchers do not state unequivocally that L2 readers can understand the meaning of proper names, and they are careful to hedge arguments concerning proper names. For example, Nation (2009), in his discussion of a vocabulary profile of an applied linguistics text, remarks that his "study assumes that most proper nouns do not need to be supported and can be included in words already known. . . . [This is a] rather shaky assumption . . . [but] not unreasonable" (p. 106). A rare mention in the literature that names may present a potential difficulty for L2 readers is found in Macalister (2010), in which he presents a preliminary study into the effects of a speed-reading course on reading authentic texts. Macalister (2010) reports that teachers in his study were instructed to conduct pre-reading activities in order to introduce any potentially problematic proper names (p. 109). However, he gives no reason why some proper names might be problematic for L2 readers. Another reference made about the potential difficulty of proper names appears in Nation and Webb's (2011) handbook for vocabulary researchers. They present an argument against assuming proper names are low-burden items: L1 readers possess knowledge surrounding names, such as which names refer to females, which refer to family names, and so on, and that this knowledge contributes to text comprehension. L2 readers, on the other hand, may not be privy to such knowledge, and thus may be burdened by unfamiliar names.

A paper that discusses in much more detail the problem of how proper names are treated in vocabulary research is Brown (2010). In his review of studies involving vocabulary coverage counts, he makes a strong case against treating proper names as known vocabulary. He notes that L2 readers have less efficient lower-level processing skills than L1 readers. Thus, when L2 readers encounter unfamiliar names, this may result in burdening the working memory and may impede overall comprehension. He also points out that while proper names may not have meaning in the usual sense as other words, L2 readers might not be privy to the connotations and associations of names. These connotations and associations are examples of Van Langendonck's (2007) presuppositional meaning that names can have (as discussed in Chapter 2, section 2.2.2). Furthermore, Brown (2010) reports that some proper names share the same form as common words, including 18 items from the first 1,000 word band, (e.g. *Young* and *White*), which could result in confusion (p. 357). Finally, he suggests that mixed proper names, for example, those consisting of a title plus name (e.g. *Governor, Saint*), may present difficulties if the status conferred by those titles is not understood (Brown, 2010, p. 357).

Only one study (Kobeleva, 2012) (that this author is aware of) directly investigated the effect of unfamiliar proper names on comprehension, though that study focused on L2 listening comprehension, not reading. Nonetheless, her findings suggest interesting implications for L2 reading. She compared the 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. However, the effect was only seen in listening for details, not global comprehension. Nonetheless, the results suggest that unfamiliar names can hinder listening comprehension. Furthermore, participants in the names known condition rated the tasks as easy to do and self-reported higher comprehension; those in the names unknown condition rated the tasks as difficult and self-reported lower comprehension. These findings suggest there is an effect of proper names on affective factors.

4.2.1 Summary

In this review, it was shown that the assumption exists in L2 vocabulary research literature that proper names are easily recognised and understood by L2 readers (e.g. Hirsh & Nation, 1992; Horst, 2013; Hu & Nation, 2000; Hwang & Nation, 1989; Nation, 2006; Webb & Macalister, 2013; Webb & Rodgers, 2009). Because of this assumption, it has become standard practice to treat proper names as known vocabulary in lexical analyses of texts, by reassigning proper names to the first 1,000-word band or removing them from analyses altogether (D. Brown, 2010). The main claims for the assumption are that context and the initial capital letter on the name will signal to the L2 reader that the item is a name.

However, this assumption warrants empirical investigation for several reasons. First, proper names are not always explained in context (Nagy & Anderson, 1984). Also, L2 readers are not as competent at using context to infer meaning as it sometimes assumed (Nassaji, 2003b). Furthermore, it should not be presumed that L2 readers are competent in lower-level processing skills (Nassaji, 2014) necessary for consistent and accurate identification of letters and words in continuous text. As was noted in Chapter 2, L2 reading is more complex than L1 reading because of dual language involvement, and L1 transfer effects are expected in L2 reading (Koda, 1996, 2005). Therefore, more processing difficulty is expected for readers whose L1 orthography and/or writing system differs from the L2 (Alderson, 2000). For these reasons, the assumption that L2 readers can recognise and understand proper names in context should be investigated. As a part of that investigation, the following study is presented which looks at how L2 readers handle proper names in continuous text.

4.3 Study: How L2 readers approach proper names in different texts

In order to investigate an assumption found in L2 vocabulary research that proper names are known or understood to some degree by L2 readers, this study looks at how Japanese L2 readers of English handle proper names while reading. The study draws on the methodology used by Carver (1994), in which he gauged reading text difficulty by the percentage of unknown vocabulary. Carver (1994) had L1 English participants (elementary school and university students) circle unknown vocabulary in reading texts of various levels (see section 4.2 for more details on Carver's (1994) study). I have

adapted Carver's procedure in order to explore the how L2 English participants treat proper names in reading texts.

The primary aim of the study is to investigate if intermediate Japanese L2 readers of English identify any proper names as unknown vocabulary in reading texts of varying difficulty. Whether the difficulty level of the text affects L2 name processing is a target variable. The rationale that text difficulty may affect L2 name processing was drawn from studies which have shown that at least 95% of words in a text need to be known for the meaning of unknown words to be inferred from context (Laufer, 1989; Liu & Nation, 1985; Nassaji, 2003b). Therefore, easy texts, in which the reader knows all vocabulary, might facilitate the processing of proper names; conversely, difficult texts, in which the reader is unfamiliar with a large percentage of vocabulary, proper names might disrupt lower-level processing with respect to lexical access. Pulido (2009) explains that a text is more comprehensible when readers know most of the vocabulary, and this results in a higher probability for successful lexical inferencing. As Carver (1994) found with L1 readers, even small differences in percentages of unknown vocabulary can render a text difficult (i.e. 1% unknown vocabulary renders a text appropriate, while 2% unknown vocabulary makes a text difficult). Thus, in a similar way that unfamiliar words affect reading level appropriateness and chances for successful lexical inferencing, the percentage of unknown words might also affect how L2 proper names are processed in a text.

A secondary aim of this study is to investigate whether Japanese L2 readers of English treat proper names as unknown vocabulary to look up in a reference source. If they do list proper names to look up, this would indicate that the L2 readers are not only unfamiliar with the items, but also that they consider understanding proper name referents necessary for text comprehension. Whether the category of proper name (e.g. personal name, family name, place name, etc.) affects whether participants list a proper name to look up is also of interest. As noted above, according to Nagy and Anderson (1984), place names are rarely explained in context. For that reason, this study looks at whether a particular category of proper names, such as geographical names, proves problematic for these participants.

The study was designed to answer the following two research questions. The first question is:

> 1. Do intermediate Japanese L2 readers of English mark any proper names as unknown in reading texts, and if so, does the difficulty level of the text affect how many proper names they mark?

It is expected that some participants will mark some proper names as unknown vocabulary, despite an assumption that L2 readers can understand proper names (see review above section 4.2). The reasoning for this prediction is that these L2 readers might not have efficient lower-level processing skills to correctly identify and infer proper name referents from context. Just as Carver's (1994) L1 readers marked proper names along with other unknown vocabulary, these L2 readers might also identify some proper names as unknown vocabulary. Furthermore, it is predicted that as text difficulty increases, more proper names will be circled, on the basis that at least 95% of words should be known to infer unfamiliar lexical items using context (Laufer, 1989; Liu & Nation, 1985).

The second research question is:

> 2. Do these readers treat proper names as vocabulary to check in a dictionary or other reference source, and if so, does the type of proper name affect whether they list the proper name to look up?

On the basis that not all names are explained in context (Nagy & Anderson, 1984), and that L2 readers are not as proficient at using context to infer word meaning as is sometimes assumed (Hu & Nassaji, 2014; Nassaji, 2003b), it is expected that some participants will list some proper names to look up. As for the type of proper names that readers might list, this aspect of the study is exploratory in nature.

4.3.1 Participants

The participants (N = 49; 33 females, 16 males) were all first-year university Japanese intermediate learners of English (average age 18 to 19). They were in a four-year university degree program (English majors), with a minimum of 8 hours/week of academic English classes. Their average TOEFL score was 455, which places them at

the cusp of A2/B1 level on the CEFR. Their mean average vocabulary size was 3,750 on the Vocabulary Size Test (Nation & Beglar, 2007) to the 6,000-word band (range 1,800; median 3,700).

4.3.2 Materials

Three expository texts were selected from academic reading course books for the purpose of ecological validity; that is, the experiment aimed to utilise a setting, stimulus and task that the participants were familiar with (Schmuckler, 2001). The texts were chosen for having a large number of proper names: 9% proper name tokens. To compare, graded readers usually have between 3 and 5.5% proper names; newspaper articles can have between 4.5 to 6% proper names (Nation 2006). So these texts were heavy with proper names. While the three texts were very similar in length, some words had to be deleted so that all the three texts ran between 550 and 552 words. They were also modified to create three levels of difficulty based on percentage of vocabulary considered known. This modification is described here. Based on the vocabulary size test results (see Participants 4.3.1 above), it was estimated these participants would have a good understanding of vocabulary to the 3,000-word (3K) level. Therefore, the vocabulary coverage for each text was determined from the 3K-level (see more below on vocabulary coverage for each text). Lexical profiling of the texts was done using the vocabulary profiler VP-Compleat, using the BNC-COCA 1-25k framework (Cobb, n.d.). Vocabprofile is an online tool that allows researchers and teachers to generate a vocabulary profile of a reading text: the output of the profile displays the words, and the percentage of words, at each 1,000-word level of the BNC-COCA 1-25K. Modifications to words beyond the 3K level were made as needed to generate the three levels of difficulty among the texts. That is, words that were beyond the 3K level were replaced with synonyms from the 1K to 3K bands. These modifications were carried out so that there were: only 2% of vocabulary items beyond the 3K list for an easy text; 5% of items beyond the 3K list for a moderately difficult text, and 10% of items beyond the 3K list for a difficult text. Proper names were ignored during the lexical profiling.

The text labelled *Easy* was a factual text describing the history of cars. It had 98.2% vocabulary coverage at the 3K level; thus, it was assumed that participants would understand 98% of the vocabulary in this text. The *Moderate* text described the history of Typhoid Mary, the first healthy carrier of the disease. The vocabulary coverage was

95.2% at the 3K level; thus, approximately 5% of the vocabulary would be unknown to the participants. The *Difficult* text described the history of smallpox and its eradication. The vocabulary coverage was 89.6% at the 3K level; about 10% of vocabulary would be unfamiliar to participants. It was also assumed that participants would have little or no familiarity with these three topics. However, participants' familiarity with the proper names in the texts was not pre-tested: checking their familiarity with the names prior to the reading task would draw their attention to these items, something I wanted to avoid.

After the three levels of difficulty were created, the texts were further modified to match the number of proper names, both types and tokens. Each text had 34 proper name types, and 50 to 51 proper name tokens, representing 9.1 to 9.3% of the text tokens. After the texts were matched for proper name types and tokens, the number of words was checked again and modified as necessary. (See Appendix 2.1 for the three texts; the proper names have been underlined for ease of reference, though they were not underlined in the participant version). All three texts had these types of proper names: personal; family; place; organisation; and other (e.g. brand names, proper name related items). (The breakdown of the types of proper names for each text is given in the Table 4.8 in section 4.4 Results). The data for word count, proper name types and tokens, and lexical frequency profile are summarised in Table 4.1.

Table 4.1

Comparison of three texts by level of difficulty

Text	Topic	Tokens	Types	Proper name tokens	Proper name types	Vocabulary coverage at 3K level
Easy	History of cars	551	231	50	34	98.2%
Moderate	Typhoid Mary	550	259	51	34	95.2%
Difficult	History of smallpox	552	284	51	34	89.6%

4.3.3 Procedure

Participants were given a consent form (in English L2 and Japanese L1) explaining what was involved in participating in the research, that their performance or decision to participate was not connected to their course grade, and that they could withdraw from the study at any time (see Appendix 2.2)[11]. All participants agreed to take part in the research.

Participants were randomly assigned one of the three texts (A4 page, Times New Roman 12). Each participant was given only one text due to task demand considerations: at 550 words in length, these texts were similar in length to the participants' course textbook readings (between 500 and 800 words). If the participants were given all three texts at one time, they would probably view this unfavourably as a heavy reading load. Similarly, if each participant was given one text over three sessions, task fatigue might set in, resulting in less than complete responses to the task. In comparison, Carver's (1994) participants received two texts of only 100 words each. The decision to give only one text to each participant precluded a within-subject comparison. However, considerations had to be given to the participants' course syllabus; there was not enough time for each participant to do each text.

Task instructions were given orally in English and also appeared above the text in English and Japanese. These instructions were taken from Carver (1994, p. 418) and read as follows: *Read this text slowly and carefully. As you read, circle any words you do not know (any words you have not seen before or do not know the meaning of).*

Participants were also told that they had received different texts of varying difficulty, so while some students would have an easy text with few or no words to circle, others would have difficult texts with many words to circle. This instruction also followed Carver (1994, p. 418). Participants were given as much time as they needed to read through the text and complete the task, though most were finished within five minutes. Then a second A4 page was handed out with these instructions (orally in English, and written in English and Japanese): *Choosing from the words you circled, put the words in order that you would look them up in a dictionary or on the Internet. Put the word that you would look up first in space 1, the word you would look up next in space 2, and so on.*

[11] This procedure meets the University's ethical guidelines for doing research. Participants were told that their data would be stored securely.

There were 15 line spaces for participants to write the words they would look up. Fifteen line spaces were given to ensure participants' involvement in the task, and so that participants would have an indication of how many words to write down. Again, they were given as much time as needed to complete the task; most were finished within five minutes. Both sets of papers were collected.

In the next class meeting, participants were given a debriefing form, thanking them for their participation, explaining the purpose of the research and reminding them they could withdraw at any time without penalty (see Appendix 2.3). If participants expressed a wish to withdraw, then the researcher would remove their responses to the task from the data set.

4.3.4 Pilot study

A pilot study was carried out with a small sample (N = 23; four males, 19 females) to identify and then resolve any unforeseen problems with the experiment design relating to materials and procedure. The participants in the pilot study were different from the participants in the main study, though they were similar in terms of proficiency levels. The pilot study participants were first-year Japanese university students (18-19 years old) in a junior college (two-year) English program. They were an intact class, which had been streamlined using the Global Test of English Communication (GTEC, a proficiency test used in Japan). The average TOEFL score of the participants was 428 (A2 level on the CEFR). Their mean average vocabulary size on the Vocabulary Size Test (Nation & Beglar, 2007), taken up to the 6,000-word level, was 3,200.

The pilot study was carried out using the materials (section 4.3.2) and the procedure (section 4.3.3) described above. The results from the pilot study were analysed. Findings from the first part of the data analysis resulted in two changes to the procedure. The first part of the data analysis involved counting words which participants had marked as unknown, excluding names. How many words the participants marked as unknown was an important indication of whether the texts differed according the difficulty levels as predicted by the lexical profile. The results from this analysis would indicate if the participants were circling all unknown vocabulary (i.e. there should be more words circled as text difficulty increases). The pilot study results did match the predictions for unknown words from the lexical profile for the Easy and Moderate texts (2% and 5% unknown words, respectively). For the

Difficult text, however, participants marked fewer words than predicted (4% unknown words vs. the predicted 10%), and fewer words than the Moderate text.

It may have been that participants did not circle all unknown words in the Difficult text simply because there were so many. The Difficult text had 31 lines of text and an average 18 words per line; 10% unknown vocabulary would translate into almost two unknown words on every line of text. In this regard, it may have been a demotivating task for participants to complete. However, the study aims to capture an accurate picture of participants' lexical knowledge as it relates to text difficulty. If it was the case that participants had not indicated all unknown vocabulary, then it may have also been that they did not mark all problematic proper names. So, in order to encourage a more complete response from participants, two revisions were made to the materials and procedure in the study.

The first revision was an addition to the English and Japanese instructions for Task 1: *Any words you do not circle may appear on a vocabulary quiz next week.* Alerting the participants to a vocabulary quiz was meant to serve as added incentive for participants to consider how well they knew the words in the text. The participants were regularly given vocabulary quizzes based on reading texts in their course, and in this regard, the added instruction was ecologically valid. For reasons of research ethics, a vocabulary quiz based on unmarked words was carried out after the debriefing session, though the results were not part of the data collection.

The second revision was the addition of a practice passage for all participants to do before the main experiment. This would ensure that all participants fully understood the instructions. Carver (1994) also had a practice passage. He eliminated data from students who did not circle three target (i.e. unfamiliar) vocabulary items in the practice passage, his rationale being that these students had not understood the task instructions. Following Carver (1994) then, a short practice passage was devised.

The practice passage was created to familiarise participants with the task instructions; it also served as a check that participants had understood the instructions. The practice passage fit on one A4 page. It consisted of Task 1 instructions and a short reading passage (136 words), followed by Task 2 instructions with line spaces for three unfamiliar words. Only three line spaces were provided because the practice passage was shorter in length than the main task (see Appendix 2.4 for the practice passage). Three vocabulary items were selected as targets that the participants would most likely

not be familiar with: *snout* (10K band), *predisposed* (7K band), and *palaeontologist* (8K band). There were also proper names in the practice passage: *Mietje Germonpre, the Royal Belgian Institute of Natural Sciences*, and *Journal of Archaeological Science*.

Revisions to the Procedure (as described in section 4.3.3) can be summarised as follows. After the consent forms were collected, the practice passage was distributed to familiarise participants with the instructions before the main experiment. Participants were monitored to ensure they had understood instructions, and had circled and added at least three words to the practice look-up list. Participants were given approximately five minutes to complete the practice, and then possible responses were checked orally as a group to ensure instructions were clear. The practice passages were collected. Then, the procedure was followed as described above in section 4.3.3 (i.e. the main experiment was carried out; the debriefing session followed a week later).

4.3.5 Data analysis

From Task 1, in which participants circled unknown vocabulary items within the text, each participant's responses were analysed, beginning with a count of any words that were circled, excluding names. While the counts of unknown words excluding proper names do not address the research questions directly, analysis of this data was important to ascertain whether the texts differed in lexical difficulty, as predicted by the lexical profile. The results would also indicate whether participants had circled all unknown words. If a word was circled more than once, it was counted only once (i.e. types were counted, not tokens). Hyphenated words were counted as one word. This gave each participant a 'count of unknown words'.

Then, for each participant, a count was taken of any proper names that had been circled. Each capitalised word in a proper name was counted as one name. For example, *Mary Mallon* counts as two words; *the World Health Assembly* as three words. The rationale was that even though the name refers to one person or entity, a participant might circle only the unfamiliar part of the name. (Similarly, Hu and Nation (2000) also counted each word in a name). This resulted in a 'count of unknown names' for each participant.

From Task 2, in which participants listed words in the order that they would look them up in a dictionary or other source, a count was kept of the number of proper names

listed by each participant. This gave each participant a 'count of unknown names to look up'. Also, a list was kept of the proper names listed in order to identify unknown names across participants.

Descriptive statistics were calculated for counts of unknown words, counts of unknown names and counts of unknown names to look up. Statistical analysis involved Chi-square test of independence to assess the relationship between unknown words (types) and text difficulty; and unknown proper names (types) and text difficulty.

4.4 Results

Descriptive and inferential statistics are presented to address research question 1 (whether Japanese readers of L2 English mark proper names as unknown vocabulary with text difficulty as a variable). Descriptive statistics are presented to address research question 2 (whether these readers list proper names as items to look up).

4.4.1 Counts of unknown words

Counts were taken of words marked unknown (excluding proper names) by each participant to check how closely the mean unknown words matched the predicted unknown words from the lexical profile (i.e. for the Easy text, 2% predicted unknown words; Moderate text, 5% unknown; Difficult text, 10% unknown). This comparison of the counts to the predictions was done to check whether the texts did indeed vary in lexical difficulty (that is, more words should be identified as unknown as text difficulty increases). The results would also indicate whether participants were in fact marking all unknown vocabulary, as per the task instructions. Table 4.2 shows the descriptive statistics for counts of unknown words by text.

Table 4.2

Counts of unknown words by text: descriptive statistics

Text	*n*	*M*	*SD*	Minimum	Maximum	Range	Skew
Easy	16	1.8	1.2	0	4	4	0.41
Moderate	17	8.2	4.7	1	20	19	0.96
Difficult	16	16.6	5.4	6	29	23	0.22

Note. n refers to the number of participants who were given each text.

The figures in the mean column show that the counts of unknown words is below the predicted levels from the vocabulary profile: for the Easy text, the mean count represents less than 1% of total types; for the Moderate text, about 3%; and for the Difficult text, less than 6%. The mean counts do demonstrate a spread in lexical difficulty across the texts, however. The figures in the maximum column are much closer to the predicted levels, suggesting that at least for some participants, the predicted text difficulty was close to accurate.

A Chi-square test of independence was calculated to test whether the three texts differed with regard to the quantity of unknown word types. A contingency table for these data is shown in Table 4.3.

Table 4.3

Contingency table: known and unknown words by text difficulty

	Easy Text	Moderate Text	Difficult Text
Known word types	4207	4829	4822
Unknown word types	33	152	266

A significant dependence between word type status (known/unknown) and text difficulty was found (χ^2 = 150.26 (2), p < .05), with an effect size of V = .10, which is a

small effect size. This indicates the texts did vary by the number of unknown words, as marked by the participants.

4.4.2 Counts of unknown names

Counts were taken of proper names that were marked as unknown by each participant. This was done to answer research question 1 (whether Japanese readers of L2 English mark proper names as unknown vocabulary with text difficulty as a variable). The descriptive statistics are presented in Table 4.4.

Table 4.4

Counts of unknown names: descriptive statistics

Text	*n*	*M*	*SD*	Minimum	Maximum	Range	Skew
Easy	16	0.3	0.6	0	2	2	2.38
Moderate	17	0.8	0.7	0	2	2	0.29
Difficult	16	0.1	0.3	0	1	1	4.0

Note. n refers to the number of participants who were given each text.

The mean column shows that some proper names in each text were marked as unknown. The figures in the mean column are small; however, note that proper names made up between 12 and 15% of total types in each text (there were 34 proper name types in each text). Thus, there were fewer proper names for participants to circle as unknown than other lexical items.

A Chi-square test of independence was calculated to test if the proportion of proper names marked, relative to the other words marked, depends on text difficulty. A contingency table for these data is shown in Table 4.5.

Table 4.5

Contingency table: unknown proper names vs. unknown non-names by text

	Easy Text	Moderate Text	Difficult Text
Proper names unknown	4	13	1
Non-names unknown	29	139	265

A significant dependence was found (χ^2 = 23.02 (2), p < .05), with an effect size of V = .23, which is between a small and medium effect size. This indicates that text difficulty and unknown word type are not independent.

4.4.3 Counts of proper names on look-up lists

Counts of proper names that participants had listed to look up in a reference source were taken in order to address research question 2 (whether these readers list proper names as items to look up). Text difficulty was not a variable, so data across all three texts was combined. The descriptive statistics are shown in Table 4.6.

Table 4.6

Counts of names on look-up lists: descriptive statistics

N	*M*	*SD*	Minimum	Maximum	Range	Skew
49	0.4	0.6	0	2	2	1.38

Note. N refers to the total number of participants.

While the figure in the mean column is small, the data shows that for some participants, some proper names were problematic enough to look up. Regarding how many individual participants listed proper names to look up, a tally of the names on the look-up lists from all texts was taken. Table 4.7 summarises this information.

Table 4.7

Tally of proper names on look-up lists

N	Names listed (tokens)	Names listed (types)	By no. of participants
49	17	7	15

Note. N refers to the total number of participants.

The percentage of participants who listed at least one name to look up was 30.6%, nearly a third.

Research question 2 also looked at whether the type of proper name affects whether participants list the item to look up. For the Easy text, these names were listed to look-up (the number in brackets following the name indicates by how many participants): *Oldsmobile* (2); *Tesla* (1) and *Cugnot* (1). In the Moderate text, these names were listed: *Typhoid (Mary)* (9)*; Irish* (2); and *Thompsons* (1). In the Difficult text, one name was listed: *Assembly* (1). In Table 4.8, the proper names are shown that were listed by participants in both the pilot study and main study (N = 72). Data from both studies are compiled in this table because the interest is qualitative in nature, not numerical. The proper names are organised according to category (personal names, family names, etc.) and text.

Table 4.8

Proper names listed to look up in pilot and main study (N = 72)

	Easy Text		Moderate Text		Difficult Text	
Proper Name Category	Types	Types listed to look up	Types	Types listed to look up	Types	Types listed to look up
Personal	8	*Gottlieb*	5		5	
Family	9	*Cugnot* *Daimler*	5	*the Warrens* *the Thompsons*	6	*Wortley*
Place	9		15		9	
Organisations	4		4		9	*World Health Assembly*
Other (brands, epithets, proper name related)	4	*the Oldsmobile* *Tesla car*	5	*Typhoid Mary* *Irish*	5	
Total	34	5	34	4	34	2

Note. Type = the total number of types from that category of proper name in each text.

Places names were the only type of proper name not listed for any of the texts. Family names were listed for all three texts. Note that in the Moderate text, the family names are in non-prototypical form (i.e. in the plural form with the definite article). The Other column shows other non-prototypical names that participants listed: *the Oldsmobile* (brand name), *Typhoid Mary* (epithet) and, *Tesla* and *Irish (*proper name related terms).

4.5 Discussion

The first aspect of participants' task responses that I looked at was how many words they marked as unknown, excluding proper names. While this does not address the research questions, I was interested in these counts for two reasons. First, because text difficulty was a dependant variable, I wanted to check if the texts had in fact differed according to the participants' vocabulary knowledge, as predicted by the vocabulary profile. Second, I wanted an indication of whether participants were marking all unknown vocabulary. As was noted in the pilot study (section 4.3.4), those participants had marked fewer unknown words in the Difficult text than the Moderate text. Adjustments had been made to the procedure to encourage participants to mark all unfamiliar vocabulary.

The results from the main study indicated that the texts did vary in difficulty for these participants. Mean counts showed that more words were circled for the Moderate text than the Easy text; and more words were circled for the Difficult text than the Moderate one. A statistically significant interaction was seen between the three texts with regard to the number of unknown words, though the effect size was small. As for whether participants were marking most unknown words, this seems likely given that the maximum counts matched the predicted levels, as well as the spread in lexical difficulty seen in the mean counts for the three texts.

The first research question focused on whether intermediate level Japanese L2 readers of English identify proper names as unknown in texts of varying difficulty. It is important to know how L2 readers treat proper names because an assumption in the L2 vocabulary literature is that L2 readers do understand proper names and therefore, can be treated as known vocabulary. Contrary to this assumption, the results in this study demonstrated that some of these readers do treat proper names as unknown vocabulary. While the participants did not identify unknown proper names at the same frequency as other vocabulary items, proper names also did not occur in these texts at the same frequency. Thus, while the counts of unknown proper names were small, it is important to remember that there were fewer proper names (between 12 - 15% of types) compared to other lexis.

A significant dependence was seen in the proportion of proper names marked unknown relative to the other words marked depending on text difficulty. Participants

marked the fewest proper names as unknown in the Difficult text. This was an unexpected finding: it was predicted that as the number of unknown words increased, more proper names would be marked as unknown. This was predicted on the basis that at least 95% percentage of vocabulary should be known in order to facilitate the use of context to infer meaning. One possible explanation for this finding is that with a larger percentage of unknown vocabulary, L2 readers have fewer attentional resources to give to proper names; conversely, as the percentage of known vocabulary increases, readers give more attention to unfamiliar proper names. A within-subject comparison may have shed more light on this phenomenon, and points to a limitation in the study design.

The second research question looked at whether participants listed proper names as vocabulary to check in a reference source. Investigating look-up behaviour was an important distinction to make from research question 1: an L2 reader may recognise that an unfamiliar word is a proper name and yet not feel the need to look it up in a dictionary. Looking up words in a dictionary takes valuable time and attention away from the text; for that reason, L2 readers try to be selective about which words they are going to bother looking up (Hulstijn, 1993). Indeed, learning to use context to correctly infer word meaning is a reading strategy often taught and practiced in L2 reading classrooms; this strategy helps avoid an overreliance on dictionary look-ups and time away from the text. Therefore, if the readers in this study did list proper names to look up, then this would indicate that context was not helpful in determining the referent of the proper name. Furthermore, it would indicate that the reader considered the proper names important enough to text comprehension to look up in a dictionary.

In this study, all the participants who marked a proper name as unknown in the text also listed that name to look up. That is noteworthy because it raises the possibility that there may have been more proper names in these texts that were unknown to the participants than the ones that were marked. Perhaps the participants were unfamiliar with some of the proper names but because they had no intention of looking up these items, they did not mark them as unknown. Only the proper names they did mark as unknown were listed to look up.

Nearly a third of participants (30.6%) listed at least one proper name to check along with other vocabulary. In placing a proper name on the look-up list, the participants indicated that understanding that particular proper name referent was important to comprehension of the text. Participants listed the fewest proper names to look up for

the Difficult text. Again, perhaps this was because when there is more unknown vocabulary, less attention is given to proper names. At any rate, given that nearly a third of participants listed a proper name to look up suggests that they were not unburdened by the proper names in the texts.

Research question 2 also explored whether the category of proper name affects whether the item is listed to look up. It would be useful for L2 teachers and test writers, for example, to be aware of what kinds of proper names are problematic. In this study, most of the proper names listed to look up were in non-prototypical form. For example, proper names listed included family names in the plural form with definite article (e.g. *the Thompsons*), and proper name related items (e.g. *Irish*). Interestingly, no place names were listed for any of the texts even though as Nagy and Anderson (1984) noted, geographical names are often not explained in context, and thus might be a source of confusion for readers. That was not seen here. However, this observation points to another limitation of this study. Only three texts were used, and therefore, generalisations cannot be drawn from the proper names associated with only these three texts. The advantage in using three texts meant that there was a control on which proper names participants encountered. The disadvantage was that this left a small pool of proper names from which it is difficult to draw any conclusions.

Another issue related to interpreting the results from this study concerns participants' familiarity with the proper names. It was assumed that participants would have little or no background knowledge of the reading topics. However, directly checking participants' familiarity with the proper names would have alerted their attention to the items, something I wanted to avoid. A possible workaround for this problem is to gauge participants' familiarity with proper names as an aspect of cultural background knowledge. That is, one might assume that L2 readers will be more familiar with proper names from their own culture than with proper names from the target language culture.

4.6 Conclusion

Much of the L2 vocabulary literature assumes that L2 readers understand names. It has become standard practice for researchers to treat proper names as known vocabulary, or to even remove them altogether from lexical analyses. Only one study (Kobeleva, 2012) investigated this assumption, looking at proper name processing for L2 listeners. As was noted above (section 4.2), Kobeleva (2012) found that L2 listeners often mistake proper names for common words, and unfamiliar names can cause

anxiety for L2 listeners. Her findings suggest implications for proper name processing by L2 readers as well. No other studies (that I am aware of) have directly investigated how L2 readers handle unfamiliar proper names.

The findings from this study demonstrate that Japanese L2 readers of English do identify proper names as unknown vocabulary. This finding contradicts the assumption that L2 readers understand all proper names. Participants in this study identified more unknown names for the Easy and Moderate texts than for the Difficult text. The participants also listed proper names as items to look up, suggesting that for some readers at least, proper names are considered important to text comprehension, important enough to look up in a secondary resource.

Several questions arise from this study. The first concerns the unexpected finding that participants listed more proper names to look up in the easier texts than in the difficult text. It was predicted that in texts with more words unknown, L2 name processing would be more difficult. The finding suggests that L2 readers tend to ignore proper names in more difficult texts. A study design that included within-subject comparisons would help to investigate this further. Secondly, based on the low number of proper names across the three texts, it is hard to draw conclusions on what kind of proper names cause difficulties for L2 learners. However, it would be informative for L2 reading teachers and test writers, for example, to know what kind of proper names are problematic for L2 readers. Lastly, one aspect of the participants' knowledge that this study did not control for was their familiarity with proper names. That is, it was not known how many of the proper names in the text were already known to the participants; this may have impacted the results. To explore the effect of familiarity, proper names will be considered as an aspect of cultural knowledge. This will be the focus of Chapter 5: how cultural familiarity with proper names can aid comprehension, and conversely, how cultural unfamiliarity with names might hinder comprehension.

Chapter summary

This chapter looked at how Japanese L2 readers of English handle proper names. From the literature review, it emerged that no empirical studies have been carried out to check the validity of an assumption that L2 readers understand proper names. Rationale was provided for checking this assumption: L2 readers might not have efficient lower-level processing skills to correctly recognise letters and words in continuous text, in particular if the L1 orthography or writing system differs from the L2.

Also, L2 readers may not be as efficient at using context to infer meaning as is sometimes assumed. Moreover, proper names are not always explained in context, and thus may present a burden to the L2 reader.

A study was carried out to investigate how Japanese L2 readers handle proper names in texts of varying difficulty. The participants did identify English proper names as unknown vocabulary, and a significant dependence was seen between the number of proper names marked as unknown and text difficulty. Furthermore, nearly a third of the participants listed proper names as items to check in a dictionary. These results do not support the assumption that L2 readers can understand proper names in texts. It seems that for some readers, proper names do present a reading burden.

In the next chapter, proper names are considered as an aspect of cultural knowledge. Several studies are reviewed that have investigated the effect of cultural background knowledge on reading comprehension. The approaches taken in those studies can be used to investigate how culturally familiar proper names affect reading comprehension.

Chapter 5: The effect of culturally familiar proper names

5.1 Introduction

The study in the previous chapter demonstrated that proper names are not always understood in context by some L2 readers. It was found that some L2 readers mark unfamiliar proper names as vocabulary to look up in a dictionary, indicating that proper names can interfere with text comprehension. Understanding more about what makes some proper names a reading strain is important for L2 reading teachers, and developers of L2 materials and tests. For example, learner dictionaries are an important resource for L2 users. Resources like dictionaries could be improved if more was known about why some proper names can disrupt reading comprehension. To unpack this question of what makes some proper names problematic, one approach is to investigate proper names as an aspect of cultural background knowledge. The aim of this chapter is to look at whether cultural familiarity of proper names aids comprehension, and conversely, whether unfamiliarity with proper names hinders comprehension. By looking at proper names with cultural familiarity as a variable, conclusions might be drawn about the potential difficulties for L2 readers.

This chapter follows the same structure as the previous chapter. A brief overview is given of relevant literature that directly informs the content of the experimental work. Then, two experiments are reported. The first study compares reading comprehension when the proper names are culturally familiar (i.e. L1 names) and unfamiliar (i.e. L2 names). The second study addresses issues found with the first study, and compares reading comprehension among three treatments: familiar names (i.e. from the L1 culture), unfamiliar names (i.e. from the L2 culture) and no names (i.e. only common noun referents). The results from both studies provide insight into how proper names might affect reading comprehension with cultural familiarity as a variable.

5.2 Cultural knowledge and its effect on reading comprehension

The cultural background knowledge that the reader brings to a text is considered an aspect of higher-level processing; that is, the reader uses her background knowledge to interact with the text to build comprehension. In order to learn more how cultural familiarity of proper names might affect comprehension, research is reviewed which

has examined the role of cultural background knowledge on comprehension. In particular, four studies are reviewed in detail: these four studies were selected in respect of their robust design, and their extension of earlier research into the effect of cultural knowledge. I will begin this section with a brief overview of schema theory because it is the dominant theory running through these four studies. Then, I will look at each study in turn, in terms of methodology and implications of the findings for proper name processing. The first two studies, from the 1970s and 1980s, were among the earliest to investigate how cultural knowledge contributes to reading comprehension. Extending this early research, the other two are more current studies from the late 2000s, the latter of which was considered a good model for replication to investigate the effect of culturally familiar proper names on comprehension. This section will conclude with various critiques of schema theory as it has been applied to reading research.

5.2.1 Schema theory

Schema theory attempts to account for knowledge acquisition and reading comprehension through activation of schemata: networks of information stored in the brain (Alderson, 2000, p. 17). The theory asserts that in order for new information to be learned, it must be connected to existing knowledge (Bartlett, 1932; D. E. Rumelhart, 1980). R. C. Anderson and Pearson (1984) presented an influential application of schema theory to reading processes. As applied to reading comprehension, schema theory assumes text comprehension is an interactive process between the reader's background knowledge of content and text structure, and the text itself (Carrell, 1983, p. 82). Schemata, often described as related sets of knowledge connected in a frame, are triggered when a text activates a particular concept. It is thought that schemata are used to aid interpretation and make inferences in support of comprehension. Grabe (2009) explains that schema theory was especially prevalent in 1980s and 1990s to explain the role of generalised background knowledge in reading.

Different subcategories of schemata are identified in the literature. Carrell (1983) distinguishes between content schemata, the knowledge related to the content area of a text, and formal schemata, the knowledge of text organisation and genres. She cautions against confounding the two types. For example, understanding how academic journal articles are structured is an example of formal schemata, while having knowledge related to the specific content of such an article is an example of content schemata. Alderson (2000) further separates content schemata into

background knowledge, which may or may not be relevant to the text, and subject-matter knowledge, which is directly relevant to the text (p. 34). For Alderson (2000), cultural knowledge is considered a particular aspect of background knowledge, and as such, may or may not be directly relevant to understanding a text. Alderson (2000) notes that because background knowledge tends be idiosyncratic, based on one's own experiences, it can be difficult to predict or control for, in this regard (p. 45). Grabe (2009), who avoids the term 'schema', takes a slightly different approach, dividing background knowledge into four subcategories: general world knowledge, cultural knowledge, topical knowledge and specialist knowledge (p. 74).

My interest here is with proper names as an aspect of cultural background knowledge and how this knowledge affects text comprehension. As Hanks (2013) points out, L1 users frequently exploit shared cultural knowledge of proper names, by using a famous name in place of a generic term. He illustrates with an example from a news article title, in which a detective is referred to as a *Sherlock Holmes* (Hanks, 2013, pp. 34, 35). In this way, familiarity with proper names can be regarded as an aspect of cultural knowledge. For insight into how the effect of cultural familiarity with proper names might be investigated, I review in detail four papers below, focusing on the approaches taken to investigate how cultural knowledge impacts comprehension. I also focus on the studies' findings for implications related to proper name processing. Because all these studies refer to schema theory to underpin their investigations, it is necessary to describe schema theory as it has been used to explore the effect of cultural knowledge on reading comprehension. I conclude by discussing the efficacy of schema theory to explain the effect of cultural knowledge on text comprehension.

5.2.2 Steffensen, Joag-Dev, and Anderson (1979)

A paper that is frequently cited in L2 reading literature as the classic study into the effect of cultural background knowledge on reading comprehension is Steffensen et al. (1979). The authors use the term 'schemata' to denote background knowledge, and they have a specific interest in cultural schemata. Where previous research had investigated the effect of cultural knowledge on only one group with culturally unfamiliar texts, the authors' aim was to present a balanced study design with two groups. That is, two groups with different cultural heritages read two texts, each text representative of those cultures. The two groups, Americans (n = 20) and Asian Indians (n = 19) living in America, were given texts (letters) that differed in content. One letter described a traditional American wedding; the other described a traditional Indian wedding. The

texts were otherwise matched for syntactic complexity and rhetorical structure. The participants were told to read for comprehension, not speed. First, they read either the American passage or the Indian passage; the order was counterbalanced. This was followed by a short retention interval (i.e. participants did a vocabulary test, unrelated to the texts). Then, participants were instructed to write about every part of the letter they could remember. This was followed by five inferential questions about the text. The procedure was repeated with the second text. The authors measured reading time and recall. They also recorded elaborations and distortions of the texts made by participants in the written recall task.

They found that both groups read faster and recalled more information from the culturally familiar texts. Both groups elaborated more on the culturally familiar text by adding culturally appropriate information not found in the text. Both groups also distorted more information about the unfamiliar texts. The authors conclude that cultural background knowledge has a profound effect on how well a text is understood and recalled. They note, however, "the precise mechanisms responsible for the effects are not well-understood" (Steffensen et al., 1979, p. 20). That is, it is not evident whether cultural background knowledge affected higher-level or lower-level processing. Steffensen et al. (1979) give the example of the longer reading time spent by American participants on the Indian passage: They query whether this was due to a lack of cultural knowledge about Indian weddings (higher-level processing) or due to a cumulative effect of unfamiliar foreign lexical items (lower-level processing) (p. 20). They note that other than proper names, the only foreign words were *dhoti* and *sari*. They do not give details as to what the proper names were or how many were included in the text. The complete texts are not provided in the appendices; however, from other examples given in the discussion section, one can identify two names, *Prema* and *Nagpur*. From excerpts of the participants' written recall tasks, one can infer that the American readers understood *Prema* referred to a woman and *Nagpur* to a place. However, there is one excerpt in which an American participant used a blank to indicate the name of the bridegroom, indicating that either the name was not given in the text or the participant could not recall the name or spelling:

(1) A37 Prema and _____ are taking a trip to the north of India . . . (Steffensen et al., 1979, p. 24)

If the name was given in the text, then this blank is indicative of a possible cognitive burden with unfamiliar proper names, which may have contributed to longer reading

time (i.e. lower-level processing was affected). On the other hand, if the name was not given in the text, then it is difficult to infer why the participant did not use a common noun referent, such as *her husband,* instead of using a blank.

Steffensen et al. (1979) seems to be robust and as noted above, is referred to in the literature as the classic study which clearly demonstrates the effects of cultural background knowledge on reading comprehension (e.g. Alderson, 2000; Carrell, 1983; Grabe, 2009; Hudson, 2007). It should be noted that the authors do not position their paper as an L2 reading study but rather as a cross-cultural study. For example, subjects in the study are not referred to as L1 or L2 users of English; instead, they are distinguished by their nationality/heritage (i.e. American or Indian). One criticism of the paper, however, concerns the lack of details regarding the participants. Certain bio-data that is missing could be relevant to interpretation of the results. For example, we are not given any details about the Indian participants' proficiency levels, or whether they were reading in their L1, L2 or even L3. This information would have useful for interpreting the slower reading times for the Indian participants. We also do not know how long they had been living in America, only that they were residents in a university community. This information would have been helpful, assuming that more time spent in a foreign community allows for more (potential) exposure to cultural norms, for example, through television, movies, and interaction with locals. Overall, the effect of cultural familiarity on reading comprehension is demonstrated in the study. However, as the authors note, it is unclear whether the participants' lower-level or higher-level processing was affected. Thus, while cultural background knowledge is used in higher-level processing, it may be the case the unfamiliar lexis (i.e. the culturally unfamiliar proper names and other foreign words) affected the readers' lower-level processing. As noted in Chapter 2, lower-level and higher-level processes are thought to be hierarchical; efficient lower-level processing must take place first in order for higher-level processing to occur (Nassaji, 2014).

5.2.3 Johnson (1981)

Another frequently cited study into the effects of cultural background knowledge on reading comprehension is Johnson (1981). In this study, the author investigated the effects of language complexity and cultural knowledge on reading comprehension in ESL students; specifically, she was interested in whether language complexity had a greater effect on comprehension than the cultural origin of the text. In this respect, she

was building on previous research, which had investigated the effects of language complexity and cultural knowledge on reading comprehension separately.

The sample size was 46 Iranian intermediate/advanced ESL students and 19 American university subjects, who served as a comparison. Two reading passages in English were used, one from Iranian folklore, a Mullah Nasr-el-Din story, and the other from American folklore, a Buffalo Bill story. Johnson (1981) states that the stories had "similar motifs which were culturally distinct yet were equivalent in plot construction" (p. 170). Two versions of each text were used, the original version and an adapted (simplified) version. In the two original text versions, the language complexity was balanced (i.e. the same number of relative clauses, compound and complex sentences, low-frequency vocabulary, etc.). The original and adapted versions contained the same number of propositions. Participants were randomly assigned: one group read adapted versions from their own culture and the foreign culture; the other group read the original texts. The text order was counterbalanced. Reading comprehension was tested with a written recall for which participants were not allowed to refer to the text. They were also given multiple-choice questions relating to explicit and implicit information from the texts. Written recalls were analysed by propositions (number and types) as well as connections between propositions, that were found in the participants' recalls (Johnson, 1981, p. 171).

The author found that for the Iranian (L2 English) subjects, cultural background of the text had a greater effect on reading comprehension than syntactic and semantic complexity. With the culturally familiar text (i.e. the Mullah Nasr-el-Din story), no differences were found in the recall of the original and adapted versions. Participants recalled more information from this text than the American text. For the culturally unfamiliar text (i.e. the Buffalo Bill story), more events were recalled from the adapted version. Also, more errors were made in the recall of the American text. By comparison, for the American (L1 English) subjects, reading comprehension was affected by both language complexity and cultural origin of the texts. Recall of plot was better for the culturally familiar American text. There were more errors in the recalls of both adapted versions. Recall was better for the original texts than the adapted versions.

Johnson (1981) concludes that the Iranian L2 English users drew on their cultural background knowledge for the Iranian original text to make up for what they lacked in language proficiency (p. 173). They had the most difficulty with the culturally unfamiliar

original text because they lacked both cultural background knowledge and language proficiency. That the American subjects' comprehension was better for the original texts may have been due a lack of text cohesion and general readability as a result of simplifying the text (Johnson, 1981, p. 174). The author also notes that use of cultural knowledge seemed to be more evident with inferential questions on the text (Johnson, 1981, p. 178).

Overall, the Johnson (1981) study is well designed and clearly demonstrates the contribution of cultural background knowledge to reading comprehension. However, there are two issues that should be noted. The first concerns a possible conflation of different types of background knowledge that may have been relevant to understanding the two texts. Carrell (1983) points out that even though Johnson (1981) says that the texts had similar motifs and plots, without the printed texts, it is not possible to say if the rhetorical structure was indeed the same in the two texts. That is, it is not readily apparent if the two texts employed the same organisational structure. Therefore, it is possible the investigation confounded formal schemata (i.e. knowledge of text organisation) and content schemata (i.e. cultural knowledge) (Carrell, 1983, p. 88).

A second issue concerns the lack of details regarding the L2 English learners' proficiency. Considering that one of variables in this study was the effect of language complexity on comprehension, more details regarding the participants' L2 proficiency would have been useful. We are only told that they were intermediate/advanced users of English. But without more details about their proficiency levels, it is difficult to say how challenging the original versions of the texts were for these readers. For example, Johnson (1981) notes that the L2 "subjects' vocabulary knowledge was low", explaining that on the vocabulary section of the Comprehensive English Language Test, their results were skewed toward low performance (p. 173). The author seems to suggest here that with respect to their overall language proficiency, they performed lower than expected on vocabulary compared to other sections of the test. If that was the case, this implies that comprehension difficulty may have originated with lower-level processing (i.e. at the semantic level), and not only with higher-level processing (i.e. cultural knowledge). Thus, just as Steffensen et al. (1979) note in their study, it is not readily apparent which level of text processing is being affected.

To recap, in both studies reviewed above, there is the possibility that cultural knowledge may have affected text comprehension at the lower-level processing level. Specifically, the authors acknowledge that unfamiliar lexical items in the text

(Steffensen et al., 1979) or a low level of L2 vocabulary knowledge (Johnson, 1981) may have impacted text comprehension. Thus, one of the challenges in investigating the effect of cultural familiarity on comprehension relates to determining which level of processing is being affected. Attempting to tease apart these factors, Alptekin (2006), reviewed below, investigates whether cultural background knowledge can compensate for deficiencies in lower-level processing. Perhaps drawing on Johnson's (1981) remark that effect of cultural knowledge was more evident with inferential questions, Alptekin (2006) compares the effects of cultural background knowledge on inferential and literal comprehension in L2 reading. He defines literal comprehension as that based on lower-level processing (i.e. lexical access and syntactic parsing); in contrast, inferential comprehension is based on higher-level processing (e.g. synthesising, summarising, making inferences) (Alptekin, 2006, p. 495).

5.2.4 Alptekin (2006)

In this study, Alptekin (2006) hopes to build on previous studies, including Steffensen et al. (1979) and Johnson (1981), in three ways. First, he criticises the previous research for selecting or writing texts specifically for research purposes, simplifying them both linguistically and conceptually. With such texts, readers are not able to activate what he calls "abstract schema": a "cognitive structure that enables readers to 'recreate' the writer's message by predicting the way in which the texts proceed" (Alptekin, 2006, p. 496). He writes that abstract schema is different from content or formal schema, but is "a logical and general embodiment of the two" (Alptekin, 2006, p. 496). He explains that abstract schema is commonly referred to as story schema; it does not depend on syntactic forms in the text but rather, allows for inferencing of elements from the script and text. Secondly, he challenges the notion that the texts used in previous research are in fact linguistically and rhetorically equivalent. He notes that readability formulas used to determine text equivalencies tend to ignore the knowledge that readers bring to a text. Lastly, the author criticises previous studies that used recall procedures to measure comprehension. He notes there are problems with learner comprehension-production interference, as well as scorer reliability. Furthermore, he notes that recall procedures focus on what is retained and retrieved, rather than what has been understood (Alptekin, 2006, p. 498).

Alptekin's (2006) study design addresses these issues. The sample was 98 advanced Turkish learners of English (38 female, 60 male). Their average age was 18; their average score on the TOEFL test was 565. Only one text (an American short story)

was used but in two versions: the original and the other "nativized", which is a "sociological, semantic, and pragmatic adaptation of the textual and contextual cues of the original story into the language learner's own culture, while keeping its linguistic and rhetorical content essentially intact" (Alptekin, 2006, p. 499). Textual cues included references to settings, characters, and occupations. Contextual cues included cultural customs, ideas and values. Some examples of how the text was nativised include changing: *New York City* to *Istanbul;* a *church* to a *mosque;* and *bathers* to *fishermen.* Comprehension was measured with a multiple-choice test of 24 questions, one for each version of the story: 12 explicit questions to measure literal understanding and 12 implicit to test inferential understanding. Literal versus inferential understanding was characterised using Pearson and Johnson's (1978) classification. Participants were assigned to the experimental or control group (n = 49). They were given 45 minutes to read the story and answer the 24 multiple-choice questions.

Data analysis (t-tests) showed that the group that read the original version did only slightly better on literal comprehension than the group that read the nativised version, but no statistical significance was seen. However, there was a statistically significant difference in inferential comprehension, with the experimental group doing better. The effect size reported is .61, which is moderate. The author concludes that cultural background knowledge affects inferential comprehension, rather than overall comprehension. Alptekin (2006) suggests that literal comprehension may not be affected by cultural background knowledge as it is data-driven (p. 502). Thus, the idea that an L2 reader's cultural background knowledge could compensate for "surface-level inadequacies in textual processing" (i.e. lower-level processing deficiencies) is not supported by this research (Alptekin, 2006, p. 502). He does caution the interpretation of these findings, however: the participants were advanced L2 users and some had graduated from English-medium high schools, which suggests they may have been familiar with American culture. He explains that advanced learners were used in this study so that the linguistic threshold, that is, the relationship between inferencing skills and L2 proficiency, was not an issue (Alptekin, 2006, p. 498). However, this proficiency level may have resulted in familiarity with the target culture as well.

Alptekin (2006) supports his claim of cultural knowledge effects on inferential reading comprehension with schema theory, which he says "accounts for the role of inferencing in comprehension based on the reader's prior knowledge of the topic" (p. 495). As noted above, he extends his discussion beyond content and formal schemata to include what he refers to as abstract or story schema (Alptekin, 2006, p. 496). Story

schema is described as a cognitive structure which helps readers to 're-create' the text by predicting how it will proceed; because story schema is not dependent on syntactic arrangements within the text, it is useful for inferencing (Alptekin, 2006, p. 496). Quoting Oller (1995), Alptekin (2006) explains that the interaction between reader and text can happen when certain words are changed to be more familiar for the reader (e.g. changing *Terry* to *Ali*): "a name recognized as pertaining to a male referent generates expectancies that will be absent if the name is not recognized as having any gender bias" (Oller, 1995, p. 297). Thus, Alptekin (2006) concludes, in absence of this story schema, where such referents are understood, it is difficult for L2 readers to identify with the content of a text.

Alptekin's argument concerning story schema and how it affects reading comprehension is woolly, however. To begin with, it is unclear how story schema is distinct from content or formal schema, and yet includes elements of both. His example of knowing the gender of a proper name (e.g. *Terry* and *Ali*) to illustrate story schema does not make it clear how this is different from cultural knowledge. For example, knowing that *Terry* can refer to males and females could be treated as an aspect of cultural knowledge. It might be that if this knowledge is applied to make inferences about what will happen next in a story, then it is to be considered 'story schema'. If this is Alptekin's intended meaning, then it is difficult to understand how researchers could ascertain whether L2 readers drew on story schema or cultural knowledge to make inferences. It seems that we are to deduce that if readers cannot make correct inferences about a text, then they do not have the appropriate story schema. Thus, his argument is weakened because it is circular. In summary, one can conclude along with the author, that inferential comprehension (i.e. higher-level processing) was affected by cultural familiarity; however, whether schema theory can be used to explain this effect will be further discussed below.

5.2.5 Erten and Razi (2009)

Erten and Razi (2009) investigate the effects of cultural familiarity and reading activities on overall reading comprehension. Erten and Razi (2009) are extending Alptekin's (2006) research, and their study is also informed by schema theory. The authors discuss the beneficial use of reading activities (pre-reading, while reading, and post reading) to activate schema (e.g. Chen & Graves, 1995; Grabe & Stoller, 2002). The research questions they address are: Does readers' familiarity with the cultural content of short stories affect their comprehension? Do reading activities used with short

stories make up for a lack of cultural schema? The authors take a positive stance toward the effects of both cultural familiarity and reading activities on reading comprehension: Readers understand a text better when they are familiar with cultural content, and reading activities designed to activate background knowledge improve reading comprehension. The authors hypothesise that while reading activities can aid comprehension, cultural familiarity will be the significant factor (Erten & Razi, 2009, p. 64).

Erten and Razi (2009) used a 2 x 2 experimental research design to study the effects of cultural familiarity and reading activities on reading comprehension. The participants were 44 Turkish advanced learners of English. These learners were divided into four groups of 11, and each group received a different treatment. All groups read a short story and answered comprehension questions on a recall test. The first group read an original American short story. The second group read the same story, and was also given reading activities for the story. The third group read an adjusted version of the original story: the content of the story was the same but textual and contextual elements had been changed to match the learners' Turkish culture. The fourth group read this adjusted version, and was also given reading activities for the story. The results from the recall test were analysed. Analysis of variance showed that comprehension was better with the adjusted version of the story. Comparing the groups that read without reading activities, the group that read the adjusted version had better comprehension (M = 69.91) than the group that read the original version (M = 60.45) with a considerable effect size (d = 0.81). Comparing the groups that read with reading activities, the group that read the adjusted version had better comprehensions scores (M = 79.18) than the group that read the original version (M = 64.55), with a large effect size (d = 1.45) (Erten & Razi, 2009, p. 69). Thus, while reading activities helped the learners in comprehending the story, cultural familiarity had a stronger effect.

The main claim that Erten and Razi (2009) are making is that when L2 readers are familiar with the cultural content of a text, their overall text comprehension is better; reading activities aid comprehension but the effect of cultural familiarity remains strong. They claim that when readers can identify with the people, places and other socio-cultural relations in a text, their reading burden is lessened (Erten & Razi, 2009, p. 61). Cultural concepts have different referents, which create different expectations on the part of the reader; for example, the concept of 'breakfast' will have different referents in different cultures, and thus create different expectations from the reader (p. 62). This

would certainly extend to proper names, which often carry culturally specific information. While the authors do not give the full texts in the appendices, they do offer some examples of changes they made (Erten & Razi, 2009, p. 66). Of the 16 example changes made, 11 of these were proper names: three character names, six place names, and of seven culture references, two were proper names or proper name related.

They base their claims about the positive effects of cultural familiarity on reading comprehension on other research as well as their own study. They refer to other studies that has shown the positive impact of background knowledge and cultural familiarity on comprehension, including Ketchum (2006), Oller (1995), Pulido (2003) and Steffensen et al. (1979). Thus, while the sample size from their study is small (N = 44), Erten and Razi (2009) provide sufficient evidence from other sources to support their claims. The authors express a fair degree of certainty when reporting that their findings "suggested a strong possibility that the students who read the nativized [i.e. adjusted] version of the story possessed relevant cultural background knowledge, which reduced the cognitive load imposed" (p. 70). They do qualify their claims by acknowledging limitations in the study: the study did not seek to test the effectiveness of the reading activities used; the groups might have been more homogeneous; and factors such as motivation and attitude were not taken into account. They also acknowledge the small sample size.

The authors do not indicate what contexts their findings might apply to. For example, they do not specify at what ability level an L2 learner might most benefit from cultural familiarity, though their study was done with advanced learners. The assumption seems to be that for any level of learner, cultural familiarity will aid reading comprehension. Nor do the authors make reference to whether differences between the reader's L1 and the L2 (e.g. orthography) would affect the role of cultural familiarity in reading comprehension. The assumption seems to be that their claim applies to all L2 readers, regardless of L1 or proficiency level.

As part of my central research aim to determine whether proper names are a burden for L2 readers, in this chapter, I am considering proper names as an aspect of cultural knowledge. Given that most of the adjustments of the short story in Erten and Razi (2009) concerned the proper names of characters and places, their methodology was considered applicable for studying the effects of culturally familiar proper names on comprehension. A replication study could also test whether their claims can be

generalised to other contexts, such as L2 readers at a different proficiency level or with different L1s. Before moving on to the replication study, this literature review will conclude with critiques of schema theory as it has been applied in reading research. As noted above, schema theory fails to adequately account for the findings in the studies reviewed: it is not clear whether the knowledge drawn on to aid comprehension was lexical or cultural knowledge, for example. In this regard, schema theory might not be useful to explain how cultural familiarity with proper names aids comprehension.

5.2.6 Critiques of schema theory

Criticisms of schema theory date back to Alba and Hasher (1983). Several problems have been noted with the theory, beginning with the vagueness of term 'schema', coupled with a lack of consistency in how it is defined (Alba & Hasher, 1983; Nassaji, 2002; Sadoski & Paivio, 2007; Sadoski, Paivio, & Goetz, 1991). Sadoski et al. (1991) note the unfortunate synonymous use of 'schema' with 'background knowledge': For while it is agreed that readers possess such knowledge and use it to construct meaning, "the question of how this knowledge is represented, organized, and used is a subject of considerable debate in cognitive psychology" (p. 465). Another difficulty with schema theory is the treatment of mental representations as fixed, rather than as reconstruction of information at each retrieval (Grabe, 2009).

There are specific criticisms of schema theory as it has been applied to reading comprehension. Carver (1992) argues that schema theory is only applicable to reading of difficult material for study purposes, not for regular reading, which he calls 'rauding'. Grabe (2009) notes that empirical support for the role of topical knowledge decreases when reading does not make specialist knowledge demands on the reader, or deal with cultural knowledge explicitly (p. 75). He cites Steffensen et al. (1979) and Johnson (1981) as two studies that do overtly deal with cultural knowledge, and thus do demonstrate the effect of background knowledge on comprehension. As was seen in the Alptekin (2006) study, the role of cultural knowledge was only observable with inferential questions (i.e. higher-level processing). However, as discussed above, relying on schema theory to explain the differences between literal and inferential comprehension in this study results in ambiguity.

Carrell (1983) also urges caution when interpreting the effects of cultural background knowledge on reading comprehension. She points out that content schemata can differ among members of the same cultural group and that lack of familiarity with content of

the text does not necessarily equate to culturally specific content: “Content schemata may be absent within as well as across cultures” (p. 89, 90). To illustrate, she points to studies done with American college students, which showed differences in content schemata in a reading comprehension test. Stories were used which could draw different interpretations. The studies found that the students chose the interpretation that matched their study major background. So, while they belonged to the same cultural group, their specialised background knowledge affected their interpretation of the readings. In conclusion, Carrell (1983) suggests that variables such as specialised backgrounds and age might have an effect on content schemata that is not culturally specific. Therefore, individual differences need to be considered in investigations of content schemata, including cultural background knowledge.

Grabe (2009) discusses the complexities of background knowledge and how it interacts with other factors like motivation, attitudes, goals and proficiency. He notes that exactly how background knowledge is retrieved and used in comprehension is not clear, and there is little agreement on how to measure this knowledge to ascertain its role in comprehension (Grabe, 2009, pp. 74, 75). Similarly, Hudson (2007) notes the “inherent problem with operationalizing the construct of *background knowledge*” (p. 160). Writing in 2009, Grabe states, “most current research . . . recognizes this complexity and tends to downplay specific claims for background knowledge effects or minimize its role in reading research studies and reading assessment” (2009, p.75).

Because of the various problems with schema theory, several authors advocate using the theory as a metaphor for representation of knowledge and memory retrieval rather than as a complete theory (Alderson, 2000; Grabe, 2009). Perfetti (1986) writes:

> Schemata are not good candidates for reading ability differences. . . . [Schemata knowledge] only demonstrates specific knowledge effects. Since one individual will have knowledge structures different from another’s, this will not help with a concept of general reading ability. This point seems so obvious that it raises the question of how “schema theory” can even be applied to reading ability. The answer is that it can’t, insofar as having knowledge is concerned. (p. 22, cited in Grabe, 2009, p. 76)

Grabe (2009) concludes his discussion on schema theory by stating “most contemporary discussions of reading abilities and reading research make no strong connections to how . . . generalized knowledge structures, or schemata would explain

reading comprehension. Contemporary overview handbooks on reading research and volumes synthesizing reading make virtually no references to schema theory" (p. 77). However, as noted above, Grabe (2009) does acknowledge that effects on reading comprehension have been seen when demands on readers' specialist knowledge or cultural knowledge are made. Sadoski et al. (1991) also suggest that schema studies which used cultural differences as independent variables may be exempt from the criticism directed at schema theory. They list several studies that show significant effects of cultural background on recall (e.g. Reynolds, Taylor, Steffensen, Shirey, & Anderson, 1982; Steffensen et al., 1979). Thus, while schema theory is mostly no longer used to explain reading comprehension, there have been empirical studies which demonstrated the effects of cultural knowledge on comprehension.

For this reason, it seems reasonable to investigate proper names as an aspect of cultural knowledge and how familiarity with proper names might affect reading comprehension; however, no reference will be made to schema theory to explain any effects found. Rather, the terms 'background knowledge' and 'cultural familiarity' will be used. A caveat is reiterated concerning investigations into the effects of cultural knowledge on reading comprehension. As noted above, Steffensen et al. (1979) acknowledge that it is difficult to determine whether lower-level or higher-level processing is affected by cultural knowledge. Similarly, Alderson (2000) cautions that it is difficult to determine if the empirical effects seen in studies investigating cultural familiarity and comprehension are due to the readers' lexical knowledge or cultural knowledge (p. 46). This point is, of course, especially relevant to proper names. As was discussed in Chapter 2, there is the significant issue of whether familiarity with proper names should be treated as lexical knowledge or encyclopaedic knowledge. Thus, it may be difficult to tease apart a reader's knowledge of proper names as lexical or cultural. If the knowledge is lexical, then one would expect to see lower-level processing affected; if the knowledge is cultural, then higher-level processing may be affected.

5.3 Study 1: Replication of Erten and Razi (2009)

As part of my exploration into whether proper names burden L2 readers, the aim of this study is to investigate whether cultural familiarity with proper names aids L2 reading comprehension, and conversely, if unfamiliarity with proper names hinders reading comprehension. The study is an approximate replication of Erten and Razi (2009). An

approximate replication is a study with changes to some of the "nonmajor variables . . . in a way that allows for comparability between the original and replication studies" (Porte, 2012, p. 8). Nonmajor variables might include a different population or a different task type, for example. In this replication study, a different population was sampled: the learners in the replication were Japanese intermediate learners of English, while in the original study they were Turkish advanced learners. Otherwise, the original study was followed as closely as possible and all other variables were kept the same, including the reading text and tasks. In this way, the methodology from the original study was applied to the replication to investigate the effect of culturally familiar proper names. Also, I was interested to see if the results from the original study could be generalised to a population with a different L1 and level of language proficiency. Further details regarding methodology follow below.

To recap, Erten and Razi (2009) investigated the effect of cultural familiarity and reading activities on L2 reading comprehension in Turkish learners of English. The basic premise was to manipulate the cultural referents, most of which were proper names, in an American short story in order to make them more culturally familiar to the participant group. Comprehension scores were compared between four groups that read: the original version with American cultural referents; the original version with reading activities; an adjusted version with Turkish cultural referents; and the adjusted version with reading activities. The authors were interested in whether reading activities could activate schema and compensate for a lack of cultural knowledge.

While the Erten and Razi (2009) study did not look at the effect of proper names specifically, the methodology was considered appropriate for investigating the effect of proper name familiarity on reading comprehension. A lot of cultural information can be expressed through proper names, and as noted, many of the adjustments made to the original short story involved proper names. For example, the names of characters in the story, and the place names where the action occurs were changed. Even some of the drink and food items are referred to by proper name related terms, such as *Scotch* and *Courvoisier*. The findings in the original study were statistically significant, and showed a large effect size on reading comprehension of the adjusted version vs. the original text. For these reasons, the study was deemed appropriate to investigate the effect of culturally familiar proper names on comprehension.

There were a few challenges in carrying out the replication due to some missing details in the publication. For example, while Erten and Razi (2009) do give examples of how

they adjusted the short story, not every adjusted element is listed, possibly due to space limits in the publication. They do not list all of the recall test items but only examples; again, this may have been due to space limitations. Answers to the short answer questions from the recall test are not included; some of these answers were not obvious to me. The authors are not explicit in how they marked the tests with regard to points awarded for each correct answer. They do not explicitly state that they converted their raw scores to percentages, though presumably this is what was done: the test had a total possible 40 points but in the table of results, mean scores range from 60.45 to 79.18 (Erten & Razi, 2009, p. 69). Other than these ambiguous areas, every effort was made to follow the original study.

The research questions from the original study (Erten & Razi, 2009, p. 63), and thus also for the replication study, are:

1. Does readers' familiarity with the cultural content of short stories affect their comprehension?
2. Do reading activities used with short stories make up for the lack of cultural schema (background knowledge)?

5.3.1 Participants

The replication study was done at a private university in Japan. The participants were students from four intact academic reading and writing classes, taught by the researcher. A total of 63 (18 male and 45 female) students participated in all stages of the experiment. At the initial request for consent to participate in the study, two students declined; the results of their posttests were removed from the data set since the experiment was conducted as an in-class activity. The participants were all in their first year of university, and the majority were aged 18 to 19, though two students were 20 and 21. They were intermediate level Japanese learners of English, having studied English for six years (from the age of 12). Seventeen of the participants were enrolled in an intensive English program at the university with the goal of studying abroad in an English-medium university for their third year of study. The other 46 participants were enrolled in a regular English program with the possibility of studying abroad at some point.

The participants were divided into four groups in order to compare four different treatments. As noted above, the participants belonged to four classes taught by the

researcher; it was decided to keep the classes intact for the experiment. Therefore, to establish group homogeneity, they were pretested with a TOEFL reading practice test. Group homogeneity was important in order to compare the groups' results with one another. Each group was randomly assigned a treatment: Group 1 received an original short story without reading activities; Group 2 received the same story with reading activities; Group 3 received an adjusted version of the story without reading activities; Group 4 received the adjusted version with reading activities. Table 5.1 shows the mean TOEFL reading scores out of a possible 30 points.

Table 5.1

Mean TOEFL reading scores for each treatment group

Name of group	*n*	*M*	*SD*
1: Original text no activities (ONA)	18	13.11	3.23
2: Original text with activities (OWA)	17	15.35	2.50
3: Adjusted text no activities (ANA)	13	14.92	4.50
4: Adjusted text with activities (AWA)	15	12.73	3.86

A one-way analysis of variance was conducted to compare the groups on TOEFL reading scores. The analysis of variance showed that the difference in reading scores between the groups was not significant (F (3, 59) = 2.170, p = .101). Thus, despite the participants being enrolled in different programs (intensive and regular), no statistically significant difference was found in their reading abilities based on the TOEFL test.

To compare with the original study, those participants were 44 Turkish advanced learners of English, training to become English teachers and in their third year of university (aged 20 to 23). The participants were grouped by their GPA scores to form four homogeneous groups. By comparison then, the replication study had a slightly larger sample size and used intact groups. Also, the intermediate level participants in the replication study were slightly younger than the advanced level participants in the original study.

5.3.2 Materials

Reading text

The short story used in both the original and replication study was "The Girls in their Summer Dresses" by Irwin Shaw (2000), originally published in 1939. The story was 2,902 words in length. Two versions were used in the replication experiment: the original version and an adjusted version. As noted above, Erten and Razi (2009) were extending the work done by Alptekin (2006) and so their criteria for adjusting the story drew on Alptekin's definition of 'nativisation': "the pragmatic and semantic adaptation of the textual and contextual clues of the original story into the learner's own culture, while keeping its linguistic and rhetorical content essentially intact" (Alptekin, 2006, p. 497). Specifically, Erten and Razi (2009) changed names of characters to Turkish. They also changed the location of the story from New York to a coastal city in Turkey that their learners were familiar with. To this end, names of streets, places and buildings were changed to suit the sequence of actions in the story. Some conceptual cues were also changed, such as what the characters ate and drank. The authors provide examples of these adjustments (2009, pp. 65, 66, 75). The elements that were changed to Turkish in their study were changed to Japanese in this replication study.

The adjusted Japanese version for the replication study was created by a native speaker of Japanese and an English-Japanese bilingual. They adjusted all the elements of the text that I highlighted, so as to match the adjustments made in the original study. The American characters' names were changed to Japanese. The city plan of New York was changed to Osaka, a city near the university and one that the replication study participants were familiar with. Here is an example from the first few lines of the original story:

(2) *Fifth Avenue was shining in the sun when they left the Brevoort and started walking toward Washington Square*. . . . *Michael held Frances' arm tightly as they walked downtown in the sunlight.* (Shaw, 2000, p. 62)

Here is the same passage in the adjusted version:

(3) *Midosuji was shining in the sun when they left Shinsaibashi and started walking toward Namba.* . . . *Takuya held Misaki's arm tightly as they walked downtown in the sunlight.*

Some conceptual cues were also changed. For example, instead of drinking Scotch with their friends, the couple drinks beer. Instead of eating a steak, they eat curry. They plan to watch a Hanshin Tigers baseball game instead of the New York Giants football game. Every effort was made to replicate the decisions made in the original study regarding adjustments; that is, whichever feature was adjusted in the Turkish version, that feature was also adjusted in the Japanese version. To compare these adjustments, Table 5.2 shows the changes made to proper names and cultural cues in both the original and replication studies.

Table 5.2

Differences in cultural referents between three versions of the text

Original short story (Shaw, 2000)	**Adjusted Turkish version** (Erten & Razi, 2009)	**Adjusted Japanese version**
(***Characters***)		
Michael (Mike) Loomis	Coskun Umutlu	Takuya (Taku) Tanaka
Frances	Özlem	Misaki
The Stevensons	Nalan & Tarik	The Yamadas
(***Places***)		
New York	Çanakkale	Osaka
Ohio	Erzurum	Nara
Alice Maxwell's house	Tarik Uyanik's house	Mariko Suzuki's house
Fifth Avenue	Kordonboyu	Midosuji
The Brevoort	Bariskent	Shinsaibashi
Washington Square	Republic Square	Namba
Eighth Street	Golf Tea Garden	Nagahoribashi
Forty-fourth Street	Küçümen	Amerika mura
Between Fiftieth and Fifty-seventh Streets	(unknown)	Between Umeda and Namba
Football game	Basketball game	Baseball game
NY Giants	Fenerbahçe	Hanshin Tigers
Cavanagh's	Albatros Fish Restaurant	Koshien's Cafeteria
Washington Square Park	Kordonboyu	Namba Parks
Sardi's	Lodos Disco	Hikakebashi
Macy's	Gima	Takashimaya

(***Culture***)		
Rolls and coffee	Baguettes and tea	Rice and miso soup
Scotch	(unknown)	Beer
An extra five pounds of husband	An extra several kilos of husband	An extra two kilos of husband
A steak as big as a blacksmith's apron	A fish as big as a man's arm	A big Koshien Curry
A bottle of wine	A big bottle of raki	A beer
A new French picture at the Filmarte	A new Turkish picture "O Simdi Asker" at the AFM cinema	A new Yoshimoto Shingigeki comedy play at Namba Kagetsu
Young women with Scotties	(item deleted)	Young women with Shibu-inu dogs
A phonograph	(unknown)	A stereo
Phonograph needles	Cakes	Jewellery
Old Italian men in their Sunday clothes	Old ANZAC tourists jogging	Old men holding their red pens and sports newspapers, heading to Wins
A little Japanese waiter	(unknown)	A little Chinese waiter
Pretzels	(unknown)	Tsukidashi
Brandy/Courvoisier	(unknown)	Shochu/mugi-shochu
The Jewish girls, the Italian girls, the Irish, Polack, Chinese, German, Negro, Spanish, Russian girls	(unknown)	The office ladies, the college girls, the foreigners
Telephone	(unknown)	Cell phone

For some of the cultural items, the sentence structures were adjusted in order to allow for a more natural sounding cultural reference, as discussed with the Japanese and bilingual informants. For example, in the original short story, there is a reference to *a steak as big as a blacksmith's apron*; in the Turkish version, this is changed to *a fish as big as a man's arm*. In the Japanese version, we changed the typical food dish to *a big Koshien curry*; Koshien is a well-known curry cafeteria at the baseball stadium. It was

felt that this was a sufficient referent, and that it would be odd in Japanese to describe curry as big as something else. Similarly, in the original, there was a reference to the couple sharing *a bottle of wine*; in the Turkish version, this was changed to *a bottle of raki*. We changed the typical drink in the Japanese version to simply *beer*, having decided it would be odd for two people to share one bottle of beer.

Some cultural cues were changed due to changes in technology since the original story was written. For example, in the original story, there are references to a *phonograph*, *phonograph needles* and a *telephone* (in a bar). In the Turkish version, *phonograph needles* were changed to *cakes*; it is not known what the other changes were, as the complete adjusted story was not published. In the Japanese version, we changed these items to *stereo*, *jewellery* and a *cell phone*, respectively. In one case, we decided to make a reference longer than what was in the original. The original referent was *old Italian men in their Sunday clothes*; we changed this to *old men holding their red pens and sports newspapers, heading to Wins*. It was felt that the additional information of the red pens and sports newspapers was needed to complete the picture in the reader's mind of this familiar site in Osaka of elderly men placing bets on games. We decided that the image created for the reader was more important than the adjustment in syntax. In another case, we needed to shorten the reference: in the original, the reference was to *the Jewish girls, the Italian girls, the Irish, Polack, Chinese, German, Negro, Spanish, Russian girls*. We felt that in Japan, a Japanese man would not make such distinctions according to nationality, but rather, would refer to all non-Japanese women as simply *foreigners*. So we decided to change this to *the office ladies, the college girls, the foreigners*. Finally, in the original version, there is reference made to a *little Japanese waiter*; we felt it would be odd to refer to a waiter in Japan as Japanese, so we changed this to a *little Chinese waiter*. It is not known how these changes were handled in the Turkish version. Further samples of the original short story and the adjusted Japanese version can be found in Appendix 3.1.

Posttest

The posttest was a recall test, that is, the participants were not able to refer back to the text while doing the test. The test included ten True/False/Not Given questions, eight scrambled events to be put in order, and ten short answer questions. Each T/F/NG question and each short answer question was worth one mark. The scrambled events task was awarded a total of 20 marks, using the Weighted Marking Protocol as described by Razi (2005). The original study had published 18 examples of the posttest

questions in the appendix of their article: five of the T/F/NG questions, five of the short answer questions, and all of the eight scrambled events. However, I contacted the researchers for the full test in order to ensure the test length and questions were exactly the same. They were happy to provide this.

One change was made to the posttest after two markers piloted the test. Neither marker knew how to answer Short Answer Question 7 from the original study: *What does Michael do when something bad happens?* Not sure what the original study's authors were looking for in an answer, I changed this question to: *What are they going to do together with the Stevensons?* Appendix 3.2 contains the posttest questions for the original short story. The groups that received the adjusted version received a posttest in which the proper names had been adjusted.

Reading activities

The researchers in the original study chose prereading, while-reading, and postreading activities that the students were already familiar with for the two groups that had activities in their treatments. The original study details these reading activities along with the approximate time limit given for each. In the replication study, I sought to follow the activities and time limits as was described in the original study. The participants in the replication study were also familiar with these same types of reading activities, so there was no need to train them. The reading activities will be described in more detail in the Procedure section.

5.3.3 Procedure

Group 1 (Original text no activities; ONA) was given 30 minutes to read the original short story and 15 minutes to complete the posttest. They were informed that there would be a posttest, and that they would not be able to refer back to story when answering the questions. Group 3 (Adjusted text no activities; ANA) was given identical timings and instructions as Group 1. These were the same time limits as given in the original study.

The procedure for Groups 2 (Original text with activities; OWA) and 4 (Adjusted text with activities; AWA) was followed as outlined in the original study. Group 2 (Original text with activities; OWA) was given the original story and a total of 45 minutes to complete the reading with activities. More specifically, prereading activities were done orally in pairs and included brainstorming (three minutes) and prequestioning (three

minutes), for a total of six minutes. While-reading activities included scanning (two minutes), and skimming (two minutes). The participants were given written scanning and skimming questions to complete, and the answers were checked as a group. While-reading activities also included clarifying (two minutes), reciprocal teaching (two minutes), inferring (two minutes), which was done in pairs orally. These while-reading activities, along with quiet time for reading the story, took 35 minutes, as was done in the original study. Postreading activities done orally in pairs, included thinking aloud (two minutes), and asking and answering questions (two minutes), for a total of four minutes. Thus, the total time on task for reading with activities was 45 minutes. The class teacher (i.e. the researcher) led the reading activities to keep each group on task for the allotted time. Group 4 (Adjusted text with activities; AWA) followed the same procedure with the adjusted version of the story. Groups 2 and 4 were also informed beforehand of the posttest, and they were given 15 minutes to complete the posttest, as was done in the original study.

5.3.4 Data Analysis

Marking the tests

Two raters marked the posttests: one was the researcher; the other was a fellow teacher at the same university. The two raters marked the tests independently for comprehension, ignoring spelling or grammatical errors, as was done in the original study. The marks given by the two raters were analysed with the Pearson Correlation Coefficient Test as done in the original study. There was a positive correlation between the two raters (r = .99, N = 63, p < .01).

After the correlation analysis was completed, the tests were examined for where marks differed between the two raters, and decisions were taken together as to acceptable answers. For example, the markers had much difficulty with scoring Short Answer Question 1: *Why does Frances want to take Michael to a football match?* The correct answer was: *Because she (Frances/Misaki) knew he (Michael/Takuya) enjoyed football/baseball.* The important point here is that Frances chose the football game not because she liked football but because she knew Michael enjoyed football. The markers debated how to deal with pronoun errors for this question. For example, one participant in Group 1 (Original text no activities; ONA) answered: “Because he knew she liked football”. It was very difficult to know if this student had made a grammatical error with the pronouns (using ‘he’ in place ‘she’, and vice versa), or in fact, had not understood the story and thus made a comprehension error. Two other students in the

same group made this same error in pronouns. A student in Group 2 (Original text with activities; OWA) and a student in Group 4 (Adjusted text with activities; AWA) also made this error with pronouns, and there again, it was difficult to know if the error was grammatical or with comprehension. In the end, the markers decided to allow the error in pronouns for the original version, thus allowing for the possibility that the error was grammatical. For the adjusted version, where the proper names of the characters were known to the participants, there should have been no confusion as to gender of the characters. So it was decided this must have been an error in comprehension for the learners with the adjusted text. It is not known if this problem occurred in the original study; the authors only mention that they ignored grammatical errors. It should also be noted that the name of the female character in the original version is problematic: *Frances* has a male homophone *Francis*. In hindsight, it probably would have been better to change the question altogether but the issue did not come up during the piloting of the test.

There were other difficulties scoring the Short Answer questions, and the markers agreed on the following points. First, for Short Answer Question 3 (*Where did Frances & Michael meet for the first time? Describe Michael's feelings at that time*.), it was agreed that this was a two-part question and therefore a successful answer would have both parts. Next, for Short Answer Question 5 (*Why does Frances feel good on that Sunday morning?)*, it was agreed that some reference needed to be made to *being with her husband*, not only *having breakfast* or *sleeping in.* Lastly, for Short Answer Question 9 (*What is the favour that Frances asks Michael to do for her?*), answers were accepted that referred to either *not looking at other women* or *not talking about other women*. Given the markers' difficulty in scoring the Short Answer questions, it might have been better to use multiple-choice questions. That way, the markers would not need to conjecture as to the participants' intended message in the answers.

The scrambled events task was worth a total of 20 points and marked using the Weighted Marking Protocol (Razi, 2005), as in the original study. This marking protocol awards partial points even if all the events were not put in the right order.

Statistical analysis

The posttest scores were analysed using ANOVA to see where group differences occurred; Cohen's *d* was used to calculate effect size on the group differences, as was done in the original study. Erten and Razi (2009) used the post hoc LSD (least significant difference) test to see where group differences occurred. However, because the post hoc LSD requires significant results from the ANOVA test, this analysis was not done in the replication study. More details follow in the Results section.

5.4 Results

The results of the mean scores from the posttest are shown in Table 5.3. As noted above, the total number of points for the posttest was 40. In the original study, this raw score was converted to a percentage score; though the authors did not explicitly state that they converted raw scores to percentages, this was inferred from how their data was presented, as can be seen in the last two columns in Table 5.3. In order to compare the results from the replication study to the original study, this conversion of scores to percentages was also done in the replication study. Table 5.3 compares the mean scores and standard deviations from posttest in the replication study and the original study.

As can be seen from the descriptive statistics, the participants in the replication study had lower scores than those in the original study. The group that was expected to do the best (adjusted text with activities) performed the worst in the replication study. Also of note is that the standard deviations in the replication study are very large relative to the mean scores. A one-way analysis of variance was conducted to compare the groups by posttest scores. The analysis of variance showed that the difference in posttest scores between the groups was not significant (F (3, 59) = 0.425, p = 0.736). Because the omnibus ANOVA was not significant, no post hoc tests were run.

Table 5.3

Mean scores by groups in replication and original studies

	Replication study			Original study (Erten & Razi 2009)	
Treatment	*n*	*M*	*SD*	*M*	*SD*
1 Original text no activities (ONA)	18	46.78	14.43	60.45	13.03
2 Original text with activities (OWA)	17	48.59	21.31	64.55	10.25
3 Adjusted text no activities (ANA)	13	52.46	24.29	69.91	10.28
4 Adjusted text with activities (AWA)	15	44.60	16.08	79.18	9.98

With regard to the research questions, unlike the results of the original study, the results from the replication study do not support the first hypothesis that cultural familiarity has a significant effect on reading comprehension. Group 4 (adjusted text with activities) did not outperform any of the other groups; in fact, the mean for this group was the lowest. Group 3 (adjusted text no activities) did slightly better than the groups with the original texts, but there was no statistically significant difference found in performance between the groups.

Unlike the results of the original study, the results from the replication study do not support the second hypothesis either, that is, while reading activities are helpful, the effect of cultural familiarity will remain strong. No significant differences were found between the mean scores of the groups that had reading activities and those that did not have these activities. The large standard deviations show that the distribution of scores was very spread out.

5.5 Discussion

The results of the replication study indicate that neither cultural familiarity nor reading activities impacted reading comprehension. These findings are not predicted by previous research that suggests a positive role of cultural familiarity on comprehension (e.g. Erten & Razi, 2009; Johnson, 1981; Reynolds et al., 1982; Steffensen et al., 1979). The findings in the replication study are also not predicted by the literature which points to the positive impact that reading activities have on comprehension (e.g. Erten & Razi, 2009; Grabe, 1991; Grabe & Stoller, 2002). The large standard deviations indicate that the groups were not as homogeneous as those in the original study, despite the TOEFL test results (described in section 5.3.1), which suggested the groups were fairly homogeneous.

In order to account for these disparate findings, differences between the original and replication study were examined. One variable that was different between the two studies was the proficiency level of the participants (i.e. advanced vs. intermediate, respectively). The text may have been too difficult for the intermediate participants, as suggested by the lower mean scores in the replication study. If the text was too difficult, then cultural familiarity might not have had an effect on comprehension. In Chapter 2, the effect of vocabulary knowledge and background knowledge on text comprehension was discussed with reference to the Schmitt et al. (2011) study. In that study, the authors found that 95% vocabulary known can result in 60% comprehension; 98-99% coverage can lead to 70% comprehension; and 100% coverage can result in 75% comprehension. With regard to background knowledge, they found that at the 90% to 93% vocabulary coverage levels, there was no advantage seen in comprehension with a familiar topic. But at the 94% - 100% levels, comprehension scores were better for the familiar topic than the unfamiliar topic. This finding indicates that background knowledge does affect reading comprehension, but only if at least 94% vocabulary is known. Thus, if participants in the replication study knew less than 94% of the vocabulary, this could account for why no effect for cultural familiarity was seen. Post hoc analysis was carried out on the lexical difficulty of the text in order to explore this possibility further.

5.5.1 Post hoc analysis

To gauge the lexical difficulty of the text, a lexical profile for the original short story version was generated with Vocabprofile on Compleat Lexical Tutor v. 8.3 (Cobb, n.d.),

using the British National Corpus and the Contemporary Corpus of American English up to the 25K band (BNC-COCA 1-25K). As described in the Chapter 4, Vocabprofile is an online tool that generates a vocabulary profile of a reading text: the output of the profile displays the words, and the percentage of words, at each 1,000-word level of the BNC-COCA 1-25K. When creating a vocabulary profile, the user of the program needs to decide what to do with proper names in the text. The program offers two options: to treat proper names as known, in which case, the profiler will place words with initial capital letters in mid-sentence in the 1K band; or to ignore proper names, in which case these items are not categorised by the profiler (i.e. the program ignores them). There is a third option for the user, one that is not directly flagged by the program: to not choose either of the two options, in which case the program will place proper names in an 'off list' category (i.e. a category for words which do not appear in the any of the 25K bands)[12]. The profile output will vary according to which treatment option for proper names is chosen. For this profile, I chose the third option (i.e. proper names to appear as 'off list') because I did not assume that the participants understood the proper names (as in the first option). Nor did I want the program to ignore the proper names (as in the second option) because proper names were a variable to reading comprehension in this replication study (i.e. as an aspect of cultural knowledge).

Schmitt et al. (2011) suggest that for independent reading, as was done in the replication study, 98% vocabulary coverage is optimal because this can result in 70% comprehension. According to the vocabulary profile, 98% coverage is not achieved in this text. The highest vocabulary coverage is 94.35% seen at the 15K level (beyond which are 'off list' words). This indicates the text was quite difficult for these intermediate level L2 readers. As noted above, Schmitt et al. (2011) found an advantage for background knowledge at the 94% vocabulary coverage level. In this profile, 94.02% coverage was seen at the 7K band. Intermediate L2 readers are unlikely to have vocabulary knowledge of the first 7,000 words; based on extensive vocabulary size testing I have done in Japan, I predicted these participants had a vocabulary size of about 3,700 words (my own unpublished research). This estimate of vocabulary size is supported by other research (Mclean, Hogg, & Kramer, 2014). Thus, it is unlikely that the participants knew 94% of the vocabulary in the short story.

[12] Since the time of writing, these options have now changed on the website. All three options appear to the user, with the first option checked by default: treat proper names as 1K.

By comparing the treatments and posttest scores from the replication study with the treatments and comprehension scores from Schmitt et al. (2011), I can estimate what percentage of vocabulary was known to the participants in the replication study. In this comparison, I will leave aside the groups that had the reading activities because Schmitt et al. did not have a reading activity treatment. I compare the treatment of familiar proper names (i.e. adjusted version of the story) with the treatment of familiar topic from the Schmitt et al. (2011) study. I also compare the treatment of unfamiliar proper names (i.e. original version of the story) with the treatment of unfamiliar topic from Schmitt et al. (2011). Table 5.4 shows the comparison of comprehension scores by treatments from the two studies.

Table 5.4

Comparison of comprehension scores by treatment

	Replication study	Schmitt et al. (2011) study	
Adjusted version with no activities (familiar names)	52.5%	52.2%	90% vocabulary coverage and familiar topic
Original version with no activities (unfamiliar names)	46.8%	50.2%	90% vocabulary coverage and unfamiliar topic

The comprehension scores by treatments are very similar at this vocabulary coverage level, so I conjecture that the participants in the replication knew about 90% of the vocabulary in the short story. If that is correct, then the text was indeed very difficult for these readers: 90% known vocabulary means that one word is unknown in every ten running words, which is a heavy vocabulary load. If this conjecture is correct, then it would account for the lower comprehension scores seen in the replication study as compared to the original study. Simply put, the intermediate readers attained lower comprehension scores than the advanced readers because the text was too challenging lexically. This may account for why the effect of cultural familiarity for proper names was not seen in the replication study: because the 94% level was not reached where an advantage for background knowledge could be seen.

While the vocabulary coverage for the advanced learners in the original study is unknown, I can make the same comparisons of comprehension scores and treatments with the Schmitt et al. (2011) study to estimate what the coverage was. Table 5.5 compares the comprehension scores from the two studies by treatment.

Table 5.5

Comparison of comprehension scores by treatment

	Erten & Razi (2009)	Schmitt et al. (2011)	
Adjusted version with no activities (familiar names)	69.9%	71.1%	98% vocabulary coverage and familiar topic
Original version with no activities (unfamiliar names)	60.5%	63.3%	98% vocabulary coverage and unfamiliar topic

Again, the similarity in comprehension scores by treatments in the two studies is marked. I would surmise that the advanced learners in the original study had 98% vocabulary coverage, thus accounting for their higher comprehension scores. In addition, this vocabulary coverage could be the reason why the effect of cultural familiarity was so strong in the original study. The 94% coverage was reached where an effect for background knowledge would be seen.

5.6 Summary

There was no statistically significant difference in reading comprehension among the treatment groups in the replication study. The group that read the version of the story in which the proper names had been adjusted to be culturally familiar did not have better comprehension than the group that read the original version. There was also no significant difference seen in the groups that read the text with reading activities and those that read without reading activities. These findings do not support those found in the original study. These findings are also not supported by previous research that has found positive effects on reading comprehension from both cultural familiarity and reading activities. For this reason, a post hoc analysis looked at the lexical difficulty of the text, treating proper names separately (i.e. as 'off list'). By comparing comprehension scores and treatments from another reading study that had looked at

the effect of vocabulary knowledge and topic familiarity on comprehension, it was estimated that the learners in the replication study knew about 90% of the words in the text. This vocabulary coverage renders a text very difficult for independent reading, and thus might account for the low comprehension scores. Furthermore, since at least 94% of the vocabulary needs to be known for cultural familiarity to have an effect on comprehension, the low vocabulary coverage might account for why no effect for cultural familiarity was seen in the replication study. Therefore, because the text used in the replication study may have been too difficult for the participants, it was decided that a second study be run, controlling for lexical difficulty of the text. Also, to extend the research, three treatments of proper names are used: culturally familiar proper names, culturally unfamiliar proper names and no proper names, only common noun referents. The second study is reported in section 5.7, followed by the Results and Discussion.

5.7 Study 2: Three treatments of proper names

The aim of Study 2 is to investigate if the familiarity of proper names facilitates reading comprehension in texts in which most of the vocabulary is known. In designing Study 2, issues with the replication study (Study 1) were addressed which may have contributed to the results that did not support the original study or other previous literature. One possible reason for the disparate results was identified from a vocabulary profile of the reading text: the vocabulary load might have been too difficult for the intermediate level participants in Study 1. So, to address the lexical difficulty of the reading material in Study 2, the vocabulary coverage is aimed at 98% to the 2K level, which should allow for adequate comprehension for intermediate L2 readers. Also, this high percentage of known vocabulary should allow for an effect of cultural familiarity to be seen on reading comprehension (see Chapter 2, section 2.3, for the discussion regarding vocabulary knowledge and reading comprehension).

Another possible issue that may have contributed to participants' low comprehension scores in Study 1 was the text length. The text in the replication study was 2,910 words. This is much longer than texts usually encountered by these students in the classroom. To compare, texts from their course textbook are an average length of 800 words. Also, a text of nearly 3,000 words is demanding on participants, in terms of attention span and motivation needed to read and understand a longer text. In view of these considerations, a shorter text length is used in Study 2. Two texts of 500 words

were selected for Study 2 (though this length was further decreased to 250 words due to changes in postreading tasks; details to follow in Materials section 5.7.2). In sum, a shorter text with high vocabulary coverage might provide conditions for an effect of familiarity with proper names on reading comprehension to be seen.

In addition to the treatments of proper names as culturally familiar and unfamiliar used in Study 1, a third treatment is introduced in Study 2: no proper names, only common nouns to refer to the person's role (e.g. *the defendant, the judge*) (hereafter, referred to as 'noun referents'). This third condition is introduced here to extend the research and investigate how the presence or absence of proper names affects comprehension. For example, proper names might aid comprehension by helping the reader establish a foothold with details found in the text. Alternatively, proper names might add to the processing load for the reader and burden the working memory, thus hindering comprehension.

The research question for Study 2 is:

> Is reading comprehension better when the proper names in a text are in from the L2 culture (here, English), from the L1 culture (here, Japanese), or when the text does not have any proper names (only common noun referents)?

It is predicted that reading comprehension will be highest when the proper names are culturally familiar to the readers, as previous studies have found an effect for cultural background knowledge on comprehension (see section 5.2 above). As for how the absence or presence of proper names affects reading comprehension, this aspect of the study is exploratory in nature, as I am not aware of other studies that have looked at this. There is the possibility that the original version of the text (i.e. with or without proper names) will be the version that generates the highest reading comprehension. Fabricated or manipulated texts might require more processing from the reader by the fact that the language or text organisation has been rendered inauthentic. This effect was seen in Johnson's (1981) study, where L1 readers had lower comprehension scores for the simplified texts.

5.7.1 Participants

A total of 111 Japanese students participated in this two-part experiment (74 females, 37 males). They were all first year university students at a private university in Japan, aged 18-19 years old. They had all studied English for six years before entering university. They were intermediate learners of English, placed at very similar levels within the university according their results on the Global Test of English Communication (GTEC, an English proficiency test used in Japan). Based on their test results, these participants were placed in the six highest levels (of 50) of the first-year program. They were all in the same program, which focuses on developing academic English skills.

The participants were grouped according to their intact classes, making this a quasi-experimental study. There were six groups of similar size, between 14 and 21 students in each. Three of the groups were taught by the researcher; the other three groups were taught by another instructor. All six groups had used the same course textbook focusing on academic reading and writing skills, throughout the academic year (two semesters of 15 weeks each), and so had been exposed to the same reading materials and task types.

5.7.2 Materials

Texts were taken from an existing course book in order to avoid a possible bias that might result from the researcher creating new, and perhaps inauthentic, texts for the purposes of the experiment. Two reading texts were taken from an intermediate course book (*Inside Reading 2: The Academic Word List in Context,* by Lawrence J. Zwier, and published by Oxford). This was a different textbook from the course textbook. Because the readings were based on academic content areas, such as literature and business, it was felt that the readings would be similar to what the participants normally encounter in the classroom, and thus offer ecological validity. The texts were between 450 and 560 words in length. Each text consisted of two accounts to illustrate a particular phenomenon: one text consisted of two accounts to illustrate Internet fraud, and the other had two accounts of where meteorological evidence had been provided in court cases. The selected reading texts were modified for length and vocabulary.

These two particular texts were chosen on the basis of how proper names were handled in the original version. Text 1 (labelled ‘Internet’), which had examples of

Internet fraud, was selected because the writer had inserted random names, presumably to protect the identity of the real people involved. For example, in introducing a character, the writer uses this convention: *One girl, (let's call her Jane). . .* Text 2 (labelled 'Meteorology'), which had examples of meteorological evidence in court cases, was selected because no proper names were used in telling the story even though there were many characters. Because the original version of each text was to be manipulated in terms of the proper name usage, it was important to use two texts in which the original version had different treatment of proper names. In this way, an effect due to the treatment of proper names might be seen in one or both of the texts.

To recap, the original version of Text 1 (Internet) had English proper names. The original version of Text 2 (Meteorology) had no proper names, only noun referents (i.e. common nouns to refer to the person's role). Three versions were created for each text: version A (English proper names); version B (no proper names); and version C (Japanese proper names). See Appendix 4.1 for the three versions of each text. Because the original version of Text 2 (Meteorology) had no proper names, to create version A (English proper names), the ten most common surnames were chosen from the United States 2000 census ("2000 Census," 2000). To create version C (Japanese proper names) for both texts, names were taken from a common surnames list ("Japan's top 100 most common family names," 2009).

To ensure that the vocabulary in the texts was at an appropriate level for the participants, each text was profiled using Vocabprofile (BNC-COCA 1-25K) on Compleat Lexical Tutor v. 8.3 (Cobb, n.d.). For Text 1 (Internet), modifications were made to the text by adding glosses for any vocabulary beyond the 1K and 2K bands, so that 98% text coverage was found at 3K level. The items modified beyond the 1K and 2K level included: *fraud* (4K), which was glossed below the text due to the length of the definition, and *ID* (5K), which was glossed in the text because the explanation was short, and thus non-disruptive to the reader. Two items beyond the 1K and 2K level were not glossed because I assumed participants were familiar with these items: *primary* (3K), from *primary school*; and *false* (3K), which I assumed students would know from test questions. The glosses were taken from an online learner dictionary (http://dictionary.cambridge.org) and can be seen in the texts in Appendix 4.1. (In the appendix, the glosses appear only in version A, though in the experiment, each version contained the same glosses).

Likewise, for Text 2 (Meteorology), modifications were made to vocabulary until 98% vocabulary coverage was found at 3K. Items beyond the 1K and 2K levels were dealt with as noted here. Items that were glossed below the text included: *meteorologist, defendant, expert,* and *courtroom*. I did not gloss *bench* (3K) because I felt the context was sufficient for comprehension: *sitting on a park bench*. Two items, *northeast* and *southwest,* were categorised as off-list by the profiler; however, I did not gloss them, assuming the students could infer meaning from the word parts because *north, south, east*, and *west* are found in the 1K band.

To test comprehension of the texts, ten True/False/Not Given questions were generated for each version of each text (refer to Appendix 4.2 for posttest questions for version A of each text). This was a recall test: students could not refer back to the text to answer the questions. A recall test was used because this format had been used in the replication study (see Study 1 section 5.3.2 Materials). The questions focused on comprehension of the main ideas of each text.

5.7.3 Piloting

Pilot study 1

Before conducting Study 2, the reading texts and comprehension test were evaluated with a pilot study. A total of 33 students from two intact classes participated in the pilot study. These students were different from those in Study 2. They were first-year students at the same university, though in a different program from the Study 2 participants. They were in an intensive program, meeting for eight 90-minute classes of academic English per week compared to the six 90-minute classes that the Study 2 participants had. Their English proficiency was similar to those in Study 2; that is, at the beginning of the academic year, they had been placed at similar levels as Study 2 participants, based on their GTEC scores.

Because of the small number of participants in the pilot study (N = 33), only two versions of Text 1 (Internet) were used: version A (English proper names) and version B (no proper names). I wanted to test if there was any difference in comprehension scores from the two versions.

The participants were given instructions in English, orally and written, that they would have five minutes to read the text, and then would answer ten True/False/Not Given

questions without being able to look back at the text. Sixteen students read version A (English proper names) and 17 students read version B (no proper names); text assignment was random.

The tests were scored and the results are shown in Table 5.6. The mean scores (out of a possible ten points) indicated that participants had an adequate understanding of the texts, and thus, the lexical difficulty of the texts matched their abilities. However, very little difference was found between the mean scores from the two versions.

Table 5.6
Pilot study 1: Comprehension scores from two versions

Text 1 (Internet) version	n	*M*	*SD*
A (English proper names)	16	7.4	1.93
B (No proper names)	17	7.6	1.50

An independent-samples t-test was conducted to compare the mean scores; the result was not statistically significant, $t(31) = -.251$, $p = .803$.

Possible reasons for the similar mean scores relate directly to the research question; that is, the presence or absence of proper names had no effect on comprehension. However, the similar comprehension scores may also have been a result of the test instrument. Because there were only ten questions, it is possible that there were not enough items to discriminate between participants with good comprehension and others who guessed correctly, for example. Also, a problem with using multiple-choice questions is that plausible distractors can alter stored information that participants have of the text (Koda, 2004, p. 239).

To address these concerns, I decided to add a free recall task: participants are asked to rewrite the story in their own words based on what they remember. This would ensure that there was no contamination from test items. Free recall is generally considered the most straightforward assessment of reader comprehension (Koda, 2004, p. 236). However, in using the free recall task, I was concerned with the length of the texts. Each text had between 450 to 560 words and consisted of two accounts: Text 1 (Internet) consisted of two accounts to illustrate Internet fraud; and Text 2 (Meteorology) consisted of two accounts to illustrate cases where meteorological

evidence was used in law courts. My concern with the text length was that there were many details from the two accounts and that participants might mix up the details between the two accounts. So I decided to cut the text length and keep only one account in each text. That left the word count of approximately 250 words per text.

A concern with conducting free recall tasks in the L2 is that the task requires productive skills to demonstrate receptive understanding. Thus, while some participants might understand the text receptively, they might not have the productive (writing) skills to demonstrate that understanding. For this reason, I decided to keep the multiple-choice questions, but administer them after the free recall task, thus providing an opportunity for students with weaker productive skills to demonstrate comprehension. Because these revisions to the texts and comprehension tasks were significant, I conducted a second pilot.

Pilot study 2

A total of 18 participants took part in the second pilot study. They were some of the same students who took part in the first pilot study but from one class only. The other class that took part in the first pilot was not available due to time constraints in the semester. As noted above, the students in the pilot study were different from the participants in Study 2; they were at a similar proficiency level to those in Study 2, though in a different program in the university.

For the second pilot, I used the other text, Text 2 (Meteorology), and piloted all three versions: Version A (English proper names), Version B (no proper names) and Version C (Japanese proper names). Only the free recall task was piloted.

Students were given instructions written in English and orally that they would have five minutes to read the text. They would then have ten minutes to write what they remembered of the main ideas of the text, without being able to look back at the text.

I scored the students' text summaries with a marking rubric: I identified five main ideas from the text and allocated each main idea two points each, allowing for partially correct answers, for a total of ten points. Four of the five main ideas necessitated reference to one or more of the characters. (See Appendix 4.3 for the marking rubric).

The comprehension scores (out of a possible 10 points) from the free recall task from the three versions are given in Table 5.7. There was a more discernable difference in the mean scores between the three versions from the free recall task than was seen in Pilot study 1.

Table 5.7

Pilot study 2: Comprehension scores between three versions

Text 2 (Meteorology) version	n	*M*	*SD*
A (English proper names)	6	7.7	2.6
B (No proper names)	6	6.5	2.7
C (Japanese proper names)	6	9.3	1.2

An ANOVA was run to compare the mean scores between the three versions. The results showed the differences in mean scores was not statistically significant, F (2, 15) = 2.40, p = .125. The non-statistically significant result may have been due to the small sample size.

I decided to use the free recall task in Study 2 because a greater difference could be seen in mean scores in the second pilot than in the first pilot study. As noted above, the multiple-choice questions were retained and used after the free recall task, so that students with weaker productive skills might still demonstrate comprehension. Also, because the tasks were presented in this order, the multiple-choice questions would not influence recall.

5.7.4 Procedure

To recap, two pilot studies had been conducted to evaluate the materials. Based on the results of these pilot studies, adjustments were made to the reading texts and tasks. The text length was reduced from 500 to 250 words per text because of a change in comprehension task: a free recall task was added, for which participants write a summary of what they remembered from the text. This summary is worth ten points for five main ideas. The True/False/Not Given questions are administered after the free recall, so that students with weaker productive skills can still demonstrate comprehension. These questions are worth ten points. Thus, the two tasks generate a total of 20 points. These revised materials were used in Study 2.

The research design of Study 2 included counter-balancing, so that each group read each text (Internet and Meteorology), though a different version (English proper names, Japanese proper names or no proper names), over two test times. Though it is unlikely that the order of task presentation would have any effect, the balance design also took this variable into account. Table 5.8 summarises the design.

Table 5.8

Study 2 research design: counter-balancing

Group	Test Time 1	Test Time 2
1	Text 1 English PN	Text 2 No PN
2	Text 1 No PN	Text 2 Japanese PN
3	Text 1 Japanese PN	Text 2 English PN
4	Text 2 English PN	Text 1 Japanese PN
5	Text 2 No PN	Text 1 English PN
6	Text 2 Japanese PN	Text 1 No PN

Note. PN = proper names

Participants were given a choice whether or not to participate, as per ethical research guidelines; they were told that the study was investigating reading skills of Japanese learners of English, and that their decision not to participate would not affect their course grade in any way. They were given a consent form to sign outlining these main points in both English and Japanese; specific details of what participating in the study would involve were also given (see Appendix 4.4 for consent form). All students agreed to participate.

The experiments were conducted during regular class time, Test Time 2 taking place one week after Test Time 1. The instructions to the participants were written in English at the top of the reading text, and also given orally: *You will have five minutes to read this short text. Then, you will have ten minutes to write a summary of the main ideas of the text. You will not be able to look back at the text when you write the summary. Then, you will answer 10 multiple choice questions about the text.* Participants were given five minutes to read, and then the reading text was collected. They were given ten minutes to write the free recall on one A4 page. This was collected, and then they were given five minutes to complete ten multiple-choice questions.

5.7.5 Data Analysis

Participants' free recall summaries were scored by two markers using the rubric in Appendix 4.3. I was one marker and the other was a teacher at the same university. After marking the summaries independently, the two markers compared scores for each summary. In 55% of cases, scores differed, never more than by four marks and in most cases by only two marks. Where scores differed, the two markers re-examined the text summary and compared it to the rubric. Clarifications to the rubric were made based on these discussions (also found in Appendix 4.3). Through discussion of the rubric points, the markers reached agreement on the final score for each summary.

Each participant's text summary score (out of ten points) was combined with the multiple-choice test score (out of ten points) for a total of 20 points. Descriptive statistics were generated for each group, by text, treatment and test time. Further analysis was carried out on descriptive statistics by text and treatment only because test time did not seem to impact comprehension scores. The means for each treatment by text were compared for inferential statistical testing.

5.8 Results

Descriptive statistics were first generated for each text by group, test time, and treatment. Table 5.9 shows the mean scores out of a possible 20 (ten points from the free recall summary and ten points from the multiple choice test).

Table 5.9

Mean comprehension scores by group, treatment, and test time

		Test Time 1		Test Time 2	
Group	n	*M*	*SD*	*M*	*SD*
1	18	11.6 EPN	3.13	12.8 NPN	5.68
2	17	11.7 NPN	3.16	13.2 JPN	4.41
3	14	12.6 JPN	2.87	14.7 EPN	2.95
4	20	14.2 EPN	3.07	12.0 JPN	3.88
5	21	14.2 NPN	3.83	14.3 EPN	3.62
6	21	12.7 JPN	3.76	12.4 NPN	3.44

Note. EPN = English proper names; NPN = No proper names; JPN = Japanese proper names

Test time did not seem to have an effect on comprehension scores. Groups 1, 2 and 3 did slightly better Test Time 2; Group 4 did slightly worse Test Time 2; and Groups 5 and 6 performed almost the same each test time. Another related issue to test time concerned the research design, which was not perfectly balanced; that is, each group did not read a version of each treatment (English proper names, Japanese proper names and no proper names), for which a third text would have been required. Because of the imperfect counter balancing and little effect seen from test time, it was decided to analyse the data by treatment and text, ignoring test time. That is, the comprehension scores from the groups that read the same versions of Text 1 were compiled; the scores from the groups that read the same versions of Text 2 were compiled. The descriptive statistics for comprehension scores by text and treatment are shown in Table 5.10.

Table 5.10

Comprehension scores by text and treatment

Treatment	Text 1			Text 2		
	n	*M*	*SD*	*n*	*M*	*SD*
English names	39	13.03	3.63	34	14.41	2.99
Japanese names	34	12.27	3.47	38	12.90	4.01
No names	38	12.05	3.30	39	13.56	4.76

Data were initially analysed by text to check that assumptions of parametric tests were met. Initial data exploration showed that the assumptions of normal distribution and homogeneity of variance were violated. Therefore, the non-parametric Kruskal-Wallis test was run to compare means between the three treatments for each text. There was no statistically significant difference seen in the treatments for Text 1 (χ^2 (2) = 1.308, p = .520), with a mean rank comprehension score of 60.67 for English proper names, 54.16 for Japanese proper names, and 52.86 for no proper names. There was also no statistically significant difference seen in the treatments for Text 2 (χ^2 (2) = 2.389, p = .303), with a mean rank comprehension score of 60.41 for English proper names, 49.58 for Japanese proper names, and 58.41 for no proper names.

5.9 Discussion

Comprehension scores in Study 2 were better than in Study 1 (the replication study), perhaps due to the 98% lexical coverage in the reading texts. In Study 2, comprehension scores for Text 1 ranged from 60 - 65%, and for Text 2 from 65 - 72%. In comparison, in Study 1, comprehension scores ranged from 45 - 53%. It is important that comprehension scores were higher in Study 2 because this suggests the texts were at a more suitable reading level for the participants, and thus, the conditions were set for an effect for cultural familiarity to be seen. However, there was no statistically significant difference seen in comprehension scores between the three treatments (English proper names, Japanese proper names, and no proper names) in either text. This finding suggests that there was no effect for cultural familiarity of proper names on reading comprehension. Furthermore, the findings suggest there was no effect for the presence or absence of proper names.

As noted above, there is a difficulty in using manipulated texts to test reading comprehension. Manipulation of elements will render the text less authentic and might in fact increase the processing load. For example, in Johnson's (1981) study, the L1 participants had lower recall scores for the adjusted versions than for the original texts. Johnson (1981) suggests that the text cohesion and general readability may have suffered in the adjusted versions. Alptekin (2006) also criticises the use of simplified texts as the manipulation may interfere with the reader's ability to draw on knowledge related to text organisation. Lexical priming theory (Hoey, 2007) also claims that because fabricated texts are not likely to maintain the natural priming of the language, this can have detrimental effects on fluency. Perhaps those negative effects can also extend to comprehension.

Thus, it may be the case that removing or adding proper names from the original versions rendered the texts more difficult to understand. The original version of Text 1 had English proper names. That version did have the highest mean comprehension score among all the versions (M = 13.03). Text 2 was originally written with no proper names, and one could argue therefore, no names were needed to understand the story. Adding names may have caused more processing difficulties. The original version of Text 2 had the second highest mean score (M = 13.56) (see Table 5.10). However, as was noted in the Results section, there was no statistically significant difference in comprehension scores, so it is difficult to draw conclusions in this regard.

One possibility for why the text versions with Japanese proper names did not result in higher comprehension than the other treatments is related to potential processing difficulties at the decoding level. The participants in Study 2 might have less processing experience reading Japanese proper names in the Roman alphabet, as Japanese proper names are usually written in Kanji (Chinese characters). Transliteration of Japanese using the Roman alphabet is done with a writing system called *romaji.* However, it was assumed that these participants would have little difficulty reading Japanese proper names in romaji because all Japanese students are taught how to read and write Japanese words using this transliteration system. Japanese proper names sometimes appear in romaji on nameplates on houses or apartments, and romaji transliteration of names is often required for web-based applications. Thus, it was assumed that these participants would have some processing experience with proper names written in romaji.

There were several limitations to Study 2 that should be considered when interpreting the results. First, there were issues related to the research design, which could have been simplified. In an attempt to anticipate all possible outcomes, the study design became unduly complicated. For example, two texts were selected in consideration of the treatment of the proper names in the original version (i.e. one original text was written with no proper names; the other drew overt attention to the proper names). There is always a concern when modifying a text for experimental purposes that the modification will render it inauthentic and difficult to read. Thus, with the original text without proper names, the concern was that inserting proper names would make it more difficult to read; and for the second text, the concern was that removing the proper names would also cause a reading burden. In trying to balance these concerns, both texts were retained and used. However, one text might have been sufficient and allowed for a simpler experiment plan. Also, the design was not perfectly counter-balanced; that is, each group did not read each treatment, for which a third text would have been required. Alternatively, the variable of the presence or absence of proper names could have been investigated in a separate study. Thus, with only one text and two treatments, the overall design would have been more straightforward.

Also, the procedure did not take into account the effect of time on task. The hypothesis was that unfamiliar proper names would cause difficulty for the reader, for example, by slowing down the reader. However, the procedure allowed the participants five minutes to read 250 words. These students had an average reading speed of 180 WPM, with the slowest readers at 100 WPM (based on results from a timed-reading course done

with the same participants over the semester). Given that the slowest readers would have only needed 2.5 minutes to read 250 words and that 98% of the vocabulary was known to them, then it seems reasonable to conclude that in five minutes, the participants would eventually solve any difficulty related to unfamiliar proper names. A study design that included recording participants' reading time would shed more light on how decoding processes were affected by culturally familiar or unfamiliar proper names.

Another issue to take into account was how comprehension was tested. While the free recall task is considered the most straightforward way to assess comprehension, it does have its drawbacks. As noted earlier, there is the problem of the L2 participants' productive skills. Also, there is a problem that the research designer chooses the main ideas for the rubric. This is problematic because what is relevant to one reader may not be to another. So, while a participant may have understood the story well, this might not have been reflected in the score, as main ideas that were relevant to her/him may not have matched the rubric. An alternative tool to check comprehension is a graphic organiser: this diagram shows organisation of the text, and the reader's task is to fill in blanks to demonstrate understanding of the main ideas. The reading texts used in Study 2 were too short to create graphic organisers, so longer texts (e.g. 800 words) would have been better for such a testing instrument.

In summary, this study found no statistically significant difference in comprehension scores between the three treatments of L1 culture proper names, L2 culture proper names, and no proper names, in either reading text. Overall comprehension scores were adequate, perhaps due to the 98% lexical coverage. However, several limitations were noted concerning the research design, in particular the materials and procedure.

5.10 Conclusion

Several well-designed, robust studies have found an effect for cultural background knowledge on reading comprehension. Four of these studies were reviewed in detail above in section 5.2. While those particular studies all referred to schema theory to account for the effect of cultural familiarity on comprehension, contemporary reading models tend to avoid the term 'schema' due to difficulties in operationalising schema theory. As was described in Chapter 2, restrictive interactive models view reading as information processing, and in this respect, emphasise the role of lower-level

processing to successful comprehension; such models also predict there is no interaction between higher- and lower-level processing but only within specific processes, such as word recognition. These models do not aim to predict how schema might be used by the reader for text comprehension. The effect seen in the studies reviewed in section 5.2 could be attributed to cultural background knowledge, though it was not clear what level of processing was affected, lower- or higher-level processing. Two studies were carried out here to investigate whether the effect of cultural knowledge could be extended to proper names. Study 1, the replication study, did not find any effect for cultural familiarity of proper names, perhaps due to the lexical difficulty of the text for the participants. Study 2 addressed the lexical difficulty issue, but other aspects of the study design became overly complicated. Thus, the reading texts and comprehension tasks may not have been optimal to investigate the effect of proper names on participants' understanding. At any rate, reading comprehension scores were very similar between the treatments used in both studies, suggesting there is no effect on comprehension regarding the familiarity with, or presence of, proper names.

However, it may be the case that the two experiments reported here did not target or effectively isolate the level of processing used for understanding proper names in a text. As noted, the studies reviewed in section 5.2 were not able to definitively answer an important aspect related to comprehension: what level of processing is affected by cultural background knowledge. Lower-level processing relates to information the reader gets from the text, and in this way, is data-driven. For example, lower-level processing involves lexical access and syntactic parsing. Higher-level processing relates to the knowledge that the reader brings to the text, including cultural knowledge. Higher-level processing includes making inferences, understanding text organisation and synthesising. The studies reviewed in section 5.2 were not able to definitively answer whether cultural knowledge affected processing at the lower level or higher level.

To recap, Steffensen et al. (1979) acknowledge that it is not clear if participants' longer reading times were the result of a lack of cultural background knowledge (i.e. higher-level processing) or the unfamiliar lexical items that were culturally specific (i.e. lower-level processing). Johnson (1981) noted that the use of cultural knowledge was more evident on inferential questions, suggesting that higher-level processing was affected. Alptekin (2006) also noted that the effect of cultural familiarity seemed to be more evident on inferential questions than for literal questions, though he makes no definitive

claim. As for Erten and Razi (2009), they only state that the effect was seen on overall comprehension, and do not specify which processing level was affected.

Because there was no effect for cultural familiarity of proper names seen in either Study 1 or 2, it cannot be determined what level of processing was affected by proper names. The comprehension questions in both studies related to overall comprehension of the main ideas, as did the free recall task in Study 2. In this regard, targeting overall comprehension might not have been an effective tool for investigating proper name processing. Proper names, as discrete lexical items, are text based. The associative and connotative meaning of a proper name depends on the context. The reader must use context to determine the referent of the name. In this regard, it is worth considering whether an effect for the familiarity of proper names might be seen on lower-level processing. Thus, an alternative approach to investigating the effect of proper names on comprehension would need to ensure that participants draw on their lower-level processing skills, such as lexical access and syntactic parsing skills, to show understanding of the proper names. Such an approach would not focus on global comprehension but on text-based processing. This is the focus of the next chapter: proper names as an aspect of lower-level processing.

Chapter summary

The aim of this chapter was to investigate whether there is an effect of familiarity with proper names on comprehension. To this end, proper names were treated as an aspect of cultural knowledge. Comprehension was compared between texts with L1 culture proper names, L2 culture proper names and, in the second study, no proper names, only common nouns. No effect of culturally familiar proper names was found on comprehension. This result may have been found because global comprehension was targeted, as opposed to inferential or detailed comprehension. Because no effect for proper names was found on higher-level comprehension processes, the next chapter considers whether there is an effect of L2 proper names on lower-level reading skills.

Chapter 6: Proper names and lower-level processing skills

6.1 Introduction

In this investigation into the potential difficulties caused by proper names for L2 readers, the previous chapter considered proper names as an aspect of cultural knowledge. Several robust studies were reviewed that have found an effect of cultural familiarity on L2 reading comprehension. However, it was noted that there is a difficulty in distinguishing whether the observed effects involve higher-level processing (e.g. background knowledge) or lower-level processing (e.g. lexical knowledge). In the two experiments reported in Chapter 5, there was no effect found for cultural familiarity of proper names. This might have been because the effect was sought on overall comprehension (i.e. higher-level processing). For that reason, the focus now shifts from the effect of proper names on higher-level processing to that of lower-level processing (e.g. word recognition and sub-skill processing).

The aim of this chapter is to investigate whether L2 readers might have lower-level processing difficulty with proper names, in particular when there is a large orthographic distance between the L1 and L2. In this respect, L1 transfer effects on L2 lower-level processing are of importance. Just as in the previous two chapters, I start by reviewing research directly relevant to the focus of this chapter: that is, how various aspects of L2 lower-level reading skills, including orthographic processing and word recognition, might impact use of context for proper name processing. Cross-linguistic transfer effects are also considered. The review in section 6.2 is followed by an experimental study that aims to target proper name processing as an aspect of lower-level reading skills. Specifically, the study looks at the extent to which L2 readers can use context to correctly identify proper names.

6.2 Proper names and L2 lower-level processing skills

From the investigation presented so far in this thesis, it has emerged that the difficulty with L2 proper names might lie with lower-level processing skills. To explore this possibility, research is reviewed that demonstrates the importance of efficient lower-level processing for reading comprehension. Next, I look at research focusing on specifically on word recognition and how it is mediated by context; several studies that

investigate L2 readers' use of context to infer word meaning are reviewed. This part of the review directly relates to the assumption discussed in Chapter 4 (section 4.2) that L2 readers can use context to understand and recognise proper names. Lastly, this section ends by summarising the areas of potential difficulties that might exist for lower-level processing of proper names.

6.2.1 Importance of lower-level processes to reading comprehension

Traditionally in L2 reading research, there has been a strong focus on higher-level processes (Bernhardt, 2005). Much of the L2 reading research from the 1970s and 1980s focused on higher-order processing, as it was informed by models such as the Psycholinguistic Guessing Game model (Goodman, 1967, 1988), and schema-driven views of reading (e.g. R. C. Anderson & Pearson, 1984). (Schema theory was reviewed in Chapter 5, section 5.2). As noted in Chapter 1, this focus was due to two beliefs: one, that reading is universal, regardless of language; and two, that lower-level reading skills develop along side L2 oral proficiency. Therefore, any problems with L2 reading comprehension were attributed to higher-level processing. However, in recent years, the importance of lower-level processes in L2 reading has been recognised, and current theories posit reading as an integration of both lower-level (text-based) and higher-level (reader-based) processes (Grabe, 2009). As was outlined in Chapter 2 (see section 2.4), there are three interconnected sets of processes necessary for reading: decoding processes (recognising and accessing meaning from printed text); syntactic and semantic processes (combining information from words into larger units for text comprehension); and text-integration processes (linking text-based information with reader's knowledge for general understanding) (Nassaji, 2014, p. 3).

In his state-of-the-art article, Nassaji (2014) disseminates current theory and research on lower-level processes in L2 reading. Lower-level processes are those "involved in recognizing words, including visual recognition of features, letters, and the use of phonological and orthographic information" (Nassaji, 2014, p. 4). His review includes research done on syntactic and semantic processes and how these processes are connected to word recognition. He also summarises studies focusing on cross-linguistic variables and the effects on L2 reading. Because my interest is in how L2 readers decode proper names in texts, as well as possible cross-linguistic effects, Nassaji's (2014) review of research into lower-level processes in L2 reading is of direct relevance. Here, I summarise key points from his article, while drawing on other research that is pertinent to L2 proper name processing.

Nassaji (2014) explains why efficient lower-level processing has come to be seen as crucial to successful L2 reading. Most contemporary reading researchers view reading from an information-processing perspective (Birch, 2015). In this view, lower-level and higher-level processes are integrated but not reciprocal: lower-level processes can be done without higher-level processes but the reverse is not true. Therefore, no matter how proficient a reader may be at higher-level processes, efficient reading comprehension is not possible without competency in lower-level processes. Nassaji (2014) notes that the importance of lower-level processes in information-processing views of reading has been recognised due to three central psychological principles of reading (p. 4). The first is the notion of 'limited attentional resources': if too much attention capacity is used for lower-level processing like word recognition, fewer resources are available for higher-order processing. The second notion is the hierarchy of skills and sub-skills in reading: in order for higher-level processing to be carried out efficiently, lower-level processing must first be carried out efficiently and accurately. The third is the principle of connectionism: this is a view of information processing as a multi-level system (from letters to words and syntax, to discourse levels). Information processing is not thought to be controlled by or occurring through higher-order processes. Rather it is "mainly text-based and occurs through the processing of visual input, building up information in a network of associations activated textually and enhanced through reciprocal interactions occurring within and between levels" (Nassaji, 2014, p. 5).

Restricted interactive reading models emphasise the importance of automaticity in lower-level processing: the more automatic the processes are, the less likely that they will be influenced by information across levels (Grabe, 2009, p. 90). Thus, these models emphasise mostly bottom-up processing. While efficient word recognition requires interaction between phonological, orthographic, and semantic processing, it does not typically require inferencing or contextual information, for example; in this way, the interaction is restricted (Grabe, 2009). The Verbal Efficiency Model (Perfetti, 1985, 2007) is an example of a restricted interactive model, which predicts efficient word recognition is key to successful reading, and any comprehension difficulties can be attributed to weak word identification skills.

Having established the importance of lower-level processes, Nassaji (2014) goes on to note that while much L2 reading research follows from and is motivated by L1 research (and he refers to research from both fields in his article), there are important

differences between L1 and L2 reading (this point was briefly mentioned in Chapter 2, section 2.4). Unlike L1 readers, L2 readers are often limited in terms of language proficiency. Another difference is that L2 readers are often proficient readers in their L1, thus rendering L2 reading a cross-linguistic activity.[13] Furthermore, L2 readers may have different degrees of socio-cultural knowledge. Nassaji (2014) concludes that because of these differences, "it is crucial to understand how L2 readers process L2 texts, what strategies they use to do so, and how the similarities and differences between L1 and L2 affect L2 reading processes" (p. 5). Similarly, the aim of this investigation is to better understand how L2 readers process proper names, what strategies they use to do so (see Chapter 3), and how differences between the L1 and L2 might affect L2 proper name processing.[14]

If it is the case that L2 readers have lower-level processing difficulties with proper names, what specific skills might be involved? Of the lower-level processes, word recognition is the most critical. Word recognition is "the skill in which readers process the visual symbol in the print in order to recognize and access its meaning in the mental lexicon" (Nassaji, 2014, p. 6). While L1 reading research has clearly demonstrated the importance of word recognition, L2 research has focused more on how these processes differ between L1 and L2 reading, or between L2 readers with different L1s. The few L2 studies that have investigated the role of word recognition in adult L2 readers have demonstrated the critical importance of word recognition efficiency, even in highly proficient L2 users (e.g. Akamatsu, 2003; Nassaji, 2003a; Shiotsu, 2009). (The Akamatsu (2003) study is critically reviewed in section 6.2.2 that follows). Nassaji (2014) notes that these findings are important because they demonstrate the importance of automaticity for L2 reading and furthermore, 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.

---

[13] An example of a robust study that looked at L2 word recognition and the effect of L1 transfer is M. Wang and Koda (2007). They compared word recognition skills of Korean (alphabetic orthography) and Chinese (non-alphabetic orthography) readers of English. Both groups performed similarly to L1 English readers, indicating an effect of L2 learning experiences. However, the Korean readers were faster and more accurate than the Chinese readers on word recognition, suggesting an effect of L1 background on L2 word recognition.

[14] For other research on L1 and L2 activation during language processing, see for example: N. Jiang (2000), Titone, Libben, Mercier, Whitford, and Pivneva (2001), Elston-Güttler, Paulmann, and Kotz (2005)

Of the sub-components of word recognition, two major ones include phonological and orthographic processes. Phonological processing refers to the ability to use knowledge of the sounds of a language to process written words. Orthographic processing refers to using visual information to process words. Nassaji (2014) distinguishes between two types of orthographic visual information in a word: word-specific information, related to the letters, shape and position, and combination of letters; and general information, the combination of letters and patterns that a word shares with other words (p. 12). He notes that orthographic processing has received less attention in L2 research than phonological processing. L1 reading research suggests that orthographic processing skills make significant contributions to word recognition separately from phonological processing; this research also indicates that orthographic processing skills develop with exposure to printed text (Cunningham, Perry, & Stanovich, 2001).

L2 readers, who have much less orthographic processing experience than L1 users, are predicted to have less knowledge of such representations. This lack of knowledge can have a negative impact on word recognition and reading comprehension (Koda, 2004). It follows then that L2 readers might be less proficient in word recognition than L1 readers, given their comparatively limited exposure to the L2 orthography, and some L2 research has indicated as much (Nassaji, 2014, p. 13). Since L2 readers will arguably have less print exposure than L1 readers, it may also be the case that they lack orthographic processing skills to efficiently recognise proper names in continuous text, in particular when there is a large distance between the L1 and L2 orthography. As noted, it has been assumed in L2 vocabulary research that L2 English readers are able to recognise proper names by the initial capital letter (see Chapter 4); however, no research has been done in orthographic processing of L2 proper names (that I am aware of).

6.2.2 Word recognition and context

Related to the potential processing difficulties that L2 readers might have with proper names is the effect of context. In reading research, context can refer to the syntax, semantic relatedness of words, and schemata (mental models) that contribute to determining the meaning of a word (Balota, Pollatsek, & Rayner, 1985, p. 365). It has been assumed that L2 readers can use context to understand and recognise proper names (see Chapter 4). Therefore, it is worth considering the importance of context for L2 proper name processing. The debate on the role of context on word recognition is ongoing: word recognition may be done through context (e.g. Goodman, 1996) or

context might serve to disambiguate at the post-lexical level rather than facilitating word recognition (Stanovich, 2000). According to Grabe (2009), research does seem to support the latter view: fluent readers do not depend on context for word recognition because it is not as fast or efficient as using decoding skills (p. 23). Context does, however, play an important role when readers encounter vocabulary that is not well known or learned (Grabe, 2009; Ryan, 1997), and presumably this would include encounters with unfamiliar proper names.

In considering the link between word recognition and context, Nassaji (2014) suggests that "L2 word recognition and its relationship to L1 orthography may be mediated by the context of reading as well as the nature of the word, including its spelling patterns and frequency" (p. 24). To illustrate, he points to the effects of context seen in two studies conducted by Akamatsu (2002, 2003). In Akamatsu's (2002) study, three groups with different L1s (Japanese, Chinese and Persian) read high frequency and low frequency words out of context. The former words were recognised faster than the latter, and there was almost no difference in performance between the three groups. However, in Akamatsu (2003), the same three language groups read connected passages of visually distorted text (i.e. cAsE aLtErNaTiOn): this time, the Japanese and Chinese readers were more adversely affected by the visual distortion of words in terms of reading time than the Persian readers. Akamatsu's (2003) study will be reviewed in detail here because the findings are illustrative of how word recognition and use of context are interactive processes, and influenced by L1 orthography. In this respect, the study's findings may have implications for L2 proper name processing.

Akamatsu (2003) investigated the effects of L1 orthographic features on L2 word recognition. The author notes that while previous research demonstrated effects of L1 orthographic features on L2 reading, this study differs in that it also takes into account the effect of context on word recognition. This is in contrast to experimental conditions, such as lexical decision tasks, where the processing of words is done without context. Because contextual clues can aid the reader in word recognition, context is an important consideration when examining L2 word processing. In addition, Akamatsu (2003) differs from other research in that pseudo-words and non-words have not been used. The author argues that an L2 user has very little processing experience with pseudo-words and non-words, so such measures might not evaluate word processing abilities but rather a skill developed through reading experience (Akamatsu, 2003, p. 211).

The study design focused on word recognition, higher-order processing skills and L1 background. First, word recognition was examined using case manipulation (i.e. cAsE aLtErNaTiOn), causing a condition where words lose shape information but retain spelling patterns. The concept is that if a reader can efficiently process individual letters, she should be less adversely affected by case manipulation. Akamatsu (2003) defines efficiency in word recognition as "the ability to process a word with sensitivity toward its constituent letters" (p. 212). Second, higher-order processing skills were tested using reading passages with comprehension questions. Third, the study compared fluent L2 English readers from three L1 backgrounds: Chinese, Japanese (both non-alphabetic L1s) and Persian (alphabetic L1). The hypothesis was that the non-alphabetic L1 readers would be more adversely affected by case manipulation than the alphabetic L1 readers because the former are more dependent on word shape information during word processing and are not as sensitive to intra-word information. This is because in logographic writing systems, it is not necessary, or always possible, to amalgamate intra-word components to retrieve phonological representations.

The sample was 49 fluent English L2 readers (18 Chinese, 16 Japanese and 15 Persian native speakers). Twelve reading passages were used (110-150 words each) with three levels of difficulty: easy, moderate and difficult. Half of the passages were presented in normal case format; the other half were presented in aLtErNaTeD cAsE. Four to six reading comprehension questions (multiple choice) were used for each passage. The experiment consisted of two separate sessions: in each session, the participant was presented with six passages, two from each level of difficulty, one in normal case and the other in alternated case. Passages were presented on a computer screen. Instructions were to read silently for comprehension as quickly as possible. Reading time spent on each passage was recorded. Once reading comprehension questions were accessed, participants could not go back and check the text. Comprehension scores were also recorded. The interval between sessions was five to ten days. No significant differences between data from the different sessions were found, so the data was combined for analysis.

The data analysis (MANOVA and planned comparisons) showed that normal case passages were read faster than alternated case passages. Easy passages were read faster than moderate ones, and moderates ones were read faster than difficult ones. Interaction between L1 and case was significant in reading time: Chinese and Japanese readers took longer to read the alternated case passages than the Persian readers, suggesting that the former were more adversely affected by case alternation

than the latter. The author interprets this finding to reflect that readers with a non-alphabetic L1 are less efficient in processing individual letters of a word. Also regarding reading time, there was interaction between case and text difficulty, suggesting the visual distortion of words creates a greater cognitive load. Analyses of reading comprehension scores showed no significant interactions between L1 and case, which the author suggests is an indication that while L1 orthographic features may affect L2 word recognition efficiency, these features might not affect reading comprehension. However, reading comprehension scores were affected by case alternation: for all groups, reading comprehension was lower for alternated case passages than for normal case.

This study seems to have a robust design, and while the sample size is small, the author was able to combine data from the two sessions for analysis, as noted above. Akamatsu (2003) notes that his findings are supported by previous studies which found that non-alphabetic L1 readers are less sensitive to L2 intra-word information (Koda, 1999; Muljani, Koda, & Moates, 1998), and that these readers are less efficient in processing individual letters in an L2 word (Akamatsu, 1999; T. Brown & Haynes, 1985). Akamatsu's (2003) main claim, that fluent ESL readers with a non-alphabetic L1 are less efficient in processing component letters of words in a text than readers with an alphabetic L1, has important implications for this investigation into L2 proper name processing. For example, his findings suggest that non-alphabetic L1 readers of English may be less efficient in processing individual letters of proper names in contextual conditions. Thus, it may not be the case that L2 readers, particularly those with a non-alphabetic L1, can consistently notice and recognise the initial capital letter of English proper names in context.

There were proper names in the twelve reading passages used in Akamatsu (2003). In a footnote, the author records that for the alternated case passages, proper names and scientific terminology were kept in normal case (p. 225). While the author does not explain the reason for this decision, and the reader is given only one sample text in the appendix, one can speculate it was because these are low frequency items and therefore, most likely unfamiliar even for these fluent ESL readers. Similar to processing of pseudo-words and non-words, the participants would have less processing experience with proper names and scientific terminology. However, given the importance of the effect of context on word recognition that the author acknowledged in his study design, it does seem that this decision warrants consideration. For instance, in the appendix, one example of a moderate level text

about trees is presented in normal and alternated case. This text contains two proper names, *Sierra* and *Big Redwood*, and each name is mentioned only once. In the alternated case passage, we see these two sentences (p. 230):

(1) MoSt Of ThE Sierra tReEs DiE oF dIsEaSe. . . . BuT nOtHiNg HuRtS tHe Big Redwood.

Akamatsu's (2003) decision to keep the proper names in normal case suggests that he deems the context insufficient to aid the L2 reader, at least with regard to proper names, and that the initial capital letter is a necessary cue for the processing of proper names.

One last comment needs to be made concerning the grouping in this study of the Chinese and Japanese readers as non-alphabetic L1 readers. In doing so, Akamatsu (2003) has perhaps inadvertently overlooked interesting data in his experiment regarding differences and similarities among the three L1 groups. Chinese are L1 logographic readers while Japanese are both logographic readers of Kanji (Chinese characters) and syllabic readers of Hiragana and Katakana (syllabic writing systems). In this way, Japanese might share processing experience with both Persian readers and Chinese readers: logography is a meaning-based writing system, while syllabic and phoneme-based writing systems, like Persian, are both sound-based systems (Cook & Bassetti, 2005, p. 5). (In fact, Cook and Bassetti define Arabic script, which is the script that Persian is based on, as consonantal and distinct from alphabetic scripts). A glance at Akamatsu's (2003) data suggests that the Japanese might share similarities with both the Chinese and Persian readers. For example, for the alternated case passages, the Persian and Japanese readers had similar mean reading times (178.3 seconds and 183.3, respectively) while the Chinese were slower (195.6). While the author's assertion that the non-alphabetic readers were more adversely affected by the case manipulation is correct, it is nevertheless noteworthy that the Persian and Japanese readers read the manipulated case passages at nearly the same mean rate. In sum, Japanese readers may have unique L1 processing experience given their two distinct writing systems (i.e. logography and syllabary). This is an important consideration to bear in mind for the findings from my investigation into proper names, given that the participant group is L1 Japanese. Having considered how a reader's L1 might impact L2 word recognition, I now turn to several studies that have investigated how L2 readers use context to infer word meaning.

6.2.3 L2 readers' use of context

As was reviewed in Chapter 4, an assumption made by L2 vocabulary researchers is that L2 readers can use context to infer the meaning or referents of proper names. To recap briefly, Hwang and Nation (1989) argue that proper names in newspaper articles are often explained in context (e.g. *Prime Minister Jacques Chirac*) (p. 324). Similarly, Hirsh and Nation (1992) reason that the function of names in stories will signal to the reader that these vocabulary items are proper names (p. 691). Nation and Wang (1999) also remark that "proper nouns could be easily understood from context and should not be counted as unknown vocabulary" (p. 358). Later studies, such as Matsuoka and Hirsh (2010), repeat the assumption from the earlier studies in the 1980s and 1990s, that L2 readers can use context to infer the meaning of proper names, as justification for treating proper names as known vocabulary. However, no studies (that I am aware of) have directly investigated this assumption. Implicit in this assumption is the belief that L2 readers have efficient word recognition skills necessary for successful higher-level processing, such as use of context and inferencing, to take place. However, as seen in the discussion thus far, it may be misleading to assume that L2 proficiency equates with efficient lower-level processing skills.

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 these 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. The learners read a text in which 95% of the vocabulary was considered known with ten target unknown lexical items. Analysis of the learners' responses showed that they were not able to successfully 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 used think-aloud protocol 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 contextual clues: when their participants had mistakenly identified a word, they did not use 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 more than non-logographic readers (Akamatsu, 2003; Ehrich et al., 2013; Koda, 2005). As was discussed above, 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.

Similarly, Ryan (1997) notes that using context to infer meaning may prove an ineffective reading strategy for Arabic readers who have difficulties at the word form level due to influence of their L1. Arabic is a consonantal language (Cook & Bassetti, 2005); word roots are based on consonants, and vowels can be, and often are, omitted in formal writing. The L1 Arabic reader learns to focus on word roots (i.e. consonant patterns) for word recognition. This lower-level processing might then be transferred to L2 English reading, which results in difficulty in distinguishing between English words with similar consonant patterns (Ryan, 1997, p. 186). Thus, for Arabic readers who are prone to such errors in lexical decoding, using context to infer the meaning of unknown lexis may be an ineffectual, and even inappropriate, reading strategy (Ryan, 1997, p. 187).

6.2.4 Summary

The findings from these various studies provide warranting for investigating whether context is sufficient for L2 readers to infer the meaning of proper names. Again, this may be especially relevant for readers for whom there is a large orthographic distance between the L1 and L2. Nassaji (2014) draws four conclusions from his state-of-the-art review of L2 lower-level processing research: first, accurate and efficient lower-level processing is critical to skilled L2 reading comprehension; second, word recognition is mainly a lower-level process and less affected by syntactic and semantic skills than previously thought; third, cross-linguistic research has found similarities and differences in reading sub-skills in L1 and L2 and there can be positive and negative effects, depending on the distance between the two languages; fourth, L1 orthography, reading experience, L2 proficiency, reading tasks and context, the nature of a word and its frequency can all impact strategies and information the L2 reader uses for word recognition (p. 24).[15]

Nassaji's (2014) review contains several important considerations relevant to my interest in lower-level processing of L2 proper names. The first concerns the assumption that word-recognition skills develop as a result of L2 language proficiency (Nassaji, 2014, p. 8). This assumption may be based on the strong relationship seen between L2 reading comprehension and L2 proficiency. This line of reasoning, driven

[15] For psycholinguistic research on top-down and bottom-up information used in word recognition, see for example: Dijkstra (2005), Cutting and Scarborough (2006), Perea, Marcet, and Vergara-Martinez (2016).

by top-down theories and schema-theoretic views of reading, contends that when highly proficient L2 users read more slowly than L1 readers, it is because they are weaker at higher-level processing skills. However, given the fact that many L2 readers are already proficient as L1 readers, it seems very unlikely that their difficulty lies with higher-level processing (Nassaji, 2014). An alternative possibility is that the weakness lies in word-recognition automaticity: more cognitive capacity is spent on word recognition, not allowing higher-level processing to occur as fluently (Nassaji, 2002). As has been noted, an information-processing view of reading postulates that efficient lower-level processing is critical to higher-level comprehension processing (Grabe, 2009; Perfetti, 2007).

The same assumption regarding the correlation between word-recognition skills and proficiency is made in L2 vocabulary literature concerning L2 readers' ability to recognise and understand proper names. For example, Webb and Macalister (2013) propose that because proper names are relatively frequent in graded readers (comprising 3% or more of the vocabulary), L2 readers "may quickly develop the skills to recognise and understand proper nouns to some degree. This is likely to allow readers to process the text more easily when they encounter unknown proper nouns rather than other unknown words" (p. 311). As noted, graded readers are simplified texts, with vocabulary graded according to the learner's proficiency level. Low to intermediate learners are the usual consumers of such reading materials. Thus, Webb & Macalister (2013) are assuming that lower proficiency readers will quickly hone their word-recognition skills to identify names because of the frequency at which they will encounter unfamiliar names. But it is not clear why lower proficiency readers possess efficient decoding skills to quickly identify names in a connected text. As was reviewed above, studies have found that even highly proficient users are lacking in efficient lower-level processing skills (Nassaji, 2003a; Shiotsu, 2009). Thus, it seems incautious to assume lower proficiency users would be skilled in decoding processes.

Another important consideration for this investigation into 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). As noted above, 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 regard 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. 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. However, 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. How well L2 readers can use context to identify proper names is investigated in the next section.

6.3 Study: L2 readers’ use of context to identify proper names

In order to learn more about the lower-level processing difficulties L2 readers might have with proper names, I conducted a study to investigate the extent to which L2 readers of English are able to use context to correctly identify target proper names. The study fills a gap left by other research literature, which for the most part, has assumed that context is sufficient for L2 readers to identify proper name referents. 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). In this way, use of context to infer meaning illustrates the interactive processes of reading: the text-based semantic and syntactic

information (lower-level processing) interacts with the reader's knowledge (higher-level processing) for interpretation of the text.

L2 readers' lower-level processing skills are targeted in the study's design. Semantic processing has been disrupted: the target proper names are semantically ambiguous in that they can be used as names but also have a common word meaning (e.g. *Rose, rose)*. Also, orthographic processing has been disrupted: all initial sentence and mid-sentence capital letters have been changed to lower case. Thus, the reader must rely on the context of the sentence to determine if the target item is a proper name or a common word.

The study aims to answer this research question:

> To what extent can Japanese intermediate L2 readers of English use sentential context to correctly identify target proper names in authentic sentences?

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 an assumption found in L2 vocabulary research that proper names can be easily recognised and 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).

6.3.1 Participants

There were a total of 61 participants (females, 51; males, 10). All participants were Japanese first-year university students and 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 women-only 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: intermediate or A2/B1 on the CEFR. The Business majors had combined TOEIC scores ranging from 225 to 400; the

Education majors had TOEFL scores ranging from 420 to 470. These scores place participants on the cusp of the A2/B1 band.

6.3.2 Materials

Twenty target names were selected that were ambiguous in meaning; that is, the target items could refer to a person's name and had another common word meaning. Twenty items were selected in order to provide an adequate sample size. The target proper name items and part of speech of the non-name items were: Rose/rose (verb); Mark/mark (noun); Jack/jack (noun); White/white (adjective); Hill/hill (noun); Brown/brown (adjective); Major/major (adjective); Bill/bill (noun); Young/young (adjective); Green/green (adjective); Grace/grace (noun); Wood/wood (noun); Frank/frank (adjective); March/march (verb); Nick/nick (noun); Cook/cook (noun); Mike/mike (noun); Bob/bob (verb); Pat/pat (noun); and Cliff/cliff (noun). These particular items were chosen for their common word frequency ranking. The first fifteen items in the list are high frequency words (i.e. the non-name forms appear in the first 3,000 words on the Common Core List of the BNC and COCA corpus (Cobb, n.d.). The last five items are off-list on the Common Core List, and as such, might be less familiar to the participants in this study.

Authentic sentences for each target item were selected from corpora in order to eliminate a bias that may have resulted if the researcher had generated the sentences. Authentic sentences were selected from either the BNC (Davies, 2004-) or the COCA (Davies, 2008-) corpora to include both British and North American registers of English. The Japanese participants in this study may have had different exposure to various varieties of English throughout their learning experiences with foreign teachers in Japan or from study abroad experiences, so these two varieties were included in the experiment. Sentences for about half the items (*mark, hill, brown, major, bill, young, grace, green, white, wood, jack, rose*) were taken from the BNC corpus; for the other items (*frank, march, nick, cook, mike, bob, cliff, pat*), sentences were chosen from the COCA corpus. The researcher selected two sentences for each target item, one which had the proper name usage and the other which had the common word meaning. The researcher selected pairs of sentences that were similar in word length; as the sentences were authentic, exact word length matches were not possible.

Two tasks (named Task A and Task B) were created using the selected authentic sentences. Each task consisted of 20 sentences with a target item appearing only once, either as a proper name or as a common word. Two tasks were used so that the

experiment could be repeated; also, this ensured that each target item would be appear only once in each task. Ten sentences in each task 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. (See Appendix 5.1 for participants' versions of Task A and Task B).

Both tasks had the same instructions to the participants, written in both English and Japanese (L1). The instructions in English 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 task sheet, illustrating the instructions to the participants:

I *N* *Y*

Example: ~~i~~ arrived in ~~n~~ew ~~y~~ork last night.

6.3.3 Procedure

As noted, participants were members of intact classes at two universities, and the experiment took place during class time. The participants were asked if they agreed to take part in the experiment. A consent form was distributed which explained that participation was not obligatory and their personal data would be kept confidential. The consent form also explained what participating in the experiment involved. This information was presented in English and Japanese. (See Appendix 5.2 for the consent form). All participants agreed to take part and signed the forms.

After the consent forms were collected, the Business majors (n = 29) were given Task B to complete; the Education majors (n = 32) were given Task A to complete. Different tasks were given to the groups to control for a practice effect (i.e. in case higher scores were seen at the second test time). Instructions were read aloud by the researcher (and teacher at the second university). 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 reminded that they could use their dictionaries. Because the sentences were authentic

and the participants had an intermediate proficiency level, it was important that they were allowed and encouraged to use dictionaries. Most participants did use their dictionaries, and most finished the task in five minutes. However, the task papers were not collected until 15 minutes had passed to ensure no one felt time pressure. Participants who had finished were encouraged to check their work and allowed to look at course materials while waiting for others to finish.

One week later, the Business majors were given Task A to complete and the Education majors were given Task B. The instructions, example and capitalisation rules were reviewed. Participants were given 15 minutes to complete the task and again encouraged to use dictionaries. When the task materials were collected, a debriefing sheet was distributed which thanked the participants for their efforts and explained the purpose of the experiment (i.e. to investigate if the context, or other words in the sentence, could help them correctly identify the proper names in the sentences). The debriefing sheet also reminded them that they were free to withdraw from the study, and provided the researcher's email address should they want to ask for more information. (See Appendix 5.3 for the debriefing sheet).

6.3.4 Data analysis

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'). Any errors relating to capitalisation of non-target items were ignored. Data analysis compared participants' responses on tasks 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 tasks was removed in this initial analysis).

6.4 Results

Descriptive statistics were compiled for correct responses to target names and target non-names in both tasks. Table 6.1 compares the means for correct responses to target names (total 10) and target non-names (total 10) in Tasks A and B. The descriptive statistics for the tasks were first analysed separately to check for any

practice effect. The scores in the mean, minimum and maximum columns show that responses were similar in Tasks A and B for both target names and non-names.

Table 6.1

Descriptive statistics: correct responses by target items and task

		Task A				Task B			
	n	*Min.*	*Max.*	*M*	*SD*	*Min.*	*Max.*	*M*	*SD*
Target names	54	0	6	2.54	1.723	0	7	3.24	1.863
Target non-names	54	6	10	9.07	.949	6	10	9.24	.845

Because the responses were similar across tasks, participants' scores from Task 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. Table 6.2 shows the descriptive statistics for the total correct target names and non-names (out of 20).

Table 6.2

Descriptive statistics: correct responses by target items from both tasks

	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 context to correctly identify proper names in authentic sentences), the correct responses to target names and target non-names were compared. 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, $p < .0005$) 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 ($r = 0.87$). (The effect sizes (r) are calculated by dividing the Z-score by the square root of the total positive and negative ranks, but not the ties). 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.

6.5 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 Task A (60 x 10 = 600); 55 participants responded to ten target names on Task 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 Task A (60 x 10 = 600); 55 participants responded to ten target non-names on Task B (55 x 10 = 550). The total data set was 1,150 responses to target non-names. Correct responses to each target names and non-names are summarised in Table 6.3.

The total row shows the correct number of responses to all target names in both tasks 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 Task A and B was different (60 and 55, respectively), a scaling was done to the number of responses to Task B (i.e. dividing by 55 and multiplying by 60) to allow for comparison.

Possible factors were identified that may have affected participants' ability to correctly identify target names: the richness of sentential context, the frequency of the target name, and the part of speech associated with the corresponding target non-name. A multiple regression was carried out to investigate if there is a relationship between correct responses to target names and the explanatory variables of: context, proper name frequency ranking and the most frequent part of speech of the corresponding non-name. Data for this post-hoc analysis was gathered in three parts.

Table 6.3

Correct identification of target names and non-names in Tasks A and B

Target Names	Total number of responses	Correct responses		Target Non-names	Total number of responses	Correct responses	
		n	%			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 (Task A)		**148**	**24.7**	**Subtotal** (Task B)		**507**	**92.2**
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 (Task B)		**179**	**32.5**	**Subtotal** (Task A)		**544**	**90.7**
Total A+B	*n*, mean %	**327**	**28.4**			**1051**	**91.4**

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 (see Appendix 5.4 for the context rater sentences). 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 6.4 shows the L1 users' response types and tallies listed by the target name (which had been replaced by a blank).

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 Task A is copied below:

(2) 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.

Table 6.4

Context rater responses by target name

Target name (replaced by blank)	**Response types and tally** N = 6 (adult L1 English)	**Context rating** 6 = rich, 0 = no context
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, unsurprisingly	5
Brown	he (5), Walter	6
(John) Major	Smith, had, Major (2), diligently, confidently	3
Bill	myself (3), her, him, Sr. Lopez	6
(Sean) Young	Connery (4), Penn (2)	6
Green	she (2), Maggie, Elizabeth, Lord Byron, I	6
(Martin) Grace	Shaw (2), Grace, Scorsese, Sheen, Short	6
(Brian) Wood	Robson, Gibson, Jones, Khan, Orser, More	6
Frank	Peter, he (3), David Suzuki, Dick	6
(the) March (hearing)	committee, court (2), arraignment, public, preliminary	0
Nick	Stephen, him (4), the DA,	6
Cook	Smith, then, inadvertently, enthusiastically, O'Toole, Pan	3
Mike	Peter, him (2), Albert, herself, Brad	6
Bob	he (3), she (2), Trump,	6
Pat	you (5), Julia	6

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). This list had been compiled as a straightforward frequency count of any word tagged as a NPO (proper noun) (personal communication, Parent, 2016). Table 6.5 shows the frequency ranking of each target name.

Table 6.5

Frequency ranking of target proper names in the BNC

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

From the table, 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 non-name 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 6.6 shows the parts of speech of the target non-name items in order of frequency.

From Table 6.6, 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 task sentences for non-name items, 14 sentences used the most frequent part of speech (*jack, grace, nick, mike, cliff, bill, wood; march, bob; major, green, frank, brown, young*).

Table 6.6

Frequency of parts of speech for target non-name items

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

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.

6.5.1 Post-hoc analysis findings

Findings from the standard multiple regression analysis are reported here. When correctly identified names were predicted, it was found that none of the predictors were statistically significant: context rating (ß = .361, n.s.) and frequency ranking (ß = .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 the occur most frequently as nouns was not predicted (ß = -.169, n.s.). Comparing items that occur most frequently as verbs to items that occur most frequently as nouns was also not predicted (ß = .008, n.s.). The overall model fit was R^2 = .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.

6.6 Discussion

The findings suggest that Japanese L2 intermediate readers of English are not very successful at using 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 < .0005), 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 infer the meaning (or referents) of proper names (see discussion in Chapter 4, section 4.2). The results from this study 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. Perhaps because most of the names were frequent, there was little differentiation seen in the model. The target items were chosen for this study because they were common words that the participants would be familiar with. However, in order to find an effect in terms of item frequency, it would have been better to include a balance of familiar names and less familiar names (like *Cliff*).

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 was seen in other studies reviewed in section 6.2 (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 6.3), 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* (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:

(3) 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. (The sentences were authentic, and there was no intention to mislead participants). 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. The target item appeared in test item #3:

(4) 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*, for which the contextual clues seem to strongly favour a non-name response:

(5) 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 (in section 6.2) 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 conceivable. The sample size in this study (N = 61) was better in terms of quantitative analysis, though certainly, qualitative investigation, such as think-aloud procedures, would have allowed more insights into the inferencing strategies used by participants.

Another limitation was that the part of speech of the corresponding non-name target items was not controlled for. As was noted in the post-hoc analysis, nine of the items occur only as nouns, or most frequently as nouns. Five items occur most frequently as verbs, and the other six items occur most frequently as adjectives. One might predict that the part of speech in which the target item appears most frequently would be the form that the L2 reader would be most familiar with (Hoey, 2007). Hence, if the corresponding non-name item appears most frequently as verbs or adjectives, this might have resulted in participants being less able to identify the target item as a name. This is because the reader is less familiar with seeing the item in the noun position. The results from the multiple regression were not statistically significant, however. In this regard, this aspect of the study design could have been better balanced, for example, by comparing an equal number of target non-name items that appear most frequently as nouns with those that appear as verbs.

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. However, some might argue that the study design did not represent an authentic reading task in that the initial capital letter that might normally provide a clue to the reader about proper names was absent. However, seen in the review (section 6.2), it might be misleading to assume L2 readers have efficient orthographic skills to process upper and lower cases equally well; this is particularly true for those L2 readers processing a different writing system (Alderson,

2000). Indeed, the assumption that the initial capital letter on proper names serves as an orthographic cue for L2 readers requires empirical support.

Some might also argue that because of the study design, the participants could use the context of only one sentence to determine which items were names; for this reason, the findings from this study might not be applicable to authentic reading. For example, in a news article or story, a proper name might be mentioned several times; at each mention, the reader has more opportunities to build up an understanding of the referent. While this is a valid point, it bears repeating that the participants were given ample time (15 minutes) to analyse the 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.

6.7 Conclusion

The focus of this chapter has been on proper name processing as an aspect of lower-level reading skills. 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. 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 from here certainly 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. Several suggestions were given how to improve the study design in order to explore potential predictors, such as context richness, frequency and parts of speech. Using slightly longer texts (i.e. more than one sentence) may also reflect a more authentic reading experience. In this way, it could be investigated how readers build meaning of proper names as they progress through a text. Qualitative approaches, such as think-aloud protocol, could also reveal more about how L2 readers make sense of the proper names they encounter.

If L2 readers cannot recognise and identify most proper names in context, then this has implications both for researchers and teachers. For example, with regard to L2 vocabulary 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) (e.g. Nation, 2006). 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.

For L2 reading teachers, there are also clear pedagogical implications. First, it should not be assumed that names are low-burden items that will take care of themselves. Overt attention should be drawn to proper names in a text, including any associations and connotations of proper names in literary texts. As a part of pre-reading activities, students can be asked to scan for names in a text and indicate what they already know about those entities. For names of which they have no prior knowledge, these cases could be used to illustrate how to exploit any contextual clues for additional information about the referent (e.g. pronouns, titles, positive or negative adjectives, etc.). Strategies for handling unfamiliar proper names could be illustrated in the classroom as well, for instance, by doing online searches or checking learner dictionaries. These sorts of activities can help L2 readers understand the importance of understanding proper name referents to aid text comprehension. Implications for the L2 classroom are discussed further in the Discussion chapter that follows.

Chapter summary

This chapter looked at proper name processing as an aspect of lower-level reading skills. In particular, the following factors relevant to proper names processing were looked at: orthographic processing, word recognition, use of context, and effects of cross-linguistic transfer. I reported on a study I conducted which looked at the extent that L2 readers can use context to identify proper names. The findings indicated that L2 readers are not very successful at identifying proper names from context, contrary to an assumption found in L2 vocabulary research. Implications of these findings for

research and pedagogy were briefly discussed. In general, researchers and teachers should not assume L2 readers can understand proper names from context.

The next chapter (Discussion) revisits the overarching aim of this investigation into the potential difficulties that proper names might cause for L2 readers. With reference to the experimental chapters (3 to 6), I look at what conclusions can be drawn and what questions remain. Implications of the findings from the five studies for language research, pedagogy and materials developers are discussed.

Chapter 7: Discussion

7.1 Introduction

The experimental work presented in Chapters 3 to 6 concentrated on some different aspects of proper name processing. In the first study (Chapter 3), interviews were conducted to explore L2 readers' perspectives on proper names (hereafter, the Interview Study). The interviewees reported various difficulties related to identifying L2 proper name referents and discussed strategies for handling unknown names. Some of the questions raised from the interview data were used as a springboard to explore proper name processing in subsequent experiments. Specifically, three issues identified for further investigation concerned: whether the findings were generalisable to a larger sample; the effect of proper name familiarity on comprehension; and the usefulness of context for identifying proper name referents. First, because the sampling for the interviews had been purposive, the next consideration was whether a random sample of participants would demonstrate similar difficulties to identify and understand proper name referents. To that end, the next experiment (Chapter 4) looked at how a larger sample of L2 readers approached proper names in a reading task for which overt attention was not drawn to issue of proper names. Participants were asked to mark unknown vocabulary in reading texts with three levels of difficulty (hereafter, the Unknown Vocabulary Study). It was found that, contrary to an assumption found in some L2 vocabulary research, some L2 readers do treat proper names as unknown vocabulary. Furthermore, they list proper names as items to look up in a reference source, indicating that for these participants, understanding proper name referents is considered important to text comprehension. Also, more proper names were listed to look up in the easy text than in the more challenging texts, which suggests L2 readers give attention to proper names when most of the other lexis is known.

A second question raised from the Interview Study concerned the participants' lack of experience with certain proper names, relating to the gender of names, personal and family names, nicknames and place names. Given that their difficulties seem to stem from their unfamiliarity with L2 proper names, the next question was whether there is an effect for proper name familiarity on reading comprehension. Specifically, the two experiments from Chapter 5 investigated whether cultural familiarity with proper names aids reading comprehension. The first study was a replication, which compared reading

comprehension of texts with culturally familiar and unfamiliar proper names (hereafter, the Cultural Study). No effect on reading comprehension was seen, perhaps due to the lexical load of the text. A second study was carried out to address the lexical difficulty of the text and compared three treatments: proper names from the L1 culture, proper names from the L2 culture, and no proper names, only common nouns (hereafter, the Three Treatments Study). No statistically significant difference was seen among the three treatments, perhaps because global comprehension was targeted, as opposed to detailed understanding.

In that no effect of proper names on higher-level comprehension processes was found, the last issue considered L2 proper name processing in light of lower-level reading skills. There had also been some indications from the Interview Study data (Chapter 3) that context is not always reliable for inferring proper name referents. For those reasons, the final experiment (Chapter 6) looked at participants' ability to identify proper names from sentential context (hereafter, the Context Study). Semantic and orthographic processing was disrupted so that participants needed to rely on context to identify proper names. The results indicated that the participants were not very successful at using context to identify proper names, providing further contradictory evidence for an assumption that L2 readers can recognise and understand names in context.

The following perspectives related to proper name processing are discussed in this chapter, based on the literature reviews and experimental work from chapters 3 to 6: how L2 readers view and approach proper names; how cultural familiarity of proper names affects comprehension; and how L2 lower-level reading skills impact proper name processing. First, the success of each of the experiments from chapters 3 to 6 in achieving the aims is considered. This is followed by an evaluation of the theoretical orientation, research design and overall success of the studies. Then, the degree to which the findings from this specific research context can be generalised to answer the central research aim of whether proper names burden L2 readers is examined. Finally, implications of the research findings for L2 vocabulary research, pedagogy, and materials development are considered.

7.2 L2 readers' perspectives on proper names

How L2 readers view proper names was the focus of the Interview Study in Chapter 3. As recounted in the introductory chapter, classroom experience contradicted

assumptions made in L2 vocabulary literature regarding proper names. Therefore, the aims of the Interview Study were to explore: how L2 readers feel when encountering new proper names; what strategies they use for unknown names; and what they find problematic about proper names. The first research question looked at what affective factors are involved in proper name processing, for example, whether readers feel apathetic or anxious about unknown proper names. How L2 readers feel about meeting unknown proper names was of interest because if they recognise the unfamiliar items are proper names, but are not concerned with understanding the referents, then perhaps it is reasonable to treat proper names as low-burden items. On the other hand, if unfamiliar names cause anxiety related to text comprehension, then this finding would have implications for testing situations, for instance.

The interviewees reported being confused about proper names with regard to the gender of names, and distinguishing between personal and family names. They also reported difficulty in identifying referents of nicknames and unfamiliar place names. Some participants noted that being familiar with proper names could aid comprehension by helping them to build images of characters in a story. They also mentioned some frustration at not knowing how to pronounce new names. This might be an indication of their inability to apply their knowledge of English phonology to words they have not encountered before (Birch, 2007). Thus, they might require training in recoding (i.e. forming auditory representations of words). Alternatively, this finding might lend support to the plausible phonology hypothesis (Brennen, 1993). This theory suggests proper names are more difficult to learn or recall because there are more plausible phonological sequences for names than other words.

Participants' strategy use for handling unknown names was the focus of the second research question. Strategy use could indicate processing difficulties. Participants mentioned using different strategies for L2 proper names including: ignoring them; doing online or dictionary searches; using grammatical or orthographical clues; and referring to or making their own name charts to keep track of various characters in a story. Given that these participants had several approaches to dealing with proper names suggests these items are not wholly unproblematic. For instance, two participants reported that when they did not know how to pronounce a name, they focused on the initial letter of the name in order to differentiate between characters. This strategy seems to be an example of L1 processing transfer to L2 reading, whereby L1 logographic readers focus on the visual appearance of unpronounceable words (Koda, 1995). Participants also demonstrated curiosity about proper names by

looking them up in dictionaries or online. In sum, their strategy use suggests that they are seeking ways to deal with an incomplete understanding of proper names for text comprehension.

Identifying decoding or comprehension difficulties related to proper names, whether perceived by the participants or demonstrated in the read-aloud task, was the aim of the third research question. Participants self-reported on various difficulties they had with proper names, such as pronouncing unfamiliar names and identifying referents of nicknames and place names. Also, in the read-aloud task, participants demonstrated problems in understanding referents for which there were not explicit contextual clues. In addition, two participants seem to have had difficulty in differentiating between two characters with similar looking names.

7.2.1 Evaluation of the Interview Study

The study explored the soundness of an assumption that L2 readers can recognise and understand proper names by probing a small sample of participants for their perceptions of, and feelings about, proper names. Typological analysis was used to analyse the data, drawing on the research questions for typologies (Hatch, 2002). Typological analysis proved to be an optimal approach for analysing this data because the participants did not deviate from the interview schedule. This made the analysis straightforward and focused on the research questions.

The data generated from the Interview Study raised several interesting questions about how L2 proper names are processed, and in that respect, the study was successful overall. First, perhaps the most obvious question was whether the findings from this small study were applicable to the larger population; that is, whether a random sample of Japanese L2 readers of English would demonstrate difficulty with proper names while reading. This question was used to motivate the Unknown Vocabulary Study, which used a larger sample to investigate how L2 readers approach proper names, utilising a reading task that did not draw explicit attention to proper names. Secondly, from the interview data, it seemed that in general, participants' difficulties resulted from a lack of familiarity with certain proper names. For that reason, the effect of familiarity with proper names on reading comprehension was identified as worthy of further investigation; this effect was explored in the Cultural Study and the Three Treatments Study. Lastly, while there were some issues with the read-aloud task, it was observed that context might not always be reliable for inferring proper name referents. This

observation sparked ideas for the Context Study, in which L2 readers' use of context to identify proper names was investigated. In sum, the Interview Study was successful in demonstrating various difficulties L2 readers can have with proper name processing. Some observations drawn from the Interview Study were used to drive the investigation into L2 proper name processing forward.

Some limitations are noted with respect to the research design and the study's success in meeting its aims. First, with regard to the participants, the sampling was purposive; that is, participants were chosen that had previously and independently remarked on having difficulties with proper names. This sampling method may have impacted the findings. Also, the sample size was small, as can be the case in qualitative research. However, as the study was exploratory in nature, it was thought that four participants would provide sufficient data. The quantity of data that was collected was limited due to the oral proficiency level of the participants. In addition, participants may have felt nervous about being interviewed and having their responses recorded. Two of the participants in particular were hesitant to elaborate on their answers. As noted, if participants had been given the interview schedule in advance, they might have had time to consider and prepare their answers. Alternatively, conducting the interviews in the L1 might also have generated fuller responses. A second limitation concerns the read-aloud task, which was not ideal for capturing difficulties related to proper name processing. Intentional oral reading requires processing and attention that can adversely impact comprehension (Birch, 2007). Also, participants were most likely anxious about reading aloud and being recorded; this anxiety would have taken away from their attention to the task. Thus, it is very difficult to draw any conclusions about the decoding and inferencing errors that were observed. Self-paced reading and eye tracking are two methods that might be used more successfully to this purpose.

7.3 How L2 readers approach proper names

Determining how L2 readers approach proper names in a reading text is important because of a widely held assumption in L2 vocabulary research that proper names are low-burden vocabulary items. Consequently, it has become standard practice in text coverage counts to either re-categorise proper names as known (i.e. 1K band items) or to remove them altogether from the analysis (see D. Brown, 2010). Since proper names can make up a considerable percentage of vocabulary in any given text (3 to 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. As noted, a more recent trend seen in some studies is to remove proper names from the analysis altogether, making it impossible to see how much of the text is comprised of proper names (e.g. Uden et al., 2014). These issues were considered in Chapter 4. The Unknown Vocabulary Study looked at how L2 readers treat proper names while reading; that is, to ascertain whether proper names are indeed understood by L2 readers, or whether proper names might present a source of processing difficulty.

The Unknown Vocabulary Study approached this aim by asking participants to mark unknown vocabulary while reading texts of varying levels of difficulty. Research question 1 looked at whether L2 readers mark proper names as unknown vocabulary, with text difficulty as a variable. That is, if all other vocabulary in a text is known, how do L2 readers handle proper names; conversely, when more vocabulary is unknown, how are proper names treated. Text difficulty was considered an important variable because if a greater percentage of words in a text are known, it seems likely that the reader will have a better chance of correctly inferring the meaning of other unknown vocabulary, including proper names. Research supports this idea, indicating that when 95% of the words in a text is known, there is greater success in making correct inferences about unknown vocabulary (Laufer, 1989; Liu & Nation, 1985). On that basis, the effect of text difficulty on proper name processing was examined. It was predicted that for texts in which most vocabulary is known, proper names might not present a burden because the reader is well-supported by context. For texts that have a larger percentage of unknown vocabulary, proper name referents might not be easily inferred, and thus present more difficulty for the reader.

The results clearly showed that for some L2 readers, not all proper names are known. Nearly a third (30.6%) of participants marked some proper names as unknown, along with other vocabulary items. What was less clear from the findings was the effect of text difficulty. It was predicted that more proper names would be listed to look up in the text which had more unknown vocabulary. In fact, the opposite was found: more proper names were listed for the text in which most of the vocabulary was known. I suggest

that this finding can be explained in this way: when the L2 reader knows most of the vocabulary, she has more attention to give to proper names. Conversely, when more words are unknown, less attention is given to proper names: the reader recognises the unfamiliar item as a proper name but is more concerned with other unknown vocabulary items, words that have content meaning as opposed to referential meaning. In this way, the reader is making judgments as to which words are more important to comprehending the text. So it might be that proper names are considered important for comprehension of the text, but only when all the other content words are known. In terms of look-up behaviour, content vocabulary is given precedence over proper names. When all other words are known, the reader turns her attention to proper names.

Whether L2 readers list proper names to look up in a dictionary, and what types of proper names prompt more look-ups was the focus of research question 2. It was found that participants who marked proper names as unknown also listed these same names to look up. The fact that they listed proper names to look up is important: it suggests that for those participants, the proper names were considered important enough for comprehension to take time away from the text to check the referent/meaning. Dictionary use requires time to look up unknown vocabulary, and thus disrupts the flow of reading; for that reason, L2 readers are likely to look up the meaning of words they perceive as necessary to achieving their reading goal, and not the words they find irrelevant (Hulstijn, 1993). Indeed, L2 reading strategy training often includes practice in how to avoid reaching for the dictionary by using context to infer the meaning of new words. Of course, there are individual differences in approaching unknown vocabulary, and there will be variance in how many words are looked up and how many are ignored (Hulstijn, 1993). But the fact that the participants in this study listed proper names to look up suggests that they considered proper names relevant to comprehension and important enough to look up.

7.3.1 Evaluation of the Unknown Vocabulary Study

The Unknown Vocabulary Study fills an important gap in the literature: There is no empirical backing for the claim that L2 readers are unburdened by proper names. In this regard, the Unknown Vocabulary Study is a successful attempt at determining how L2 readers treat proper names. The results from the study offer contradictory evidence to the claim that proper names are known vocabulary: nearly a third of participants

marked at least one proper name as unknown, and furthermore, they listed them as items to look up in a dictionary. While the sample size was small (N = 49), the results indicate that further research should be conducted with other L2 readers, for example, those of different proficiency levels and different L1s. The results also indicate that researchers should reconsider how best to handle proper names in vocabulary analyses. For example, one possible approach would be to maintain the proper name ranking according to its frequency in the language, rather than re-categorising them as 1K; this might produce a more accurate picture of the potential burden for L2 readers. Proper names that are very frequent in the language might be identified and recognised more easily than proper names that very infrequent. However, this might not hold true for learners in EFL contexts who might be exposed to a different pool of proper names than learners in English as a Second Language (ESL) contexts.

The study design had the advantage of ecological validity: the materials, method and setting were all familiar to the participants. Types of academic texts were used that these participants would very likely encounter in the classroom, in terms of topic and length. The task was also familiar: as L2 students of academic reading, the participants were familiar with marking unfamiliar vocabulary while reading. Similarly, deciding which words to look up in a dictionary is also routine. As for setting, the experiment took place during class time with the classroom teacher; while the students were aware of their participation in a study, they were in a familiar setting, completing a routine task. Thus, the study was strong in terms of ecological validity, and for that reason, one can be fairly certain that the participants' responses to the task were authentic.

One aspect of the study design that was not controlled for was whether participants were familiar with the proper names in the texts. The texts used in the study were expository, describing historical events, and so in this respect, the proper names were authentic and unaltered. While it was predicted that these participants would have little background knowledge related to the topics, it is likely that participants were familiar with at least some of the proper names in the texts. However, checking which proper names were familiar was not done: there was a concern that this would draw unwanted attention to the proper name items, something to be avoided. So, when interpreting the results, it is important to note that the participants' prior knowledge of the proper names was not known.

As for the overall success of the Unknown Vocabulary Study, the research questions were partially answered: these L2 readers do mark English proper names as unknown

vocabulary, and they do list them to look up in a dictionary. Aspects of the research questions that were not fully answered included how the text difficulty and the types of proper names impacted look-up behaviour. First, with respect to text difficulty, it was noted in the discussion of Chapter 4 that if each participant had completed the task for each level of text difficulty, patterns in individual reader differences might have been detected. That is, it might be that some readers are more inclined than others to pay attention to proper names. In that case, text difficulty might not have an effect; rather, individual differences might emerge in the data. Another way to explore the variable of text difficulty is to conduct post-task interviews so that participants can explain why they marked certain proper names as unknown and not others. So, while the study was less successful in explaining the effect of text difficulty on whether L2 readers mark proper names as unknown, there remain other approaches to investigate this effect.

A second limitation concerns part of the second research question: whether the category of proper name has an effect on look-up behaviour. For instance, it might be the case that place names or company names are looked up more often than other types of proper names. Because place names are not often explained in context (Nagy & Anderson, 1984), one might expect that readers will have more difficulty with unfamiliar place names. In this study, however, no place names were listed. And this finding points to a limitation of the study's conclusions: because only three texts were used, it is difficult to draw any generalisations about the proper names associated with only those three texts. The names that were listed to look up in this study were mostly non-prototypical names: family names in plural form with definite article, epithets, and brand names. Further research with a larger corpus would be better positioned to discover patterns in categories of proper names that trigger look-up behaviour.

Tracking L2 readers' look-up behaviour during extensive reading with graded readers, for example, might reveal a discernible pattern in categories of proper names that trigger look-ups. Preliminary analysis from my own unpublished data indicates that some categories of proper names might indeed cause more look-ups than others. The data was collected informally as classroom research throughout an extensive reading program run at a university. The students were required to keep a list of new vocabulary from their graded readers that they would later check and study. The task of maintaining a list was designed to motivate their vocabulary learning through extensive reading. I collected the vocabulary-to-learn lists every week over a 15-week semester, and for exploratory purposes, noted any proper names that appeared on the

lists. From this data, a discernible pattern emerged in the proper names that students listed to look up: the largest category was place names (23 items listed); the second largest was 'Other', which included brand names, religions, and planets (19 items). To a lesser extent, family names (10 items) were listed, and female and male personal names (7 items). While the data was collected informally as classroom research, it illustrates that a pattern in look-up behaviour might be found using a larger corpus. Computer programs that monitor and record look-up behaviour could be useful in conducting such research, in comparison to the pen and paper method described here.

There are other ways to extend the research; for example, participants with different proficiency levels and L1s could also be investigated. As the findings here indicate, proper names are not wholly unproblematic for intermediate level Japanese readers of English. While it remains unclear which types of names might be difficult or how text difficulty affects proper name processing, it is evident from the findings that to assume L2 proper names are unproblematic is misleading.

7.4 Cultural familiarity of proper names

Having established that proper names can be problematic for some L2 readers, the aim of Chapter 5 was to examine the effect of cultural familiarity of proper names on reading comprehension. As was discussed in Chapter 2, proper names can be considered an aspect of cultural knowledge; members within a cultural group will exploit this shared knowledge, for example, by using proper names in place of generic terms (e.g. referring to a detective as a *Sherlock Holmes)* (Hanks, 2013). L2 readers might not be privy to this cultural knowledge. As Hanks (2013) remarks, if you do not know who Shakespeare was or where London is, you are not a fully-fledged member of the English speaking culture (p. 34). Indeed, several studies have demonstrated an effect of cultural knowledge on reading comprehension (e.g. Johnson, 1981; Steffensen et al., 1979). Thus, the two studies in Chapter 5 considered whether cultural familiarity with proper names aids reading comprehension, and conversely, whether cultural unfamiliarity with names hinders comprehension.

The Cultural Study replicated the original study (Erten & Razi, 2009) as closely as possible in terms of materials and methodology. The variable that differed was the participants: in the original study, the participants were advanced Turkish readers of English while the replication study used intermediate level Japanese readers of

English. Cultural referents, most of which were proper names, were manipulated in an American short story in order to make them more culturally familiar to the participant group. Reading comprehension was compared between four treatments: the original version with English proper names; the original version plus reading activities; an adjusted version with L1 culture proper names; and the adjusted version plus reading activities. While a large effect for cultural familiarity on comprehension was seen in the original study, there was no effect found in the replication study. Post-hoc analysis was carried out to determine possible reasons for these disparate results. A vocabulary profile of the reading material indicated that the text might have been too difficult in terms of lexis for the intermediate Japanese readers. When a large portion of vocabulary is unknown (in this case, perhaps up to 10% unknown), no effect for cultural knowledge will be seen. Research has indicated that for background knowledge to have an effect on reading comprehension, at least 95% of the vocabulary needs to be known (Schmitt et al., 2011). Results from the post-hoc analysis indicated that the lexical difficulty of the text was likely a contributing factor as to why no effect found. In this regard, the research question (i.e. does cultural familiarity of proper names aid comprehension) was not successfully answered because the reading material was too challenging for the participant group.

For that reason, a follow-up study was run, this time taking into account the lexical difficulty of the text, in order to investigate the effect of culturally familiar proper names. In the Three Treatments Study, two shorter reading texts were used; vocabulary profiles indicated participants would know 98% of the words, a level at which an effect of cultural background knowledge might be seen. Three experimental treatments were compared: culturally familiar proper names (i.e. L1 names), culturally unfamiliar names (i.e. L2 names), and no proper names, only common nouns (e.g. *the witness, the defendant*). Comprehension was tested with a free recall task, followed by multiple-choice questions. No effect for culturally familiar proper names was seen on comprehension. I suggest this might have been due to the comprehension tasks, which focused on overall comprehension. Previous research into the effect of cultural knowledge has shown that it is difficult to determine what level of processing is affected by this knowledge (Steffensen et al., 1979). Some studies suggested that the effect might be seen mostly on inferential comprehension, though it is difficult to draw definitive conclusions (Alptekin, 2006; Johnson, 1981). Perhaps this is why no effect was seen in the Three Treatment Study: global comprehension was targeted by the summarising task, not inferential. This points to an area for further research: the effect

of culturally familiar and unfamiliar proper names on inferential or detailed reading comprehension.

Cultural familiarity of proper names was treated in these two studies as an aspect of background knowledge, but not within schematic theory. The two studies drew on previous research that had used schema theory to underpin the investigations. As was discussed in Chapter 5, schema theory has largely fallen out of favour in contemporary reading models: there is vagueness and disagreement in defining the term 'schema', and very little is known how such mental representations are drawn on and used to aid reading comprehension (Alba & Hasher, 1983; Sadoski & Paivio, 2007). Also, 'schema' is sometimes used to mean 'background knowledge'; however, these two terms should not be used synonymously (Sadoski & Paivio, 2007). The reason for not conflating the two terms is that research has demonstrated an effect of background knowledge on comprehension in cases when the reader is required to draw on specialist or cultural knowledge (Grabe, 2009). So, while schema theory is referred to less often in contemporary literature, the concept of cultural knowledge was considered valid to investigate the effect of proper names on comprehension.

7.4.1 Evaluation of Cultural and Three Treatments studies

The hypothesis that cultural familiarity with proper names would aid comprehension was well supported by previous research (Alptekin, 2006; Erten & Razi, 2009; Johnson, 1981; Reynolds et al., 1982; Steffensen et al., 1979). However, as noted, it is not clear from the research if cultural knowledge affects higher- or lower-levels of processing. As Grabe (2009) points out, while the effect of background knowledge on reading comprehension has been demonstrated, it remains unclear how the reader draws on this knowledge to aid comprehension. Carrell (1983) also points out that members of the same cultural group might draw on different specialist knowledge when making interpretations of a text. Such specialist knowledge might override cultural knowledge for text interpretation. So, while an effect of cultural knowledge on reading comprehension has been shown in several robust studies, it is difficult to determine where the effect is seen and how this knowledge is used for text comprehension. While several robust studies have shown an effect of cultural knowledge on reading comprehension, these two studies did not.

Even though the overall success of the two studies presented here was middling, the results from the two studies do provide further support regarding the importance of

lexical knowledge to reading comprehension. In the Cultural Study, it was estimated that participants knew about 90% of the vocabulary; their comprehension scores ranged from 45% to 53%. In the Three Treatments Study, for which texts were modified with glosses so that participants knew 98% of vocabulary, their comprehension scores ranged from 60% to 72%. These findings support the notion that a larger percentage of known vocabulary leads to better reading comprehension; other research has also indicated as much (e.g. Carver, 1994; Nation, 2006; Schmitt et al., 2011).

The methodology used in the Cultural Study closely followed that of the original study. As a result, a limitation to the findings was created with regard to the instruments and the participants' proficiency level. While it is true that one of the purposes in conducting replication research is to explore whether the conclusions are generalisable to other contexts, in this case, the reading material proved too difficult for the participants. When more than 5% of vocabulary is unknown, it is unlikely that an effect of background knowledge on comprehension will be seen (Schmitt et al., 2011). The lexical difficulty of the text was evident from the participants' comprehension scores, which ranged from 45% to 53% in the different treatment groups. This level of comprehension is considered minimal (Laufer, 1989), while 70% comprehension is considered sensible to aim for (Schmitt et al., 2011). One possible work-around would have been to gloss all the words in the original text that were at the 3K-band and above. Then, the intermediate level participants might have known at least 95% of vocabulary, thus setting the conditions to test for an effect of cultural background knowledge on comprehension. However, modifying the text with glosses might have made it difficult to compare results with the original study.

The Three Treatments Study, which was conducted to address the lexical issues from the Cultural Study, had some limitations to its success in meeting the aims. First, the design was somewhat complicated in trying to address several possible outcomes. For example, two texts were used instead of one, in consideration of the proper name treatment in the original text versions. Also, using two texts and three treatments resulted in a design that was not perfectly counterbalanced. Another limitation related to methodology was 'time on task': participants were given ample time to read and presumably work out any difficulties related to proper names. This limitation points to possibilities for further research: inclusion of a timed element to determine how familiar or unfamiliar proper names can impact processing speed. There were also some limitations with regards to the testing instruments. As was noted in Chapter 5, the texts

were shortened to the point (250 words) where it was difficult to generate comprehension questions. A related problem was with the multiple-choice task: it was not necessary to understand the proper name referents in order to answer these questions. Multiple choice questions can be problematic in that they can be misleading or even alter existing information (Koda, 2012). For that reason, a free recall task was used in addition to the multiple-choice task. Summary writing tasks focus on global comprehension, and in this respect, the free recall task might not have been sensitive enough to tease apart the effects of proper name familiarity on comprehension. Thus, alternative ways of assessing comprehension, such as using graphic organisers, might be used in future research. Graphic organisers require more cognitive processing than basic tasks; they can also complement other comprehension tasks (X. Jiang, 2012; Schmitt et al., 2011).

In sum, the main limitation to the Cultural Study was the lexical difficulty of the text for the participants; the main limitation to the Three Treatments Study concerned the level of processing being targeted through the comprehension tasks. The limitations to these two studies illustrate the difficulty in determining how cultural background knowledge affects reading comprehension. Other studies that have found an effect of cultural knowledge on comprehension also report that it is unclear whether the effect is on lower-level or higher-level processing. Reading researchers also acknowledge that while background knowledge affects comprehension, it remains unclear how the reader uses the knowledge for text interpretation (Carrell, 1983; Grabe, 2009). In conclusion, while other research has demonstrated an effect for cultural familiarity on comprehension, these two studies did not find an effect for cultural familiar proper names on global comprehension.

7.5 Lower-level processing of L2 proper names

Given the challenge in determining the precise effect of proper names on higher-level processing, the next consideration was proper name processing as an aspect of lower-level reading skills. The aim of the Context Study (Chapter 6) was to investigate the extent to which L2 readers can use context to identify proper names. The reading task was one in which the orthographic clue for proper names (i.e. initial capital letter) was absent, and the semantic information was ambiguous (i.e. target items have both a common word meaning and a proper name usage). Thus, the reading task in this study required participants to draw on contextual clues to identify proper names. Through this

study's substantive aim, the validity of an assumption found in L2 vocabulary research is examined: whether L2 readers can identify proper names in context.

The Context Study was successful in answering the research question. The findings indicated that participants were able to use context to identify proper names in only 28.4% of cases. Compared to identification of target non-name items, the difference was statistically significant and the effect size was large (r = .87). In this respect, the substantive aim was clearly answered: L2 readers are not very successful at using context to identify proper names. This finding contradicts an assumption that L2 readers will easily recognise and understand proper names from context.

The findings from this study support those from other research that looked at L2 readers' ability to use context to infer word meaning. Those studies reported very similar results, in that context aided correct lexical inference in approximately 25% of cases (Bensoussan & Laufer, 1984; Nassaji, 2003b). Also, there were some suggestions in the data that indicated these participants, as L1 logographic readers, might have been relying on word recognition skills over contextual clues to identify names: there were three target non-name items that were misidentified as proper names by a significant percentage of participants. Other research has also found that L1 logographic readers tend to rely on bottom-up processing by nature of how they read in their L1 (Huckin & Bloch, 1993; Parry, 1996). However, it is conjecture that the participants in this study were using word recognition skills to identify proper names because the participants' inferencing strategies remain unknown. No posttask interviews were conducted, for instance, that might have shed light on their inferencing strategies, so it remains speculative whether they were relying on word recognition skills.

7.5.1 Evaluation of the Context Study

Investigating proper name processing as an aspect of lower-level reading skills has sound theoretical grounding. While L2 research and pedagogy has traditionally focused on development of readers' higher-level processing skills (Bernhardt, 2005), it is likely that L2 readers' difficulties are not related to higher-level processing (Nassaji, 2002, 2014). Rather, processing and comprehension difficulties might be attributed to a lack of L2 lower-level processing experience (Segalowitz, Poulsen, & Komoda, 1991). This may be especially true for L2 readers who are reading in a writing system that differs from their L1 (Alderson, 2000), as is the case in this research context (i.e. Japanese L2

readers of English). Recall that in the Unknown Vocabulary Study (Chapter 4), it was ascertained that some L2 readers treat proper names as vocabulary to look up, suggesting a difficulty with lexical access. Also, the Cultural Study and Three Treatments Study (Chapter 5) found no effect of cultural familiarity of proper names on higher-level processing; that is, there was no effect on overall comprehension due to familiarity or unfamiliarity with the proper names. Thus, it seems reasonable to investigate whether any difficulties with proper names occur with lower-level processing.

As has been noted, contemporary reading models recognise the importance of both levels of processing to reading comprehension (Birch, 2007; Grabe, 2009; Nassaji, 2014). It is thought that processing occurs interactively within levels. Restricted interactive models predict that for successful higher-level processing to occur, lower-level processing must be automatic and efficient. For example, in Perfetti's (1985) Verbal Efficiency Model, word recognition skills are key; word identification occurs through interactive orthographic, phonological, semantic and syntactic processing (i.e. sub-skills of word recognition). The model predicts that when lower-level processing is efficient, cognitive processes are freed up for comprehension. The model attributes any difficulties with higher-level processing to inefficient word identification skills. While Perfetti's model is not an L2 reading model, the implications for L2 reading seem relevant: L2 readers have much less experience with word recognition processing than L1 readers, so it seems likely that any difficulties they have would occur at the lower level (Segalowitz et al., 1991). Because L2 readers might not process proper names efficiently and automatically due a lack of experience, this inefficiency in proper name recognition might interrupt comprehension processes, such as using context and inferencing skills. Therefore, with respect to the theoretical orientation, the Context Study's focus on proper name processing as an aspect of lower-level skills seems well grounded.

The research design was successful in the creation of a reading task in which participants needed to rely on contextual clues to identify proper names: the orthographic clue was removed (i.e. initial capital letter), and the semantic information was ambiguous (e.g. *rose/Rose*). Presumably, the participants had phonological representations of these high frequency items. However, the disruptions to orthographic and semantic processing seemed sufficient to motivate contextual inferencing. In this regard, the methodology seemed successful in meeting its aim. Another strength of the design was that authentic sentences were selected from a

corpus; this was done in order to limit researcher bias that might have resulted from generating sentences for the purpose of the experiment.

The study was less successful in identifying which factors affected participants' ability to identify some proper names and not others. None of the predictors in the multiple regression were significant, though richness of context made the largest contribution to the model. First, the selected sentences were not initially piloted to test for richness of context. As part of the post hoc analysis, it was found that L1 users rated most of the sentences as rich in context to indicate that the target item was a proper name; however, not all were rich in context. Three sentences had moderate context and one sentence had no context to indicate a proper name. Thus, the effect of context could have been better examined in the multiple regression if richness of context had been controlled for prior to running the experiment. For example, it might have been informative to use a mix of sentences rich in context and others moderate in context.

Another limitation related to the research design concerns the frequency of the selected target items. In the initial planning, it was desired that all target items were high frequency to ensure that the intermediate level participants were familiar with them. However, item frequency might have been a predictor in participants' ability to identify them. Connectionist theory of learning predicts that lexical knowledge is based on prior experience and statistical frequencies (Ellis, 2002; Grabe, 2009). Therefore, it seems likely that frequency would have had an effect on how well participants were able to identify proper names as such. Unfortunately, this effect was not identifiable in the multiple regression because most target items were high frequency. An alternative approach would be to have a mix of high and low frequency items, in order to test for an effect of frequency.

A final limitation to be noted also relates to the selected target items. An alternative analysis could consider the part of speech of the corresponding non-name items, which was not controlled for in the initial experiment planning. The part of speech of the non-name target items might affect participants' ability to correctly identify the proper names. As Hoey (2007) argues in his theory of lexical priming, we are primed by every encounter we have with a word, keeping a subconscious record of the context and co-text in which the word was seen or heard. He goes on to assert that while there are many types of priming, grammatical category is the most fundamental priming of words: "The point here is that words are not nouns, verbs or whatever, they are typically *primed for use* [emphasis added] as nouns, verbs etc." (Hoey, 2007, p. 35).

Therefore, if a target non-name item occurs most frequently as a noun, then the reader is primed to identify the item when it is used as a noun. If an item occurs most frequently as a verb or adjective, then the reader is not primed to identify it as a noun (or in a noun slot). Thus, an alternative approach would be to select only nouns for non-name items, or contrast only two categories, such as nouns and verbs, to control for the part of speech variable.

In summary, the Context Study was successful in determining the extent to which L2 readers can use context to identify proper names. The study's design aimed to disrupt normal word recognition by removing the orthographic and semantic information. In doing so, the participants were required to rely on contextual clues to identify target proper names. The results indicated that these intermediate L2 readers were not very successful in using context to identify proper names: they were successful in less than a third of cases. These findings are supported by other research into L2 readers' use of context for lexical inferencing (Bensoussan & Laufer, 1984; Nassaji, 2003b). Importantly, the findings from this study are at variance with an assumption in L2 vocabulary research that L2 readers can identify proper names from context.

7.6 Generalising the studies' findings

In this next section, the degree to which the findings from these five studies can be generalised to answer the central research aim of whether proper names present a strain to L2 readers is considered. The findings from the five studies are briefly summarised: first, participants self-reported having difficulties with identifying referents and pronouncing unfamiliar proper names, and reported using various strategies to deal with unknown names; second, it was found that proper names can disrupt reading in that some L2 readers treat them as vocabulary to look up in a reference source; third, cultural familiarity of proper names does not seem to affect global comprehension when most of the other vocabulary in a text is known; and lastly, these participants were not very successful at using sentential context to identify proper names.

Two tentative conclusions can be drawn from these findings regarding L2 proper name processing. First, difficulties that L2 readers have with proper names might occur with lower-level processing, in particular, the sub-skills of semantic, phonological, syntactic and orthographic processing. These difficulties were indicated in the findings from the Interview Study (phonological and semantic processing), the Unknown Vocabulary

Study (semantic processing), and the Context Study (syntactic and semantic processing) (Chapters 3, 4 and 6 respectively). Second, L2 proper names might not directly impact higher-level comprehension processes, as seen in the Cultural and Three Treatment studies (Chapter 5). These conclusions are supported by restricted interactive models of reading, which predict that inefficiencies in reading occur with lower-level processing skills (Grabe, 2009; Perfetti, 1985, 2007). It is important to note that these conclusions are drawn from this specific research context, which is subject to certain considerations. Specifically, how the variables of participants and reading materials used in these studies impact the generalisability of the findings is discussed below.

7.6.1 Participants

Several factors related to the participants in this research context, including their proficiency and L1, should be considered as having potential influence on the findings. To recap, the participants in all of these studies (N = 288) were intermediate level Japanese readers of English, studying in their first year of university (average age, 18 to 19 years old). Their majors varied, but they all had similar language learning backgrounds in that they had studied English for at least six years upon entering university. Koda (1996), in her proposed framework for L2 reading research, suggests that the following factors related to the participants be taken into account: the amount of L2 reading experience; the orthographic distance between the L1 and L2; and transfer of L1 processing to L2 reading (p.453). Drawing on this framework, the following factors are considered as having possible influence on the findings: the participants' L2 orthographic processing experience; their non-alphabetic L1 (i.e. Japanese); and how their L1 processing (i.e. of logography and syllabary) might have impacted their L2 processing (i.e. of alphabetry).

First, the participants' L2 orthographic processing experience may have influenced the results in these studies. As has been noted, lower-level reading skills, including orthographic processing, are thought to develop through processing experience (Koda, 1996). Inefficient lower-level skills can disrupt higher-level comprehension processes, such as using context to make inferences (Perfetti, 1985, 2007). The participants in this research context probably did not have a great amount of L2 reading experience as intermediate level users of English. There is some research which suggests that prior to entering university, the average Japanese student will be exposed to only 100,000

English words over six years of study in junior high and high school (Waring, 2014). As a result of this limited exposure to English text, lower-level processing skills would certainly be underdeveloped.

In any case, L2 proficiency should not be conflated with lower-level processing skills (Nassaji, 2014). Recall that even advanced, proficient L2 users read slower in their L2, most likely due to inefficiencies in lower-level processing (Akamatsu, 2003; Nassaji, 2014; Segalowitz et al., 1991). Because processing experience is important to developing efficient and automatic lower-level skills, it seems these participants' limited processing experience may have impacted their performance in the Context Study, for instance. This possibility points to an area for further research: to investigate how readers with more processing experience are affected by and process proper names in continuous text. Research approaches such as eye tracking might be used to investigate this further.

Another important factor with respect to how the participants may have impacted the findings is their non-alphabetic L1. The participants in this specific research context are unique because as L1 Japanese users, they are both logographic and syllabic readers. It is also important that as L2 readers of English, they are processing a distant orthography from their L1 (Alderson, 2000; Koda, 1996). As was noted in Chapter 2, L2 reading is a dual-language process (Koda, 2005). Transfer effects from the L1 are expected, and these can be positive or negative, depending on the distance between the L1 and L2. For example, cross-orthographic research has investigated how similarities and differences between L1 and L2 orthography aid or hinder L2 reading. Nassaji (2014) cites several studies (e.g. M. Wang & Koda, 2007; M. Wang et al., 2003), which have demonstrated that differences between L1 and L2 orthography can have a negative effect on L2 word recognition: L2 readers use L1 orthographic processing strategies for L2 texts, and when the orthographies differ greatly, this can negatively impact L2 word recognition.

Ehrich et al. (2013) cite evidence that suggests different writing systems (i.e. logographic, syllabic and alphabetic) are not processed in the same way. In particular, they present evidence from behavioural and neuroscience research that supports the hypothesis that logography is processed differently than alphabetry. The arguments that Ehrich et al. (2013) present have direct implications for the findings from these studies because the participants were L1 Japanese (i.e. logographic and syllabic) readers of English (i.e. alphabetry). A key point Ehrich et al. (2013) make is that visual

processing is more important for reading logography while phonological processing is more important for alphabetry. Because there is L1 processing transfer to L2 reading, this can be problematic when the L1 and L2 writing systems emphasise different access routes (i.e. phonological vs. visual).

To back their assertion, Ehrich et al. (2013) first present the main differences between writing scripts. Logography is more visually complex, and while there are sound and meaning associations in characters, the radicals (components of each character) are often not reliable clues to pronunciation. That is, in processing of Chinese characters, it is not possible to access phonology prior to word identification (Ehrich et al., p. 42). In contrast, alphabetry is less visually complex and has a phonemic association but graphemes (letters) alone have no meaning association. How alphabetic readers identify words has been much debated, whether it is pre-lexical, that is, relying on phonology to access meaning, or post-lexical, that is, retrieving meaning before phonology, if phonology is accessed at all. Syllabary (e.g. Japanese Kana) is similar to alphabetry in that the graphemes (syllables) are relatively simple in structure, have sound associations but no meaning associations. The main difference between a syllabic system and an alphabetic one is graphemes: in a syllabic system, the basic graphic units correspond to syllables while in alphabetry, graphemes correspond to phonemes (Perfetti & Liu, 2005).

Providing evidence that logographic and alphabetic scripts are processed differently, Ehrich et al. (2013) draw on research from a range of disciplines. They cite experimental psychology studies that suggest reading logography is a cognitive process more closely related to picture processing than reading alphabetic script (e.g. Chen & Tsoi, 1990). The authors point to L1 reading studies which demonstrate the importance of phonological awareness and phonological processing skills for English word identification; in contrast, visual skills are important for logographic reading (Hanley & Huang, 1997). They also refer to L2 reading studies which indicate L1 logographic background readers depend more on visual processing than phonological processing (e.g. Akamatsu, 2003). Other studies have shown how L1 logographic readers are less sensitive to intra-word analysis (analysing individual/constituent letters) than L1 alphabetic or syllabic readers (Koda, 1999). Cognitive psychology studies have shown L1 logographic participants outperform L1 alphabetic participants on visuo-spatial working memory tasks (e.g. Demetriou, Kui, Spandoudis, Kyriakides, & Platsidou, 2005). Evidence from neuroscience includes studies involving fMRI scans, such as Tan et al. (2001), which showed activity in mid-dorsal and prefrontal regions of

the brain for logographic readers; these areas are associated with spatial and verbal memory and are not usually active when reading alphabetry. Another significant finding involving fMRI scans (Nakada, Fujii, & Kwee, 2001) is that alphabetic and logographic scripts are processed in distinct areas of the brain, and that readers use the same areas of the brain when reading in their L1 and an orthographically different L2. In their conclusion, Ehrich et al. (2013) note how reading models like Goodman's (1988) emphasise top-down processing (i.e. higher-level processing); their criticism of such models is that logographic readers may rely more on bottom-up processing (i.e. lower-level processing) by virtue of how they read in their L1. A universal reading model would need to take into account such cross-linguistic variation in reliance on bottom-up and top-down processing (Ehrich et al., 2013, p. 48).

The evidence that Ehrich et al. (2013) present for processing variation in different writing systems has implications for the assumptions made about L2 proper name processing. For example, it is often assumed that L2 readers can easily identify proper names in texts from the initial capital letter (see Chapter 4, section 4.2). For L1 logographic readers who have highly developed visual skills, one might predict that such readers are skilled at noticing the initial capital letter on proper names. However, recall that alphabetry is visually much simpler than logography. Thus, it might be that for letters which look similar in upper and lower case (e.g. Cc, Kk, Oo), the differences in case size might not be as obvious to L1 logographic reader. That is, alphabetic letters might be deceptively simple to the logographic reader who is more familiar with processing complex characters. Conversely, such readers might easily notice upper case letters that look very different from the lower case (e.g. Qq, Aa, Bb). This prediction would need empirical backing, for example, by looking at reaction times of processing similar looking case letters and different looking case letters, using a grapheme decision task (e.g. similar to a lexical decision task: *are these letters the same grapheme?*).

Another example of how processing differences in the L1 and L2 writing systems can impact proper name processing concerns the possible transfer of L1 reading strategies to L2 reading. Brain imaging studies with Japanese-English bilinguals have shown that they use the same area of the brain, the left occipito-temporal region, to read their L1 Kana syllabary and L2 Roman alphabet Nakada et al. (2001). In addition, Nakamura, Dehaene, Jobert, Le Bihan, and Kouider (2005) have shown with neural imaging that when Japanese read Kanji characters (Chinese characters), there is more activation on the midline temporal regions of both the left and right hemispheres. The differences in

area of activation are the result of different emphasis on phonological versus lexico-semantic routes from the different scripts (Nakamura et al., 2005). The implications of these findings for proper name processing are unclear. In Japanese, proper names are most often written in Kanji script. And recall that in logographic writing systems, the character itself will often not provide any reliable clue to pronunciation. Indeed, in Japanese, it is a common problem with proper names that if the reader does not know the Kanji character involved, the name will remain meaningless (Cook & Bassetti, 2005). Because it is impossible for the reader to guess the pronunciation, the reader must either research the pronunciation or continue reading without a phonological representation. If the Japanese reader extends this approach to reading L2 English names, that is, not trying to create a phonological representation of the name, then this might have a negative impact on overall comprehension. For as Koda (2004) notes, it is important to have a phonological representations of a new word: without one, it is harder to keep the new item in working memory or move it to long-term memory. Overall comprehension of the text may then be impeded because of competition for working memory capacity. One might predict that in a text with the proper names of several different people or places, this would create a burden for the working memory.

Thus, in considering how proper names affect L2 reading comprehension, it is important to consider the reader's L1 for possible transfer effects. As Ehrich et al. (2013) note, current reading models fail to take into account variation in processing of different writing scripts. The processing variation of different writing systems should also be considered when making assumptions about whether L2 readers understand certain lexical items, such as proper names. Certainly, the fact that the participants in these studies were L1 Japanese readers of L2 English is an important factor for the findings. Because participants' L1 will impact how L2 proper names are processed, this offers directions for further research, for example, a comparison of how L1 logographic and L1 alphabetic participants process orthographically distant L2 proper names.

7.6.2 Reading materials

Another consideration for how applicable the findings are to determining the strain of proper names for L2 readers concerns the length and types of reading texts used in these studies. First, text length might impact applicability in so far as the text resembles an authentic experience of proper name processing. For example, longer texts with many proper names might place a greater burden on working memory. On the other

hand, shorter texts can create a different processing strain in that a proper name might occur only once in the text. In these studies, texts of different lengths were used to different purposes. The text lengths ranged from sentence-long utterances to a short story of nearly 3,000 words. There were also several shorter texts, ranging from 250 to 550 words. The short story from the Cultural Study (Chapter 5), as the longest text, did have many proper names. With the shorter texts used in the Three Treatment study (Chapter 5), one problem that arose was the difficulty in generating comprehension questions for such short texts. Another issue with shorter texts of around 300 words is that the readings may have been too short to have any salient text structures and thus were perhaps inauthentic; a reader can use such discourse structures to aid comprehension (X. Jiang, 2012). In comparison, a text of 700-800 words might demonstrate the writer's organisational framework in a more conspicuous way. As for the sentence-long utterances used in the Context Study (Chapter 6), these 'texts' might raise the criticism that they do not represent proper name processing in an authentic way: a reader would normally have several encounters with a proper name in a longer text, and presumably, she would build meanings for that proper name from each meeting. In these various ways, the length of text might have impacted the applicability of the findings to a wider context.

The types of texts used may have also influenced the findings. In these experiments, several of the readings used were academic style expository texts (Unknown Vocabulary and Three Treatments studies, Chapters 4 and 5). It was felt beneficial for purposes of ecological validity to use text styles that the participants were familiar with from their academic reading classes. However, the purpose in reading academic texts is to learn new information. In this regard, academic texts might not have been ideal to investigate proper name processing in that the primary reading goal is to learn and retain information. If a proper name is not critical to that new information, then the name might be overlooked or ignored. In that regard, literary fiction might be a more suitable genre to investigate proper name processing: this type of text demands the reader's attention for names of people and places to follow the plot. If the reader does not pay attention to the people and places in a fictional story, then arguably, little will be understood. For example, the excerpt taken from a graded reader for the read-aloud task (Interview Study, Chapter 3) had several proper names requiring the reader's processing attention. The short story used in the Cultural Study (Chapter 5) was also suitable in this regard: it contained many proper names that the reader had to pay attention to, to understand the story. As noted, however, the original version of that

story was authentic, and as such, was too difficult for the participants in terms of lexical load.

Certainly, when using authentic texts, lexical difficulty is an important consideration. At least 98% of the vocabulary needs to be understood to allow for 70% comprehension (Schmitt et al., 2011). That would make most authentic texts unsuitable for intermediate level L2 readers. However, some researchers (e.g. Alptekin, 2006) argue against modifying or creating texts for experimental purposes: the modification renders the text artificial and perhaps even difficult to read. Such an effect on general readability may have been seen in Johnson's (1981) study, where the L1 readers had lower comprehension scores for the adjusted text than the original version. A possible workaround when using authentic texts with participants of lower proficiency levels is to allow the use of dictionaries, as was done in the Context Study (Chapter 6). Also, glossing vocabulary above a certain band level would help beginner or intermediate readers, as was done in the Three Treatments Study (Chapter 5).

A type of text that might be ideal for investigating proper name processing is graded readers. One advantage of graded readers is that vocabulary level of the material can be easily matched to the participants' vocabulary knowledge. Also, graded readers are extended pieces of texts, and as such, their length would be useful to investigate the load of proper names on working memory. Indeed, one concern with graded readers is that often, too many characters are introduced in the first few pages, especially for simplified versions of original novels (Hill, 2013). Some might argue that graded readers are not authentic material in that they are modified or simplified. One can refute the claim, however: graded materials offer the L2 reader an authentic reading experience in that most of the vocabulary is known, similar to L1 reading (Nation, 2018). Newspaper articles represent another text genre that might be used to investigate proper name processing. Newspaper articles contain a high percentage of proper names, up to 6% of text tokens (Nation, 2006). However, as noted above, the issue of authenticity and lexical difficulty would need to be addressed for low proficiency participants.

In sum, the reading materials used in these studies may have affected the findings in terms of text length and genre. Certainly there remain many possibilities to explore how L2 readers are affected by proper names by using other text types. As noted above, graded readers might be ideal in that they are longer pieces of extended writing, and the vocabulary has been graded to match the reader's proficiency level.

7.7 Implications of the research findings

This final section will present implications of the findings from this research for the areas of: L2 vocabulary research; pedagogy; and materials development, including textbooks, graded readers, and testing. These implications are discussed in light of the findings from the studies presented in Chapters 3 to 6, as well as other literature.

7.7.1 Vocabulary research implications

The findings from this research have direct implications for L2 vocabulary research, specifically, for lexical analyses of texts in which proper names are treated as known vocabulary. In particular, the findings from the Unknown Vocabulary Study (Chapter 4) indicated that some L2 readers do look up proper names in a reference source while reading. This suggests proper names are not known vocabulary and are considered important enough for comprehension to warrant dictionary use. Also, participants self-reported difficulties in identifying proper name referents in the Interview Study (Chapter 3). Moreover, it was shown in the Context Study (Chapter 6) that L2 readers are not very successful at using context to identify proper names, even when they use dictionaries to help with other lexical items.

Given these findings, it seems misleading to assume that context is sufficient for an L2 reader to understand proper names and their referents. Therefore, if the purpose of a lexical analysis is to assess the potential difficulty of a text for an L2 reader, it might be more accurate to categorise proper names according to their frequency in the language, as is done with other lexical items. Just as with high frequency vocabulary, L2 readers might be more familiar with proper names that are common in the target language. However, this might not be the case in EFL contexts. For example, the most common names in Britain might be relatively unfamiliar to EFL learners in Japan. More research would be needed in this area, for example, by asking L2 readers to self-report on proper names they are familiar with from the target language. In short, it seems that treating proper names as known vocabulary or to remove them altogether from lexical analyses does not capture the potential burden proper names can present for some L2 readers.

7.7.2 Pedagogical implications

Given that L2 readers do treat proper names as vocabulary to look up (Unknown Vocabulary Study, Chapter 4), and that they are not very successful at identifying proper names in context (Context Study, Chapter 6), these findings have implications for the L2 classroom. It is important that adequate attention be given to proper names in classroom reading texts. Proper names do contribute to the meaning of a sentence or text (Allerton, 1987). And as the participants in these studies have demonstrated, it is not likely that L2 readers will understand every proper name they come across. Learners can support themselves to an extent. For example, learner dictionaries usually have entries for proper names that are common in the target language. For proper names that are not as common, Internet searches will usually provide that information, assuming online searches are possible in the classroom. However, the classroom teacher also can take several proactive approaches to proper names to support learners.

Teachers should check what students understand about the proper names in any given text, before, during or after reading. This comprehension check should extend to any titles that appear with proper names, as students might not understand the significance of such titles. To take an example from the text about smallpox used in the Unknown Vocabulary Study (Chapter 4), several participants thought that the title *Lady* (from *Lady Mary Wortley Montagu)* meant a young, beautiful girl; they were not aware of the meaning of *Lady* as an aristocratic title. In literary texts, there are other aspects of proper names that learners need support with, for example, nicknames and the use of different names to refer to the same person. Participants remarked on these specific difficulties in the Interview Study (Chapter 3).

To briefly illustrate how nicknames and different name-calling of characters can create processing challenges for L2 readers, I will present a few examples from a short story I have used in the L2 classroom. In Ernest Hemingway's "Indian Camp", the main character is Nick, and his father is a doctor. Nick calls his father *Dad* while other characters in the story refer to him as *Doctor,* which leads some students to wrongly infer that in addition to Nick's father, there is another character, a doctor. Later in the story, after a suicide has been discovered, Nick's father calls him *Nickie*; the significance of the diminutive form *Nickie* is one that L2 students are not usually aware of. In the same way, after the suicide, Nick refers to his father as *Daddy*, suggesting a reversion to childhood. While an L1 college reader might easily pick up on the

significance of such name-calling, it is mistaken to assume that an L2 reader will also understand the nuances behind these alternate name forms. Of particular importance to proper names in “Indian Camp” is that while there are several Indian characters, none of them are given names, causing referential strain for the reader. Hemingway, of course, would exploit this device of not naming certain characters to great effect in later stories.

There was one proper name in “Indian Camp” that I failed to note the first time I taught the story. A student drew my attention to it after class. The line is spoken by the doctor:

(1) *The nurse should be here from St. Ignace by noon and she’ll bring everything we need.*

Note that there is little context to help the reader determine the referent. Also, the pronunciation of *saint* is not apparent from how the name is presented. In answer to the student’s enquiry, I said it referred to the name of a hospital; I explained that hospitals are sometimes founded and run by the Catholic Church, and that the name refers to a saint. Later, I looked it up and learned that *St. Ignace* is in fact a town in Michigan, where the story is set. This is a nice example of how place names are not explained in context. What was especially interesting to me was that the student had noticed the name, even though the line was not of great importance to understanding the plot. She could not find the name in her dictionary and took time to see me after class to ask what it referred to. In short, there are many facets of proper names that are not obvious to the L2 reader, and sometimes L1 readers can make errors with proper names. If classroom teachers are attuned to the difficulties that proper names can cause, they are in an ideal position to support learners.

Depending on the type of reading text and the reading goal, it might be helpful for learners if proper names are pre-taught. For example, with L2 listening, Kobeleva (2012) found that if proper names were pre-taught, learners did better on detailed comprehension questions; however, there was no difference on global comprehension between the group for whom names were pre-taught and the group for whom names were unfamiliar. Kobeleva (2012) also found that for the listeners in the names unknown treatment, they were able to derive less than 50% of information about the proper name referents from context, after two listening attempts (p. 94). Furthermore, the listeners who were pre-taught proper names perceived the comprehension tasks as easier to do, and self-reported higher comprehension than the unfamiliar name

treatment group. These findings suggest that proper names do contribute to the meaning of a given text, and if these items remain unknown, learners might feel frustration with the task. Thus, it seems reasonable that L2 readers would also benefit from pre-teaching of names, in particular if the reading goal focuses on detailed comprehension. Pre-teaching of proper names might also result in less anxiety related to text comprehension.

Another proactive approach teachers can take in the classroom is to ensure that students have phonological representations of proper names. As was seen in the Interview Study (Chapter 3), several participants mentioned frustration at not knowing how to pronounce unknown names. Not being able to pronounce names led participants to focus on the first letter of names, which is not always a useful strategy, especially when two names start with the same letter. As Brennen (1993) suggests in his plausible phonology hypothesis, there is a wide range of phonology that is plausible and acceptable for proper names. His theory suggests that the learning of new phonologies is done more often for proper names than for other types of words. This is where teachers can support learners, by training them how to apply phonological rules of English (or other target language) to unknown names. Having a phonological representation is important for moving the name to long-term memory. As Hulstijn (2001) notes, when new lexical information is processed more elaborately (e.g. both phonological and orthographic information), this will lead to higher retention than if only one aspect is given attention (p. 270). Thus, the more aspects of proper names that teachers can draw to their learners' attention will help ensure that this knowledge can be applied in future encounters.

Teachers can also help their students with proper names processing by developing lower-level skills using word recognition practices. These are activities specifically designed to improve learners' word recognition skills, and thus enhance overall reading efficiency and automaticity (Grabe & Stoller, 2002). An example of a word recognition activity is to present target words on the left side of the page; next to the target words are four or five distractors of real words that have very similar spelling. Here is an example from Crawford (2005, p. 37):

(2)	**fluent**	fluid	flaunt	flute	flutter	fluent
	reading	reeling	raising	rising	reading	reaping

The reader works as quickly as possibly to identify the item that matches the target; usually 20 or 25 such items are presented on one page as a timed task. Including proper names on such word recognition activities would give learners processing experience to identify proper names with initial capital letters.

Such word recognition activities could also be used for letter case processing. For instance, either the target word or the distractors can be presented in different case conditions (i.e. all capital letters or all lower case) to give learners practice in processing case conditions. This practice is especially important for learners who are processing a different L2 orthography. An example of case processing difficulties is evident in a strategy I observed with an intermediate level Japanese student working on a crossword puzzle (another type of word recognition activity). Before starting the puzzle, the student converted all the target words, which had been listed in upper case letters, to lower case letters. She explained that she found it difficult to read words in all upper case letters, an anecdotal example that confirms Alderson's (2000) assertion that words in all upper case are harder to process than lower case or a mix of cases (p. 75). The difficulty in processing words written in all upper case lies in the fact that upper case letters are all the same height: if the letters are masked and presented in profile only, words would be indistinguishable. In contrast, lower case letters are more distinctive with ascending and descending strokes; if masked and presented in profile, the shape alone can lead to successful word recognition (Perea & Rosa, 2002, p. 786). Thus, for learners who are processing an L2 with a different orthography, it is important that they regularly practice word and letter recognition.

Lastly, teachers can support their learners by teaching patterns found in English names. For example, English family names will more likely end with an *-s* than a personal name (e.g. compare the personal names *Edward* and *Adam* with the family names *Edwards* and *Adams*). Some family names have prefixes (e.g. *O'* as in *O'Connell*) and suffixes (e.g. *-son* as in *Alderson*). Similarly, there are a few common suffixes for places names (e.g. *-ville*, *-shire).* Such strategies for analysing proper names are necessary so that learners can apply them out of the classroom. In summary, teachers can support learning by drawing students' attention to these various aspects of L2 proper name usage. What might seem obvious to L1 users is not necessarily obvious to L2 users whose writing system has different rules for proper names. Therefore, teachers can greatly benefit their learners by taking classroom time to provide regular practice in proper name recognition.

7.7.3 Implications for L2 material developers

In this last section, implications of the research are discussed concerning the development of L2 materials, such as textbooks, graded readers for extensive reading, and tests. These implications draw on the findings from the studies presented in this thesis, as well as other literature. Beginning with textbooks, these materials can be influential on teachers, often informing their pedagogical practice. Most contemporary L2 reading textbooks incorporate reading strategy training for learners. Material developers might incorporate proper name awareness into such reading strategy building. For instance, pre-reading activities can be used to draw learners' attention to proper names and titles through activities, such as scanning for specific information. Tasks can be set for learners to do online searches of place names, and to say how that knowledge enhances text comprehension. As a while-reading activity, learners can be tasked with circling proper names as they read, to build orthographic processing skills. Alternatively, proper names can be dealt with as post-reading activities. For example, readers can be tasked with making a list of the people, places and/or companies mentioned in the article, and to say what they know about each. Learners would need repeated practice with such proper names processing strategies, so that strategy use becomes automatic.

Secondly, some publishers of graded readers do provide support for readers in terms of characters' names. As seen in the Interview Study (Chapter 3), two participants mentioned using the names charts or family trees that publishers provide at the beginning of books. Showing the relationships between different characters can be helpful, especially when a character might be referred to by more than one name. As Hill (2013) remarks, introducing several characters in the first few pages of a story creates a processing burden for the reader; she must try to keep different characters in mind. This burden could be addressed by limiting the number of characters introduced in a short space. Also, not having a working pronunciation of unfamiliar names can cause a strain on working memory. Offering pronunciation guides will be useful to learners who are encountering these proper names for the first time.

Lastly, test writers need to consider the reading burden that proper names might place on test takers. Kobeleva (2012) suggests that for listening tasks in testing situations, proper names should be easily inferable from the context; if proper names are not easily inferred, she recommends not using any names at all due to the anxiety that unfamiliar names can cause for test takers (p. 96). For reading assessment, test

writers should also strive to support proper names with rich contextual clues. No test-taker can prepare by learning all possible proper names; therefore, test-takers should not be unduly tasked by the presence of proper names in the text. Proper names should be used to aid comprehension, not hinder it.

Chapter summary

In this discussion chapter, the five empirical studies presented in Chapters 3 to 6 were summarised to determine to what degree the research aims were successfully answered. Theoretical orientation, research design and overall success were evaluated for each study, along with a summary of the findings. Next, the degree to which the findings in these studies could be generalised to other contexts was considered. It was noted that because the participants were L1 Japanese and processing a distant orthography, this had a significant impact on the findings. Then, implications of the research findings for L2 vocabulary research, classroom practice, and materials development were considered. In the concluding chapter that follows, I briefly summarise what I set out to achieve in this investigation, and how the aim was achieved. Several suggestions are made for future research into L2 proper names.

Chapter 8: Conclusion

8.1 Introduction

The aim of this investigation and how that aim was achieved is briefly summarised in this concluding chapter. Recall from the introductory chapter the anecdotal evidence presented that L2 readers are not always able to recognise and understand proper names from the form and context. For example, students had problems inferring referents of proper names like *Craig* and *Jack*, even though these names were presented in L2 reading materials, rich in context at the sentential level, with pictures to support the reader. This anecdotal evidence motivated the research in that it is at variance with a prevalent assumption in L2 vocabulary research that proper names are unproblematic for L2 readers and can be treated as known vocabulary. Underlying this assumption seems to be a treatment of proper names as encyclopaedic knowledge, and in this regard, accords with the widely held philosophical view that proper names have reference but no sense (Lyons, 1977). However, as was discussed in Chapter 2, there are strong arguments from linguistic perspectives that proper names can represent lexical knowledge. Proper names can have categorical meaning, albeit minimal (J. M. Anderson, 2007; Van Langendonck, 2007). For this reason, L2 readers might analyse proper names as they do other L2 lexis, and so "in the learning of a language the distinction between names and common nouns may not always be clear-cut" (Lyons, 1977, p. 219). This investigation has sought to determine how L2 readers analyse proper names and in doing so, assess the soundness of assuming that L2 readers can recognise and understand proper names. After summarising how the aim of the investigation was achieved in section 8.2, several directions for future research into L2 proper name processing are presented.

8.2 Summary of the research

The central aim of this thesis was to determine whether proper names present a burden for L2 readers. In order to address that aim, a series of experiments were conducted to examine the potential burden of L2 proper names from different perspectives. First, proper names were considered from the vantage point of L2 readers; that is, how they view and approach proper names while reading. Then, the effect of proper names on higher-level comprehension processes was investigated.

Lastly, proper name processing was considered from the aspect of lower-level reading skills.

To learn about L2 readers' perspectives on proper names, interviews were conducted with a small sample to gauge L2 readers' feelings towards, difficulties with, and strategies for unfamiliar proper names (the Interview Study). The findings from this exploratory study suggested that some L2 readers do find proper names problematic; for instance, participants self-reported difficulties in identifying proper name referents, distinguishing between family and personal names, and pronouncing unfamiliar names. Some of the questions raised from the Interview Study were used to move the investigation forward. In particular, the following points for further investigation were identified from the Interview Study. First, are the difficulties that the four participants mentioned having with proper names representative of a larger population? Second, because the participants' difficulties seem to stem from a lack of familiarity with proper names, is there an effect of proper name familiarity on reading comprehension? Third, given that the participants were not always able to infer proper name referents from context, how useful is context for L2 readers to identify proper names? These three questions were used to direct the studies that followed.

A larger sample of L2 readers was used in the Unknown Vocabulary Study (Chapter 4). To determine how this larger sample approached proper names, participants were asked to identify unknown vocabulary in reading texts, and list any words they would look up in a dictionary. It was found that nearly a third of participants treated proper names as unknown vocabulary to look up in a dictionary. This finding is important because L2 readers tend to look up words they consider important to comprehension, and ignore words that are considered irrelevant to the reading goals (Hulstijn, 1993). This finding provides counter-evidence for the assumption that L2 proper names can be treated as known vocabulary.

Because the findings from these first two studies indicated that a lack of familiarity with proper names could cause strain for L2 readers, the next consideration was whether there was an effect of proper name familiarity on reading comprehension. Proper names were treated as an aspect of cultural knowledge in the Cultural Study and the Three Treatments Study (Chapter 5). In this way, the focus shifted from the L2 reader's perspective to the effect of proper names on higher-level processing. While some robust L2 reading studies have found an effect of cultural background knowledge on comprehension (Johnson, 1981; Steffensen et al., 1979), it is not clear at what level of

processing the effect was seen: at the lexical level (i.e. lower-level) or with inferencing (i.e. higher-level). Neither study from Chapter 5 found an effect for culturally familiar proper names. This may have been because global comprehension was targeted by the comprehension tasks. Thus, an understanding of proper names might not be important to global comprehension. This finding supports results from a study with L2 listeners (Kobeleva, 2012), where understanding of proper names was beneficial for detailed comprehension, but had no effect on overall comprehension.

Given that no effect on higher-level comprehension processes was found, along with indications from the Interview Study that context is not always useful to identify proper name referents, the focus progressed to proper names as an aspect of lower-level processing. In the Context Study (Chapter 6), participants' semantic and orthographic processing was disrupted, constraining them to rely on contextual clues to identify proper names. It was found that the participants were not very successful at using context to identify proper names; they were able to use context to identify proper names in less than a third of cases. This finding challenges the assumption that L2 readers can understand proper names from context. The implications of the findings from these five studies for L2 vocabulary research, pedagogy and materials were presented in the Discussion chapter. In short, researchers, teachers and material developers should not assume that L2 readers can easily recognise and understand proper names from context. Furthermore, the difficulties that L2 readers have with proper names seem to occur with lower-level processes, not higher-level.

In terms of generalising the findings from the five studies presented here, there are two important variables to note in this research context. First, the participant group, as Japanese intermediate L2 readers of English, have impacted the findings with respect to their L1, and their L2 orthographic processing experience. Thus, while the findings indicate that proper names can present a burden to L2 readers, it important to qualify these results with regard to these participants. It may be that different results are obtained for readers whose L1 is not orthographically distant from their L2, or for readers with more L2 processing experience. This points to further opportunities for research in this area, for example, with L2 readers who have different L1 orthographies and proficiency levels. Secondly, the types of reading texts and comprehension tasks used may have also influenced the findings. For example, understanding of proper name referents may be less important to comprehension of academic texts than of literary texts or newspaper articles. Also, familiarity with proper names may be less important for global comprehension than for detailed comprehension.

8.3 Prospects for future research

The studies in this thesis have examined L2 proper name processing from some different perspectives. A few possibilities to further this research have been briefly mentioned in the discussion sections of Chapters 3 to 6, as well as in the Discussion chapter (Chapter 7). For instance, it is worth considering: whether some types of proper names more problematic than others; how different reading contexts might affect proper name recognition; and individual learner differences in proper name processing. However, there remain many different avenues for future research, avenues that were not directly explored in this thesis. Here, several proposals are presented for future research into how L2 readers process proper names. These proposals are meant to demonstrate to the reader various possibilities that exist. Specifically, more research is needed in the areas of: orthographic processing; semantic processing (i.e. how L2 readers build information about proper names); and the efficiency and automaticity of L2 proper name processing.

8.3.1 Orthographic processing of proper names

One area that has not received much attention in L2 reading research is orthographic processing (Nassaji, 2014). This aspect of proper name processing was not directly examined in this thesis. Since L2 readers arguably have less print exposure than L1 readers, it seems informative to look at the role of orthographic processing in L2 proper name recognition. 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). To summarise their findings, the authors found that proper names with an initial capital letter are recognised faster than common nouns with an initial capital letter, and faster than proper names and common nouns with lower case first letter. Proper names and common nouns with the first letter in lower case are recognised at the same speed, and at the same speed as common nouns with an initial capital letter. The findings suggest a "facilitation effect for proper names with the first letter capitalized" (p. 106). Because no advantage was seen 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).

It is interesting to consider whether this facilitation effect also exists for L2 users. As was reviewed in Chapter 4 (see section 4.2), part of an assumption in L2 vocabulary research is that the initial capital letter on proper names will help L2 readers to easily recognise names. However, it is not known whether the initial capital facilitates proper name recognition for L2 users whose L1 employs a different writing script or orthography (e.g. in German, all nouns are capitalised while in English, only proper names and related adjectives are capitalised).

Proficiency level may also be a variable for how efficiently L2 readers orthographically process proper names in continuous text. There is evidence that as L2 proficiency increases, orthographic knowledge is used more than phonological knowledge (e.g.Chikamatsu, 2006; Nassaji, 2003a). Thus, it might be the case that L2 users with limited processing experience do not process upper case and lower case letters as efficiently and automatically as L1 users. For example, letters that look similar in both cases (e.g. Cc, Oo, Zz) might be more difficult to identify in continuous text than letters that are dissimilar (e.g. Aa, Bb, Ee). As Alderson (2000) notes, "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, the assumption that L2 readers can easily identify proper names in continuous text by the initial capital letter needs to be confirmed with empirical data. And this might be particularly relevant for readers whose L1 employs a different writing script.

One possible approach to investigating L2 orthographic processing is a timed response to the task: *are these two letters the same grapheme?* (Peressotti et al., 2003, p. 90). Responses to combinations of letters are timed and recorded as affirmative (e.g. *a A*) or not (e.g. *e A*). It is possible that an effect is seen for similar looking letters. For example, Thompson (2009) refers to data with L1 children that indicates letters are learned faster when cases are similar looking (e.g. Yy, Cc, Jj) than for letters that look dissimilar, though there were exceptions (i.e. Dd had a small lag; Xx and Uu had a large lag). It is difficult to predict how L2 users with a non-alphabetic L1 background process differences in case letters. Their processing experience is very different from L1 children. Thompson (2009) notes that L1 children learn upper case letters faster,

and lag behind in learning lower case letters; no current research can explain why (p. 53). He suggests one possibility that upper case letters are more salient on the page. In sum, investigations into L2 orthographic processing are underrepresented, and several options exist to fill this gap by looking at orthographic processing of L2 proper names.

8.3.2 Semantic processing of proper names

Part of the debate surrounding proper names is whether they have meaning (see Chapter 2). For some linguists (e.g. Allerton, 1987; J. M. Anderson, 2007; Van Langendonck, 2007), proper names do have associative sense or meaningfulness, and in this way, they contribute to the meaning of the sentence or text. For Van Langendonck (2007), whether proper names have meaning is the wrong question: a better question is how the meanings of names are construed by the reader and function in the text (p. 38). Under this conceptualisation, it would be informative to investigate how L2 readers process and build meanings of proper names from each encounter in a text. This aspect of proper name processing was also not directly examined in this thesis.

By tracking L2 readers' look-up behaviour, one could investigate how readers build understanding of proper names. Different tools can be used to track look-ups. For example, 'reflash' is a text modification tool (B. Wang, 2017), which was designed to assist readers in using context to infer lexical meaning. Target words in a text are attached with reflash, which is a set of right and left pointing arrows; it operates similarly to the navigation function (Ctrl+F) in Microsoft Word. If the reader chooses, she can click on the reflash left button on a target item to see how the word was used previously in the same text; if she presses the right arrow, she will see the next occurrence of the same word. The reader can easily return to where she was reading by pressing the left arrow again. Unlike glosses, a more common form of text modification, reflash does not indicate to the reader the meaning of the target item. Rather, reflash aids the reader by quickly showing all the occurrences of the target item in a given text. Thus, the reader's use of reflash for proper names could be tracked and recorded to investigate how proper names are construed. Follow-up interviews could be done to check what was understood about particular names.

There are less intrusive ways to track proper name look-up behaviour. Some computer programs do not draw attention to particular vocabulary; rather, they only record any items that are clicked on. An example of one such tool is 'the social reader' (Garner, 2018), which allows teachers and researchers to track the vocabulary their learners look up. Teachers can also choose which definitions to include on their particular text, which in the case of proper names might be an interesting feature. If a large amount of reading was done in this way, patterns in look-ups of proper names might be identified. Similar to the Unknown Vocabulary Study (Chapter 4), low-tech approaches can also be used to investigate look-up behaviour.

Another methodology that could be used to investigate how readers build meaning of proper names is self-paced reading. Self-paced reading is normally used to investigate syntactic parsing, and it tests claims about comprehension processes (McDonough & Trofimovich, 2012). Because syntactic parsing is affected by lexis and semantics, this method might be used to explore how proper names are processed and understood in context. The tasks used in self-paced reading focus on comprehension; so, to investigate proper names, the task could focus on how the reader progressively builds meaning of proper names. There are some limitations to self-paced reading methodology: it is unfamiliar and slow-paced reading, and the reader may have extra time to engage other processes that are not normally used (Rayner, Pollatsek, Ashby, & Clifton, 2012, p. 221).

8.3.3 Efficiency and automaticity of L2 proper name processing

As noted, one of the limitations to the studies presented here is that none of the reading tasks included a timed element: that is, participants were given ample time to read the texts and complete the tasks. Therefore, it remains unknown how efficiently and automatically L2 proper names are processed. Research has shown that even advanced bilinguals read 30% more slowly in their L2 than L1 readers; furthermore, this research suggest 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 et al., 1991). For this reason, it would be informative to investigate how efficiently L2 proper names are processed. Segalowitz et al. (1991) report on methods used for measuring word reading time, such as primed lexical decision tasks, and an adapted version of the word superiority effect paradigm

to investigate letter processing (p. 19). Such methods might be employed to explore automaticity of L2 proper name processing.

Another possible method for testing automaticity of proper name processing is a dual task paradigm (Akamatsu, 2008). In this paradigm, a primary task occurs continuously (e.g. deciding whether a proper name appears in a sentence). A secondary task occurs randomly (e.g. pressing a button when a sound is heard). The timed response to the secondary task is taken as an indication of how fully automatised the primary task is processed. If the primary task is carried out automatically, then the second task should be responded to faster because fewer attentional resources are needed for the first task. The performance on the secondary task alone can used as a baseline to measure against the performance on the secondary task in the dual task mode (Akamatsu, 2008, pp. 176, 177). These are some possibilities for learning more about how efficiently L2 proper names are processed.

8.4 Importance of research into L2 proper name processing

I have shown in this investigation that some L2 readers are encumbered by unfamiliar proper names, and that this strain seems to occur with lower-level processing. This finding is significant because of a misleading assumption in L2 vocabulary research that L2 readers can easily understand the proper names they encounter. Treating proper names as known vocabulary in lexical analyses can have a significant impact on a profile of a text. And because all proper names in a text are not likely to be understood or correctly inferred from context, treating proper names as known vocabulary gives an inaccurate indication of the readability or difficulty of a given text. Furthermore, treating proper names as known vocabulary in research contexts may give practitioners and materials writers an imprecise impression that proper names are low-burden items for L2 readers. Therefore, the key findings from this investigation, that L2 readers can have difficulties with proper names, in particular with lower-level processing, have far-reaching implications for research, pedagogy and materials development.

As processors of a second culture and language, L2 readers will certainly encounter unfamiliar proper names. As Hanks (2013) reports, “in some large lexical databases, aiming at full coverage of a language, over 70% of lexical entries already are proper names, and this percentage continues to increase” (p. 35, 36). This statistic is a powerful indicator that L2 readers will most definitely meet unfamiliar proper names in

the target language. Rather than assuming proper names are known, further research is needed in how L2 readers process these special kinds of words. In conclusion, this investigation has shown that the difficulties L2 readers can have with proper names likely lie with lower-level processing, and in doing so, a big step forward has been taken with regard to the assumption in L2 vocabulary research that proper names are easily understood. Because the difficulty L2 readers can have with proper names seems to lie with lower-level processing, this points to a need for further research and a pedagogical focus on lower-level processes in L2 reading.

References

2000 Census. (2000). Retrieved October 28, 2014, from United States Census Bureau http://www.census.gov/genealogy/www/data/2000surnames/index.html

Akamatsu, N. (1999). The effects of first language orthographic features on word-recognition processing in English as a second language. *Reading and Writing: An Interdisciplinary Journal, 11*, 381-403. doi:10.1023/A:1008053520326

Akamatsu, N. (2002). A similarity in word-recognition procedures among second language readers with different first language backgrounds. *Applied Psycholinguistics, 23*, 117-134. doi:10.1017.S0142716402000061

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

Akamatsu, N. (2008). The effects of training on automatization of word recognition in English as a foreign language. *Applied Psycholinguistics, 29*(02), 175-193. doi:10.1017/s0142716408080089

Alba, J., & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin, 93*, 203-231. doi:10.1037/0033-2909.93.2.203

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

Alptekin, C. (2006). Cultural familiarity in inferential and literal comprehension in L2 reading. *System, 34*(4), 494-508. doi:10.1016/j.system.2006.05.003

Anderson, J. M. (2007). *The grammar of names*. Oxford: Oxford University Press.

Anderson, R. C., & Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading comprehension. In P. D. Pearson (Ed.), *Handbook of reading research* (Vol. 1, pp. 255-291). New York: Longman.

the Arctic. (n.d.). Retrieved February 26, 2016, from Google Translate https://translate.google.com/

Balota, D., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology, 17*, 364-390. doi:10.1016/0010-0285(85)90013-1

Bartlett, F. (1932). *Remembering: A study in experimental and social psychology*. Cambridge: Cambridge University Press.

Bassetti, B. (2008). Orthographic input and second language phonology. In T. Piske & M. Young-Scholten (Eds.), *Input matters in SLA* (pp. 191-206). Clevedon, UK: Multilingual Matters.

Bassetti, B., & Atkinson, N. (2015). Effects of orthographic forms on pronunciation in experienced instructed second language learners. *Applied Psycholinguistics, 36*(1), 67-91. doi:10.1017/S0142716414000435

Batuman, E. (2010). *The possessed: Adventures with Russian books and the people who read them*. London: Granta.

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

Bernhardt, E. (2005). Progress and procrastination in second language reading. *Annual Review of Applied Linguistics, 25*, 133-150. doi:10.1017/S0267190505000073

Birch, B. M. (2007). *English L2 reading: Getting to the bottom* (2nd ed.). New Jersey: Lawrence Erlbaum Associates.
Birch, B. M. (2015). *English L2 reading: Getting to the bottom* (3rd ed.). New York: Routledge.
Brennen, T. (1993). The difficulty with recalling people's names: The plausible phonology hypothesis. *Memory, 1*(4), 409-431. doi:10.1080/09658219308258246
Brown, D. (2010). An improper assumption? The treatment of proper nouns in text coverage counts. *Reading in a Foreign Language, 22*(2), 355 - 361.
Brown, T., & Haynes, M. (1985). Literacy background and reading development in a second language. In H. Carr (Ed.), *The development of reading skills*. San Francisco: Jossey-Bass.
Bultena, S., Dijkstra, T., & Van Hell, J. G. (2015). Language switch costs in sentence comprehension depend on language dominance: Evidence from self-paced reading. *Bilingualism: Language and Cognition, 18*, 453-469. doi:10.1017/S1366728914000145
Carrell, P. L. (1983). Some issues in studying the role of schemata, or background knowledge, in second language comprehension. *Reading in a Foreign Language, 1*(2), 81-92.
Carver, R. (1992). Commentary: Effect of prediction activities, prior knowledge, and text type on the amount of comprehension: Using rauding theory to critique schema theory research. *Reading Research Quarterly, 27*, 164-174. doi:10.2307/747685
Carver, R. (1994). Percentage of unknown vocabulary words in text as a function of the relative difficulty of the text: Implications for instruction. *Journal of Reading Behavior, 26*(4), 413 - 437. doi:10.1080/10862969409547861
Chen, H. C., & Graves, M. F. (1995). Effects of previewing and providing background knowledge on Taiwanese college students' comprehension of American short stories. *TESOL Quarterly, 29*, 663-686. doi:10.2307/3588168
Chen, H. C., & Tsoi, K. C. (1990). Symbol-word interference in Chinese and English. *Acta Psychologica, 75*, 123-128. doi:10.1016/0001-6918(90)90082-Q
Chikamatsu, N. (1996). The effects of L1 orthography on L2 word recognition: A study of American and Chinese learners of Japanese. *Studies in Second Language Acquisition, 18*(4), 403-432. doi:10.1017/S0272263100015369
Chikamatsu, N. (2006). Developmental word recognition: A study of L1 English readers of L2 Japanese. *The Modern Language Journal, 90*, 67-85. doi:10.1111/j.1540-4781.2006.00385.x
Coates, R. A. (2006). Properhood. *Language, 82*(2), 356-382. doi:10.1353/lan.2006.0084
Cobb, T. (2010). Learning about language and learners from computer programs. *Reading in a Foreign Language, 22*(1), 181-200.
Cobb, T. (n.d.). VP-Compleat [computer program] (Version 2). Retrieved from https://www.lextutor.ca/vp/comp/
Cook, V. (2002). Background to the L2 user. In V. Cook (Ed.), *Portraits of the L2 user*. Clevedon: Multilingual Matters.
Cook, V., & Bassetti, B. (2005). *Second language writing systems*. Clevedon: Multilingual Matters Ltd.
Coulmas, F. (2003). *Writing systems: An introduction to their linguistic analysis*. Cambridge: Cambridge University Press.
Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*(2), 213-238. doi:10.2307/3587951
Crawford, M. (2005). Adding variety to word recognition exercises. *English Teaching Forum, 43*(2), 36-41.
Crystal, D. (2006). *Words, words, words*. New York: Oxford.

Cunningham, A. E., Perry, K. E., & Stanovich, K. E. (2001). Converging evidence for the concept of orthographic processing. *Reading and Writing: An Interdisciplinary Journal, 14*, 549-568. doi:10.1023/A:1011100226798

Cutting, L. E., & Scarborough, H. S. (2006). Prediction of reading comprehension: Relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured. *Scientific Studies of Reading, 10*(3), 277-299.

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/

Day, R. R., & Bamford, J. (1998). *Extensive reading in the second language classroom*. Cambridge: Cambridge University Press.

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

Dehaene, S. (2009). *Reading in the brain: The new science of how we read*. New York: Penguin.

Demetriou, A., Kui, Z. X., Spandoudis, G., Kyriakides, L., & Platsidou, M. (2005). The architecture, dynamics, and development of mental processing: Greek, Chinese, or universal? *Intelligence, 33*, 109-141. doi:10.1016/j.intell.2004.10.003

Dijkstra, T. (2005). Bilingual visual word recognition and lexical access *Handbook of bilingualism: Psycholinguistic approaches* (pp. 179-201).

Dornyei, Z. (2007). *Research methods in applied linguistics*. Oxford: Oxford University Press.

Douglas, S. R. (2015). The relationship between lexical frequency profiling measures and rater judgements of spoken and written general English language proficiency on the CELPIP-General Test. *TESL Canada Journal, 32*(9), 43-64. doi:10.18806/tesl.v32i0.1217

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

Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition, 24*, 143-188. doi:10.1017.S0272263102002024

Elston-Güttler, K. E., Paulmann, S., & Kotz, S. A. (2005). Who's in control? Proficiency and L1 influence on L2 processing. *Journal of Cognitive Neuroscience, 17*(10), 1593-1610.

Erten, I. H., & Razi, S. (2009). The effects of cultural familiarity on reading comprehension. *Reading in a Foreign Language, 21*(1), 60-77.

Eskey, D. (1988). Holding in the bottom: An interactive approach to the language problems of second language readers. In P. L. Carrell, J. Devine, & D. Eskey (Eds.), *Interactive approaches to second language reading* (pp. 93-100). Cambridge: Cambridge Univesrity Press.

Francis, W. N., & Kucera, H. (1982). *Frequency analysis of English usage*. Boston: Houghton Mifflin Company.

Fraser, C. A. (1999). Lexical processing strategy use and vocabulary learning through reading. *Studies in Second Language Acquisition, 21*, 225-241.

Frege, G. (1892 [1952]). On Sense and Reference. In P. Geach & M. Black (Eds.), *Translations from the Philosophical Writings of Gottlob Frege*. Oxford: Blackwell.

Garner, J. (2018). The social reader. Retrieved from http://vocabkitchen.com/

Goodman, K. S. (1967). Reading: A psycholinguistic guessing game. *Literacy Research and Instruction, 6*, 126-135.
Goodman, K. S. (1969). Analysis of oral reading miscues: Applied psycholinguistics. *Reading Research Quarterly, 5*(1), 9-30. doi:10.2307/747158
Goodman, K. S. (1988). The reading process. In P. L. Carrell, J. Devine, & D. Eskey (Eds.), *Interactive approaches to second language reading* (pp. 11-22). Cambridge: Cambridge University Press.
Goodman, K. S. (1996). *On reading*. Portsmouth, N.H.: Heinemann.
Gough, P. B. (1972). One second of reading. In J. F. Kavanaugh & I. G. Mattingly (Eds.), *Language by ear and by eye: The relationships between speech and reading* (pp. 331-358). Cambridge, M.A.: MIT Press.
Grabe, W. (1991). Current developments in second language reading research. *TESOL Quarterly, 25*, 375-406. doi:10.2307/3586977
Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. Cambridge: Cambridge University Press.
Grabe, W., & Stoller, F. L. (2002). *Teaching and researching reading*. Harlow, England: Pearson Education.
Grabe, W., & Stoller, F. L. (2011). *Teaching and researching reading* (2nd ed.). England: Pearson.
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.
Hall, D. G. (1996). Preschoolers' default assumptions about word meaning: Proper names designate unique individuals. *Developmental Psychology, 32*(1), 177-186. doi:10.1037/0012-1649.32.1.177
Hanks, P. (2013). *Lexical analysis: Norms and exploitations*. Cambridge: MIT Press.
Hanley, J. R., & Huang, H. S. (1997). Phonological awarenss and learning to read Chinese. In C. K. Leong & R. M. Joshi (Eds.), *Cross-language studies of learning to read and spell: phonologic and orthographic processing* (pp. 361-379). Dordrecht: Kluwer.
Hatch, J. A. (2002). *Doing qualitative research in education settings*. New York: State University of New York Press.
Hill, D. R. (2013). Survey review: Graded readers. *ELT Journal, 67*(1), 85-125. doi:10.1093/elt/ccs067
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.
Horst, M. (2013). Mainstreaming second language vocabulary acquisition. *The Canadian Journal of Applied Linguistics, 16*(1), 171-188.
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
Hu, M. H., & Nation, I. S. P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language, 13*(1), 403 - 430.
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.
Huddleston, R. (1984). *Introduction to the grammar of English*. Cambridge: Cambridge University Press.

Hudson, T. (2007). *Teaching second language reading*. Oxford: Oxford University Press.
Hulstijn, J. H. (1993). When do foreign-language readers look up the meaning of unfamiliar words? The influence of task and learner variables. *The Modern Language Journal, 77*(2), 139-147. doi:10.1111/j.1540-4781.1993.tb01957.x
Hulstijn, J. H. (2001). Intentional and incidental second language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), *Cognition and second language instruction*. Cambridge: Cambridge University Press.
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.
James, L. E., & Fogler, K. A. (2007). Meeting Mr Davis vs Mr Davin: Effects of name frequency on learning proper names in young and older adults. *Memory, 15*(4), 366-374. doi:10.1080/09658210701307077
Japan's top 100 most common family names. (2009). *The Japan Times*. Retrieved from https://www.japantimes.co.jp/
Jiang, N. (2000). Lexical representation and development in a second language. *Applied Linguistics, 21*(1), 47-77.
Jiang, X. (2012). Effects of discourse structure graphic organizers on EFL reading comprehension. *Reading in a Foreign Language, 24*(1), 84-105.
Johnson, P. (1981). Effects on reading comprehension of language complexity and cultural background of a text. *TESOL Quarterly, 15*(2), 169-181. doi:10.2307/3586408
Ketchum, E. M. (2006). The cultural baggage of second language reading: An approach to understanding. *Foreign Language Annals, 39*, 22-42. doi:10.1111/j.1944-9720.2006.tb02247.x
King, J. (2013). Silence in the second language classrooms of Japanese universities. *Applied Linguistics, 34*(3), 325 - 343. doi:10.1093/applin/ams043
Kobeleva, P. (2012). Second language listening and unfamiliar proper names: Comprehension barrier? *RELC Journal, 43*(1), 83-98. doi:10.1177/0033688212440637
Koda, K. (1994). Second language reading research: Problems and possibilities. *Applied Psycholinguistics, 15*, 1-28. doi:10.1017/S0142716400006950
Koda, K. (1995). Cognitive consequences of L1 and L2 orthographies. In I. Taylor & D. Olson (Eds.), *Scripts and literacy* (pp. 311-326). The Netherlands: Kluwer.
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. (1999). Development of L2 intraword orthographic sensitivity and decoding skills. *The Modern Language Journal, 83*(1), 51-64. doi:10.1111/0026-7902.00005
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.
Koda, K. (2012). How to do research on second language reading. In A. Mackey & S. Gass (Eds.), *Research methods in second language acquisition: A practical guide*. UK: Wiley-Blackwell.
Koda, K. (2013). Second language reading, scripts, and orthographies. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics*: Blackwell Publishing.
LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*, 293-323. doi:10.1016/0010-0285(74)90015-2

Laufer, B. (1989). What percentage of text-lexis is essential for comprehension? In C. Lauren & M. Nordman (Eds.), *Special language: From humans to thinking machines* (pp. 316-323). Clevedon, England: Multilingual Matters.
Laufer, B., & Paribakht, T. (1998). The relationship between passive and active vocabularies: Effects of language learning context. *Language Learning, 48*(3), 365-391. doi:10.1111/0023-8333.00046
Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language, 22*(1), 15-30.
Liu, N., & Nation, I. S. P. (1985). Factors affecting guessing vocabulary in context. *RELC Journal, 16*, 33-42. doi:10.1177/003368828501600103
Lyons, J. (1977). *Semantics* (Vol. 1). Cambridge: Cambridge University Press.
Macalister, J. (2010). Speed reading courses and their effect on reading authentic texts: A preliminary investigation. *Reading in a Foreign Language, 22*(1), 104 - 116.
Matsuoka, W., & Hirsh, D. (2010). Vocabulary learning through reading: Does an ELT course book provide good opportunities? *Reading in a Foreign Language, 22*(1), 56 - 70.
McDonough, K., & Trofimovich, P. (2012). How to use psycholinguistic methodologies for comprehension and production. In A. Mackey & S. Gass (Eds.), *Research methods in second language acquisition: A practical guide.* UK: Wiley-Blackwell.
Mclean, S., Hogg, N., & Kramer, B. (2014). Estimations of Japanese university learners' English vocabulary size using the Vocabulary Size Test. *Vocabulary Learning and Instruction, 3*(2), 47-55. doi:10.7820/vli.v03.2.2187-2759
McPhee, J. (2015, March 9). Frame of reference. *The New Yorker*. Retrieved from https://www.newyorker.com/
Mill, J. S. (1865). *A system of logic, ratiocinative and inductive* (6th ed. Vol. 1). London: Longmans, Green and Co.
Muljani, D., Koda, K., & Moates, D. (1998). The development of word recognition in a second language. *Applied Psycholinguistics, 19*(1), 99-113. doi:10.1017/S0142716400010602
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
Nakada, T., Fujii, Y., & Kwee, I. L. (2001). Brain strategies for reading in the second language are determined by the first language. *Neuroscience Research, 40*, 351-358. doi:10.1016/S0168-0102(01)00247-4
Nakamura, K., Dehaene, S., Jobert, A., Le Bihan, D., & Kouider, S. (2005). Subliminal convergence of Kanji and Kana words: Further evidence for functional parcellation of the posterior temporal cortex in visual word perception. *Journal of Cognitive Neuroscience, 17*(6), 954-968. doi:10.1162/0898929054021166
Nassaji, H. (2002). Schema theory and knowledge based processes in second language reading comprehension: A need for alternative perspectives. *Language Learning, 52*, 439-482. doi:10.1111/0023-8333.00189
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. (2008). *Teaching vocabulary: Strategies and techniques*. USA: Heinle Cengage Learning.
Nation, I. S. P. (2009). New roles for L2 vocabulary? In V. Cook & L. Wei (Eds.), *Contemporary applied linguistics: Language teaching and learning* (Vol. 1). London: Continuum.
Nation, I. S. P. (2018). Reading a whole book to learn vocabulary. *ITL - International Journal of Applied Linguistics, 169*(1), 30-43. doi:10.1075/itl.00005.nat
Nation, I. S. P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher, 31*(7), 9-13.
Nation, I. S. P., & Heatley, A. (2002). Range: A program for the analysis of vocabulary in texts [software]. Retrieved from http://www.vuw.ac.nz/lals/staff/Paul_Nation
Nation, I. S. P., & Wang, K. M. (1999). Graded readers and vocabulary. *Reading in a Foreign Language, 12*(2), 355 - 380.
Nation, I. S. P., & Webb, S. (2011). *Researching and analyzing vocabulary*. Boston: Heinle Cengage Learning.
Oller, J. W. (1995). Adding abstract to formal and content schemata: Results of recent work in Peircean semiotics. *Applied Linguistics, 16*, 273-306. doi:10.1093/applin/16.3.273
Pang, K. S. (2010). Eponymy and life-narratives: The effect of foregrounding on proper names. *Journal of Pragmatics, 42*, 1321-1349. doi:10.1016/j.pragma.2009.09.023
Parent, K. (2016). [BNC proper noun frequency list]. Unpublished raw data.
Park, H. S. (2006). Structural characteristics of proper nouns in Korean-Swedish discourse. *International Journal of Bilingualism, 10*(1), 17-36. doi:10.1177/13670069060100010201
Parry, K. (1996). Culture, literacy and L2 reading. *TESOL Quarterly, 30*(4), 665-689. doi:10.2307/3587929
Pearson, D. P., & Johnson, D. (1978). *Teaching reading comprehension*. New York: Holt, Rinehart, and Winston.
Perea, M., Marcet, A., & Vergara-Martinez, M. (2016). Does top-down feedback modulate the encoding of orthographic representations during visual-word recognition? *Experimental Psychology, 63*(5), 278-286. doi:10.1027/1618-3169/a000327
Perea, M., & Rosa, E. (2002). Does "whole-word shape" play a role in visual word recognition? *Perception & Psychophysics, 64*(5), 785-794. doi:10.3758/BF03194745
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. (1985). *Reading ability*. New York: Oxford University Press.
Perfetti, C. (1986). Cognitive and linguistic components of reading ability. In B. Foorman & A. Siegel (Eds.), *Acquisition of reading skills: Cultural constraints and universals* (pp. 11-40). Hillsdale, N. J.: L. Erlbaum.
Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*(4), 357-383. doi:10.1080/10888430701530730
Perfetti, C., & Liu, Y. (2005). Orthography to phonology and meaning: Comparisons across and within writing systems. *Reading and Writing, 18*, 193-210. doi:10.1007/s11145-004-2344-y
Porte, G. (Ed.) (2012). *Replication research in applied linguistics*. Cambridge: CUP.
Proper noun. (n.d.). *English Oxford Living Dictionaries*. Retrieved from https://en.oxforddictionaries.com/definition/proper_noun

Pulido, D. (2003). Modeling the role of second language proficiency and topic familiarity in second language incidental vocabulary acquisition through reading. *Language Learning, 53*, 233-284. doi:10.1111/1467-9922.00217
Pulido, D. (2009). Vocabulary processing and acquisition through reading: Evidence for the rich getting richer. In Z. Han & N. J. Anderson (Eds.), *Second language reading research and instruction: Crossing the boundaries* (pp. 65-81). Ann Arbor: The University of Michigan Press.
Quirk, R. (Ed.) (1995) Longman dictionary of contemporary English (3rd ed.). Harlow, England: Longman.
Rayner, K., Pollatsek, A., Ashby, J., & Clifton, C. (2012). *Psychology of reading* (2nd ed.). New York: Psychology Press.
Razi, S. (2005). A fresh look at the evaluation of ordering tasks in reading comprehension: Weighted marking protocol. *The Reading Matrix, 5*(1), 1-15.
Reynolds, R. E., Taylor, M. A., Steffensen, M. S., Shirey, L. L., & Anderson, R. C. (1982). Cultural schemata and reading comprehension. *Reading Research Quarterly, 17*, 353-366. doi:10.2307/747524
Rumelhart, D. E. (1977). Toward an interactive model of reading. In S. Dornic (Ed.), *Attention and performance* (Vol. 6, pp. 573-603). Hillsdale, NJ: Erlbaum.
Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. E. Brewer (Eds.), *Theoretical issues in reading comprehension*. Hillsdale, NJ: Lawrence Erlbaum Associates.
Ryan, A. (1997). Learning the orthographical form of L2 vocabulary - a receptive and a productive process. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 181-195). Cambridge: CUP.
Sadoski, M., & Paivio, A. (2007). Toward a unified theory of reading. *Scientific Studies of Reading, 11*(4), 337-356. doi:10.1080/10888430701530714
Sadoski, M., Paivio, A., & Goetz, E. (1991). Commentary: A critique of schema theory in reading and a dual coding alternative. *Reading Research Quarterly, 26*, 463-484. doi:10.2307/747898
Sanford, A. J., Moar, K., & Garrod, S. (1988). Proper names as controllers of discourse focus. *Language & Speech, 31*, 43-56.
Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal, 95*, 26-43. doi:10.1111/j.1540-4781.2011.01146.x
Schmuckler, M. A. (2001). What is ecological validity? A dimensional analysis. *Infancy, 2*(4), 419 - 436. doi:10.1207/S15327078IN0204_02
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.
Seidenberg, M., & McClelland, J. (1989). A distributed, development model of word recognition. *Psychological Review, 96*, 523-568.
Shaw, I. (2000). *Short stories: Five decades*. Chicago: University of Chicago Press.
Shcherba, L. V. (1940 [1995]). Opyt obshchei teorii leksikografii. [Towards a general theory of lexicography]. *International Journal of Lexicography, 8*(4), 314-350.
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.
Stanovich, K. E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers*. New York: The Guilford Press.
Steffensen, M. S., Joag-Dev, C., & Anderson, R. C. (1979). A cross-cultural perspective on reading comprehension. *Reading Research Quarterly, 15*(1), 10-29. doi:10.2307/747429
Strawson, P. F. (1950). On referring. *Mind, 59*(235), 320-344.

Tan, L. H., Liu, H. L., Perfetti, C. A., Spinks, J. A., Fox, P. T., & Gao, J. H. (2001). The neural system underlying Chinese logograph reading. *NeuroImage, 13*, 836-846. doi:10.1006/nimg.2001.0749

Thompson, G. B. (2009). The long learning route to abstract letter units. *Cognitive Neuropsychology, 26*(1), 50-69. doi:10.1080/02643290802200838

Titone, D., Libben, M., Mercier, J., Whitford, V., & Pivneva, I. (2001). Bilingual lexical access during L1 sentence reading: The effects of L2 knowledge, semantic constraint, and L1-L2 intermixing. *Journal of Experimental Psychology: Learning, Memory and Cognition, 37*(6), 1412-1431.

Tracey, D. H., & Morrow, L. M. (2017). *Lenses on reading: An introduction to theories and models*. New York: The Guilford Press.

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.

Van Langendonck, W. (2007). *Theory and typology of proper names*. Berlin: Mouton de Gruyter.

Van Langendonck, W., & Van de Velde, M. (2016). Names and grammar. In C. Hough (Ed.), *The Oxford handbook of names and naming*. Oxford: Oxford University Press.

Wang, B. (2017). A novel type of text modification and its implications for vocabulary learning through reading. *Oxford Education Cloud, 1*(1), 93-107.

Wang, M., & Koda, K. (2007). Commonalities and differences in word identification skills among learners of English as a second language. *Language Learning, 57*, 201-222. doi:10.1111/j.1467-9922.2007.00416.x

Wang, M., Koda, K., & Perfetti, C. (2003). Alphabetic and nonalphabetic L1 effects in English word identification: A comparison of Korean and Chinese English L2 learners. *Cognition, 87*, 129-149. doi:10.1016/s0010-0277(02)00232-9

Waring, R. (2014). *Priniciples of vocabulary: Matsuyama JALT*. Presentation. http://www.robwaring.org/presentations.

Webb, S., & Macalister, J. (2013). Is text written for children useful for L2 extensive reading? *TESOL Quarterly, 47*(2), 300-322. doi:10.1002/tesq.70

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

West, M. (1953). *A general service list of English words*. London: Longman, Green & Co.

Wray, A., & Bloomer, A. (2012). *Projects in linguistics and language studies* (3rd ed.). Great Britain: Hodder Education.

Appendices

Appendix 1: L2 readers' perspectives on proper names

1.1 Invitation to interview

Dear (student),

I'm doing research on how Japanese learners read in English. I was wondering if you would like to take part in an interview to help me with this research.

The interview would be in my office, and it would last 30 minutes at most. It would involve you answering my questions about how you read in English. I would keep your name and answers confidential (secret).

If you would rather NOT, that's no problem. It will NOT affect your grade in anyway whether you participate in the interview.

If you would like to do the interview, let me know a day and time that is convenient for you.

Regards,

1.2 Text excerpt for read-aloud task

Larry's Idea

After July came the cold winds and the grey skies of August. My family had all their usual illnesses. My mother and I had bad colds. My brother Leslie had a problem with his ears. My sister Margo's spots were worse than ever. Only my oldest brother, Larry, was healthy, but he found the rest of us very difficult to live with.

'Why do we stay in England in this weather?' he asked Mother. '*They're* ill and *you're* looking older every day.'

'I'm not,' Mother replied. She was reading at the time.

'You are,' Larry said. 'We need sunshine…a country where we can grow.'

'Yes dear, that's a good idea,' Mother answered, not really listening.

'George says Corfu's wonderful. Why don't we go there?'

'If you like, dear.' It was important to keep Larry happy.

'When?' asked Larry with surprise.

Mother realized her mistake and put down her book. 'Perhaps you can go first and look at the place,' she said cleverly. 'If it's nice, we can all follow.'

Larry looked at her. 'You said that last time. I waited in Spain for two months and you didn't come. No – if we're going to Greece, let's go together.'

'But I've only just bought this house!' Mother answered.

'Sell it again then!'

'That's stupid, dear,' said Mother. 'I can't do that.'

So we sold the house and ran from the English summer.

We travelled by train with our clothes and our most important belongings: Mother's cook books, Leslie's guns, something for Margo's spots, Larry's books, my favourite insects and Roger, my dog.

From Italy we caught a boat. We slept when the boat left and then, very early the next morning, we watched for Corfu. The sea turned blue, then purple, and suddenly there was the sleeping island in front of us. We sailed nearer and, above the noise of the ship, we could hear the high, clear sounds of the insects.

1.3 Consent form

Consent form for participating in research

I (Kimberly Klassen) would like to conduct research on reading skills of Japanese learners of English. If you agree to participate in this research, your performance will not affect your grades for this course in anyway. Also, your decision to participate in this research or not will not affect your grade in this course.

Your name will NOT appear in published results of this research. Likewise, your class number will NOT appear in published results. In sum, there will be no way to identify you personally as a participant.

Taking part in the project will involve answering questions about your attitudes and strategies as a Japanese reader of English. You will also be asked to read a short text aloud and talk about your strategies as you read.

You can withdraw from this study afterwards without giving a reason. You are free to ask questions at any time. The data collected from this research will be held confidentially.

日本人学習者の英語のリーディング・スキルに関する調査のご協力をお願いします。この調査での内容が授業の成績に影響することは一切ありません。また、この調査に参加するか否かが成績に影響することもありません。

協力して頂いた方の名前や学籍番号は研究結果の中に記載致しませんので、あなた個人がこの調査の参加者として特定されることはありません。

このプロジェクトに参加するにあたり、どの様な姿勢や戦略を持って、日本語読者が英語読んでいるかについて答えていただきます。また、短い英文を声に出して読んでもらい、どの様な戦略で読んだのかについて質問させていただきます。

どの様な質問をしていただいても構いませんし、参加後でも、理由を述べることなく辞退しても構いません。この調査で集められたデータは機密で管理されます。

If you agree to take part in this research, please sign the consent below:

Student Name: _____ Student No. _____

Date: _____ Signature: _____

1.4 Interview transcriptions

Interview #1 with Student A
December 4, 2014 4.40 - 5.00 (approx. 20 min.)
Interviewer's office at university

I: So thanks A for coming. So today is Dec. 4, and A and I are sitting in my office. And it's 4.30, after class. So as I said, the purpose of my research is to look at how language learners deal with names of people and places when they are reading. So when you are reading something in Japanese, does it ever happen you, like maybe you are reading a newspaper, does it ever happen to you that you come across a Kanji you don't know, like a name of a person? Yeah? What do you do when that happens?

A: I don't care. Because I want to know the article, contents of article. So I don't care if I couldn't read the person's name correctly.

I: So if it's not important to understanding the article, you'll just keep reading. So when you are reading something in English, like maybe a graded reader or textbook or something, what do you do, does it happen you come across names you don't know? Yeah? So what do you do then?

A: Ah (.) sometimes, for example, the last book I read, there are two persons, the names begin with A. I very confused. So in such situation I really concentrate on the name, very look carefully but usually I don't sure how pronounce especially English name or Indian name, it's very difficult. But I can understand the who is he if I can't pronounce so I not so care.

I: Do you think it would help you if you could pronounce the name?

A: Yes, yes. Because the book like Romeo and Juliet, it's very famous, so it's easy to understand ah, this is Romeo's phrase and if I could pronounce the name so I think it help me to read.

I: So this last book that you read, it had two names starting with A, so what was your strategy, what did you do to try to distinguish between the two?

A: Ah, they were related, one of them is a girl, and the other is a man, an uncle of her so often they appear at the same time. The main character called the girl she or her, but uncle is he or him so I could organise the two person. Hmm (.) Yeah. [laughs]

I: Yeah, okay. How do you feel when you are reading a book or something, maybe a textbook, and there are lots of names you aren't sure about? How do you feel?

A: Ah, uncomfortable. Because like today's test, there are reading 1's author (.) *eto* (Japanese: well) (.) ah, when I read the textbook I am not so much care about the person's name. And so in the test, the question says blah blah says and I a little bit confused, uh which reading is it?

I: Right. So you're not paying attention to the author's name. Yeah, uh huh. Good point. How would you describe your level of knowledge of English names, like would you say you know a lot of names in English, or you know a few names in English?

A: A few names.

I: A few names. Ok. Do you think it's important to know about names?

A: Yeah, because I have exchange student, name Tom. Actually his name is Thomas. Often the American or other foreign people say Thomas, Tom and I didn't know such things before I come to here. So sometimes when I'm talking with other people I'm not sure who is he. So it's important.

I: Yeah, we like to use nicknames. That's called nicknames. So my real name is Kimberly but my nickname is Kimi. So I guess if you were reading a book, and you didn't know that that was the same person, it'd get confusing. When you are reading, do you recognise names quickly or easily?

A: No.

I: How do you know, how do you know when there is a name? How do you know it's a name?

A: How, ah, ah, while I'm reading whole books, I can understand a little bit the relationship and then I can organise the name. But at first, new character appeared at first, I can't understand who is he or who is she. And (.) Mmm (.)

I: So what do you do then if you don't know?

A: Sometimes it's very complex relationships I take memo, and name and so, brother or sister.

I: So you do that while you are reading a graded reader. Very good. So you draw like a chart, who is who. That's a great idea. Do you think it's easier to recognise names in English or in Japanese?

A: [Laughs]. Japanese because it's my first language. And yeah, of course I think pronounce is very important. Because if I know how to pronounce the name, it's easy to memorise because I think I can understand from eyes and ears. And I (.) Japanese is Kanji so Kanji usually can read only one or two ways. So one letter has only one or two ways to read. But English has no words so (.) Like Kanji is if I wrote like this (gestures on hand) it's meaning 'one'. But English is o-n-e and English makes the word use many letters so it's difficult. Because 'ph' is pronounced (.) so

I: So the Kanji you could at least make a guess but in English it could be anything really. One reason I wanted to ask you for this interview is because last week when we did timed writing, we talked about learning to read in English. And you mentioned that names (.) if there are lots of names in a story, it's difficult, so I was interested. Do you remember writing that? No? Yeah, we've been so busy lately. So I thought, oh I should ask her about that. Um, is there anything that I didn't ask you yet about names that you think I should have asked you? Is there anything important for you when you are reading, when you come across names, anything you would like to add?

A: Hmm (.) when I'm reading, last month I read the book, the character was British maybe and someone called his first name but her, his, ahh (.) There are three persons: the main character was a girl, and second is her brother, older brother. And he is the friend of the brother. And she called her brother, brother. But he called ah, name, his name, and so sometimes he is, friend is talking about her brother but I confused who is he. Because usually, almost (.) she, the main character call her brother brother, brother, brother in her conversation but sometimes, suddenly appeared the name. So.

I: Ah, so when a character has two names. Yeah, that's interesting. I didn't even think of that. Yeah good. Can I ask you do one more thing then. I wanted you to read just one page of this graded reader. It's called *My Family and Other Animals.* It's Level 3, so kind of normal what you would read. So I wanted you to read just this page aloud, and every time you come to a name, just tell me, what does that refer to. Okay? I will give example. I will give you example. So if I'm reading this book. And then every time I come to a name, I will say who that refers to. (Reading). The train rushed along angrily. Guy was thinking. Guy, that must be a man's name, was thinking about Mariam, Mariam is a girl's name. He saw her pink round face, her cruel mouth, he started to hate her. Perhaps Mariam, that's the girl, doesn't want a divorce, Guy thought. Ah, so maybe they are married. Guy thought unhappily, but she's pregnant and it's not my child but she must want to marry the father. Why does she want to see me though? Okay? So every time you come to a name, tell me what does it refer to. Yeah? So just from this page until here. So can you read aloud for me?

A: (Reading) After July came the cold winds and the grey skis [sic: skies] of August. My family had all their usual illnesses. My mother and I had bad colds. Huh, I describe? no?

I: Yeah, okay, mother? Okay.

A: (Reading) My brother Les, Les, Leslie, he is the brother of main character, had a problem with his ears. My sister Margo's, she is also her younger or older sister, spots were worse than ever. Only my oldest brother, Larry, Larry is older brother of main character, was healthy, but he found the rest of us very difficult to live with. Why do we stay in England in this weather? he asked Mother. They're ill and you're looking older every day. I'm not, Mother replied. She was reading at the time. You are, Larry said. Larry is older, oldest brother and he is healthy. We need sunshine, a country where we can grow. Yes dear, that's a good idea, Mother answered, not really listening. George, George is brother, hmm (.) George? George, says Corfu's wonderful (.) Corfu. George and Corfu, heh (.) [laughs] Why don't we go there? Maybe they are family, members of the family. If you like, dear. It was important to keep Larry happy. Larry is the healthy, oldest brother. When? asked Larry with surprise.
Mother realized her mistake and put down her book. Perhaps you can go first and look at the place, she said cleverly. If it's nice, we can all follow. Larry, the oldest brother, looked at her, her is mother. You said that last time. I waited in Spain for two months and you didn't come. No, if we're going to Greece, let's go together. But I've only just bought this house! Mother answered. Sell it again then! That's stupid, dear, said Mother. I can't do that. So we sold the house and ran from the English summer. We travelled by train with our clothes and our most important belongings: Mother's cook books, Larry's [sic: Leslie's] Larry, Larry is brother who has illness. Huh? Doko (Japanese: where) Hmm (.) Ah yeah, with ears. Larry's guns, something for Margot's spots, Margot is sister, Larry's books, my favourite insects and Roger, (sic: Rogger) my dog, dog. From Italy we caught a boat. We slept when the boat left and then, very early the next morning, we catched [sic: watched] for Corfu, the family. The sea turned blue, then purple, and suddenly there was the sleeping island in front of us. We sailed nearer and, above the noise of the ship, we could hear the high, clear sounds of the insects.

I: Great thank you. Good. So there were quite a few family members.
How about when you see place names, are place names ever difficult for you? Names of places?

A: If I know the place, for example, Greece, or England or Spain is okay. But sometimes appear the place name I don't know, then it's difficult to pronounce because

I can't accept, expect what, how to pronounce. So if the letter was a little bit difficult, I can guess what the country is. But the unknown place is difficult.

I: You've mentioned it a few times now, that if you don't know how to pronounce the name of the person or of the place, it's difficult to remember. So what is your strategy then, like, do you try to find out how to pronounce it? Or what do you do, do you make your own pronunciation in your head?

A: Or I memorise only the form.

I: Okay, interesting. So you know what it looks like. But in your head, you don't have a sound for it.

A: Hmm, hmm.

I: Yeah interesting. How about your dictionary? Does it have any names in it?

A: English dictionary? Ah, sometimes.

I: Have you ever looked up a name and found it there?

A: Ah, maybe I thought the person's name is the place name. I misunderstood it is a place name. So I searched in dictionary. But it was person's, man's name. So dictionary said it's a man's name (.) It has very popular names.

I: And then does your dictionary have pronunciation then of the name? Or no?

A: Maybe not.

I: So it has some popular names. Oh that's good. Because there's a lot of information in these style of dictionaries. Ok good, that's all I wanted to ask you. So that was interesting for me. Anything you want to say? Any comments?

A: Hmm (.) names (.) Especially Indian names is very difficult.

I: Have you read something recently?

A: Yes, I have. Grey Owl.

I: Grey Owl. Ah! Native American Indians. Ah, yes. So why do you find them difficult?

A: The Indian's name is difficult to pronounce and I thought the very similar names.

I: Okay. And do you think they are very long, or? Or what's?

A: Ah, not so long.

I: Okay, just the pronunciation. Grey Owl, okay I'll have to look for that book. Okay, thanks very much, thanks for coming. That was about 20 minutes, so a little bit shorter than I thought. So that's good.

Interview #2 with Student T
December 5, 2014 3.10 – 3.30 (18 min. 13 sec.)
Interviewer's office at university

I: Okay, T and I are in my office and it is December 5th. And it's 3.10. So thank you T for coming.

T: Fine. [laughs]

I: Ah, okay, so as I said, the purpose of my research is to look at how language learners deal with names when they are reading. So the questions I'm going to ask you are about reading and names. Okay? So, let's start with this one. When you read something in Japanese, like a newspaper article or something, does it ever happen to you that you see a name in Kanji that you don't know, you've never seen before?

T: Ah sometimes.

I: Okay. And what do you do when that happens?

T: Ah (.) nothing. Just read. Just keep reading.

I: So even though you don't know the name, that doesn't interfere with your comprehension of the story?

T: Yes but sometimes the name, rare name confuse us. Or I can't read smoothly. Sometimes. But as I read, as I see that name again and again, that makes us feel that's normal. And at last, not influence so much.

I: Ok, so how about in English? When you are reading something in English and you see a name you don't know?

T: Ah, that really confused. Because we don't know which one is family name or first name. So yes, very difficult.

I: Do you think that's important to know which one is family name and which one is first name?

T: Ah, yes, yes, because I think there are some (.) there are specific or style or form already decided, and once we get to know that, we can guess what is the family name or first name.

I: Okay, what do you do when that happens? You are reading something like textbook or a graded reader, what do you do when you come across a name that you don't know? Do you do anything or do you just keep reading?

T: Ah. Just keep reading.

I: Okay. Umm. (.) How would you describe your level of knowledge of English names? Do you know a lot of English names, a few?

T: A few, I think, few.

I: Enough or not enough?

T: Not enough.

I: Okay, so you think it's important to know about English names, for example, which names are for men, which names are for women?

T: Yes.

I: Would that help you understand (.) the thing you are reading?

T: Hmm, yes I think so. Because the more we get to know the name, the more we can get easily remember and remember the story too. So yes it's important I think.

I: So is it important to remember names?

T: Er (.) Yes.

I: So do you make a point of trying to remember new names?

T: No. [laughs]

I: Okay, but it might help you. Okay. So let's say you are reading something and there is a new word you don't know, how do you know that new word is a name?

T: Ah, that's a problem.

I: What's your strategy?

T: Strategy. When we see something, the new word?

I: Yeah, like you see a new word you don't know, umm, might be a name, you're not sure. What's your strategy?

T: Ah, look the words like 'the' or 'a'. We don't say 'a' before the name. So first that's a strategy. And check the dictionary. And then sometimes we can see the name on the dictionary but sometimes not.

I: What kind of names does your dictionary have?

T: Mike. Very typical names. Or Michael.

I: So it has very common names.

T: Yes, only common names.

I: And does it tell you the pronunciation? Does it say how to?

T: Ah, yes yes. But I can't often read the name, read the foreigner's name. I don't know to pronounce.

I: So what do you do if you can't pronounce, you just?

T: Hmm, just read as it is.

I: Because if you are reading like a graded reader, and the name appears many times, you just…

T: Yes, yes. There are often case that I can't understand, I can't pronounce and finish reading. I can't pronounce and finish reading. This is often case for I think not only but Japanese.

I: So do you ever ask someone?

T: Nooo.

I: No, just go through the book. You look for 'the' or 'a'. And then you might check your dictionary. Anything else?

T: Ah, capital word, capital letter. That's it.

I: Does your dictionary have place names, names of places?

T: Ah yes, very famous one.

I: So do you check those ones?

T: Hmm, hmm, sometimes.

I: Are you ever confused by place names when you see them?

T: Umm, yes. It's difficult for us to recognise name from place.

I: Hmm. Why?

T: Why. Because I think it's close. The meaning, the form, it looks like, the name and the place looks like close. Both start with capital letter.

I: Oh, like a name and a place name are the same, okay.

T: And in addition to that, as I said before, we don't check so usually I think we just keep reading. So that's why we can't recognise.

I: So in general, do you think it would be helpful then if you knew names, would it help the reading or this is something you just.

T: Ah, yes, helpful. Because in the story, so for example, in the test, in the story, in the novel, we have to remember or we have to imagine for ourselves what the character looks like. And that's of course connected to name. And under the name, we make character for our own. But if we don't know name, that's we often forget what this character doing or not.

I: Yeah. No I agree. When I'm reading, for example, a Russian novel, all the characters' names are Russian, it's hard for me to pronounce and I can't remember. Yeah, good. Can I ask you to read just one page for me from this book. It's called *My Family and Other Animals*. It's a Level 3 book so similar to what you would read. Yeah, and I'd like you to read aloud and every time you come to a name, I want you to say what does that name refer to.

T: Hmm. What does that name (.) refer to. Hmm.

I: Yeah, should I give you example? Yeah? Okay, so let's say I was reading this book. So. (Reading) The train rushed along angrily. Guy, okay this must be a name because

there's no must be a man. Guy was thinking about Mariam. He saw her, so Mariam must be girl, he saw her pink round face, her cruel mouth, he started to hate her. Perhaps Mariam doesn't want a divorce, Guy thought. Oh, so maybe Mariam is his wife because he is talking about divorce. Guy, is the character, thought unhappily but she is pregnant and it's not my child, okay? Do you see what I mean? Whenever you come to a name, tell me what does it refer to or who. Yeah? So just one page, this page til here, yeah?

T: Okay.

I: You can start there.

T: (Reading) Larry's idea. After July came the cold windows [*sic*: winds] and the grey skies of August. My family had all their usual illness [*sic:* illnesses]. My mother and I had bad colds. My brother Leslie had a problem with his ears. My sister Margo's spots were worse than ever. Only my oldest brother, Larry, was healthy, but he found the rest of us very difficult to live with.

I: Hmm. So any names so far? Any names so far? Which ones are the names?

T: Er (.) Leslie.

I: Who's Leslie?

T: Leslie is brother. Leslie. And Margo.

I: Uh huh. Who's that?

T: Sister. And Larry.

I: Uh huh. Who's Larry?

T: Larry is (.) oldest brother. And keep going?

I: Sure.

T: (Reading) Why do we stay in England in this weather? he asked Mother. Should I refer this 'he'?

I: No. Just the names.

T: He asked Mother. They're ill and you're looking older every day. I'm not, Mother replied. She was reading at the time. You are, Larry said. We need sunshine a country where we can grow. Yes dear, that's a good idea, Mother answered, not really listening. George says Corf's [*sic*: Corfu's] wonderful. Why don't we go there? If you like, dear. It was important to keep Larry happy.

I: Ok so any names so far?

T: Larry's name. Larry is (.) oldest brother.

I: Ok. Anything else?

T: (.) That's it.

I: Ok. Continue.

T: (Reading) When? asked Larry with surprise. Hmm. Larry. Mother realized her mistake and put down her book. Perhaps you can go first and look at the place, she said cleverly. If it's nice, we can all follow. Larry looked at her. You said that last time. You said that last time. I waited in Spain for two months and you didn't come. No if we're going to Greece, let's go together. But I've only just bought this house! Mother answered. Sell it again then! That's stupid, dear, said Mother. I can't do that. So we sold the house and ran from the English summer.

I: Good. Can I stop you there? Umm, any names? Place names or people names?

T: Spanish. Greece. Hmm. That's it.

I: Okay.

T: (Reading) So we sold the house. (.) We travelled by train with our clothes and our most important belongings: Mother's cook books, Leslie's guns, something for Margo's spots, Larry's books, my favourite insects and Roger, my dog. From Italy we caught a boat. We slept when the boat left and then, very early the next morning, we watched for Corfu. The sea turned blue, then purple, and suddenly there was the sleeping island in front of us. We sailed nearer and, above the noise of the ship, we could hear the high, clear sounds of the insects.

I: Okay. Names in that last part?

T: Corfu.

I: What's Corfu?

T: Corfu (.) name of (.) can I check? (.) Name of man (.) ah no Corfu is place.

I: Okay, any other names in that last part you read?

T: Other names (.) Italy, this ones ah people's name. Margo (.) ah sister (.) Larry, oldest brother. Roger, Roger (.) ah my dog.

I: Yeah, good. That's it. Yeah good. Okay so is there anything that I didn't ask you that I should have asked you about names. Is there anything else you can think of to tell me?

T: Tell me?

I: Yeah, whether you think names are helpful when you are reading, or are they kind of a problem.

T: Ah, hmm so (.) nothing.

I: Okay. I thought maybe there was a question I should have asked you but I didn't ask you. Yeah, no that's it. That's very helpful.

T: Oh really? [laughs]

I: Yeah. Because I just want to get students' opinions, you know, about names. And how they feel, their feelings.

T: Klassen, Kimberly.

I: Uh huh. What's my first name?

T: Hmm…

I: Or what's my family name?

T: I call you Kimi, Kimberly, Kimi. So Klassen is family name.

I: Yeah, Kimi is my nickname. It's short.

T: Yeah, Kimberly is full name. Kimberly Klassen.

I: Yeah so in English we say first name first. Kimberly Klassen. Japanese is opposite.

T: Do you know my name?

I: No because I know your nickname is T. Ah…

T: (Says his full name) Yeah, it's difficult.

I: (Repeats his name) Yeah that's difficult.

T: Which one is the family name?

I: Umm (says his family name).

T: Yeah, good.

I: (Says his first name) is your first name. T is your nickname.

T: So what's the purpose, what's the purpose or?

I: Of this?

T: Yes, of this research.

I: Oh of everything or this interview? Well this interview is just to find out students' opinions of names. But the purpose of my research is to find out if names are difficult for students and if they are, how?

T: If they are how difficult.

I: If they are a problem for students, then how are they a problem for students.

T: Ah, hai hai hai (yes, Japanese). And now you analyse.

I: Hmm. So I'm going to stop this. So thank you very much.

T: [laughs]

Interview #3 with Student K
December 9, 2014 4.40 – 5.00 (23 min. 15 sec.)
Interviewer's office at university

I: Ok, today I'm here with K, it's Dec. 9 and it's 4.40 after class. So thank you K for coming. As I told you, the purpose of my research is to look at how language learners, like you, deal with names when they are reading texts. So the questions I will ask you are about reading and names in texts, names of people, names of places. Ok? So I will keep the information confidential. Ah, so your first language is.

K: Japanese.

I: Japanese. And how long have you been studying English?

K: For six years.

I: For six years. Okay. Umm. So my first question is when you are reading something in Japanese, like Japanese newspaper or something, does it ever happen to you that you see a Kanji that you don't know?

K: Ah, sometimes.

I: Sometimes. And then what you do when that happens?

K: I guess the meaning from the context.

I: Okay, and anything else? Like you do try to find out how to pronounce.

K: Ah (.) when it happens, I often asked my mother or grandmother how to read. They will know about it.

I: Ok, so you'll ask someone else. So how about when you are reading English, does it ever happen that you are reading something and yeah?

K: Yeah, so often.

I: Yeah? Okay, so what do you do then?

K: When it happens, I will check it on the Internet and because spell is very complex and I really don't know how to even pronounce it. So I copy the spell, on Internet, it's easy for me. In Internet, I can know in Japanese because there are many facts on Internet dictionaries.

I: So what website do you use, like a dictionary?

K: Ah no, Google.

I: Google? So Google Japan, you will type in the name and then it will come up.

K: Like Wikipedia or Google map.

I: Ok great, yeah good. That's a great resource, right, the Internet, when you don't know something. That's great. How do you feel when you are reading a text or a graded reader for example, and there's lots of names you don't know? How does that make you feel?

K: Hmm. Uncomfortable. But when names often appear, commonly the book has the page to list the names.

I: Like the characters?

K: Hmm, the characters. At first or at last. So when I confused, really confused, I check it, return back to the page and check it, who it is.

I: Well that's helpful. What if the book doesn't have this chart?

K: Ah, I took memo.

I: Okay, do you make your own chart, like who is who?

K: Ah, when I really confused, I make memo and hmm (.)

I: Yeah, that's a good idea. Because if there are many characters in the book, it gets confusing right?

K: Now I'm getting used to remember who it is.

I: Oh yeah? Okay. So how do you remember who the characters are?

K: Character of course has characteristics, like she often says some kind of things, and others don't say such a thing so (.) Like that I distinguish the person. Like Sherlock Holmes, Holmes often says cold things, cool things. But Watson always confusing. [laughs] Like that.

I: Yeah, that's a good example. Do you think you know a lot of English names or a few?

K: Few, very few I think.

I: Enough or not enough?

K: Not enough.

I: Do you think it's important to know.

K: I think it's really important because when I can't pronounce it, I can't do anything. So it's important I think. And I don't know which is family name or first name, so for the essay or something, we usually use the last name or and it's complex for me to which name I should use. So I think it is important, it is essential for using English.

I: Hmm, yeah, because English names they are opposite of Japanese, right? In Japanese you have family name first right and then in English, yeah, so it gets confusing.

K: If it is only a part of name, I can't understand which it is.

I: Yeah because you don't know if the text is maybe referring to the family name or could be first name. How about men's names and women's names, are those confusing?

K: Ah, yes sometimes I misunderstand. First I thought that person is man but after I realize that it's a woman's name. [laughs] It is embarrassing for the person so I want to know about names.

I: That happens to me too in Japanese. I think Mizuki is a girl's name but now I have a male student same name.

K: Ah I know.

I: So when you are reading, you're reading something in English, how do you this new word is a name? For example, the reason I wanted to talk to you is, you know when we do the extensive reading in class and you make a list of vocabulary, a few times you have written down a word and then in brackets you wrote "maybe name?" So how do you know, can you say more about that?

K: Sometimes it appears big letter. First is big letter. Or I can't get the result from Internet. And sometimes it's meaning appeared in the Google but not fit in the context. So like three ways, I do.

I: Interesting. Yeah, 'cause I remember one book you read was *The Secret Garden* and there was a name "robin".

K: Ah, yes.

I: And you wrote it with a small 'r'.

K: So then I confused it is small 'r'. But later I found that was name of bird.

I: Oh it's the name of a bird? I think it's also the name of the boy isn't it?

K: Yes, like Robin Hood.

I: Yeah exactly.

K: I first confused but later, the name 'robin' appeared with the picture of a bird. And also then later, significantly, it's the words of bird like 'tchtchtch'. Robin says, tchtchtch. [laughs] So it must be bird.

I: Oh, because Robin can be a boy's name or a bird. So I guess if you googled that then you might get both.

K: On the Internet, I could only found the boy's name, Robin is boy's name. So I couldn't find the meaning of bird. So then I confused.

I: So when you do a search on Google and you are looking for a name, do you put the capital letter first? Or you just type it?

K: No, just type it, just type it, not capital.

I: But it still comes back with the name, yeah? Interesting okay. Uh, so that's why I wanted to ask you about names. At the beginning (when we were walking to my office), you said you find names difficult. Can you tell me more about that? Why do you think it's difficult, names?

K: Because the pronunciation of name is quite different from the spelling. I learned a lot of words from junior high school and there are many words and almost pronunciation is same as spelling. So I can pronounce the word I don't know with the spelling. But name is quite different from spelling. Name pronunciation is different so I can't guess how to pronounce it. Like ah like sorry I don't have an example.

I: No, I understand though, 'cause names can have much more possibility for spelling so could be anything so that makes it more difficult.

K: Ah, like Miss Zoë Jenkins. Her name Zoë is difficult for me. Because the z, o and e (gestures a dieresis on the 'e'). And also Jenkins is difficult for me, maybe for Japanese to pronounce. I only could know its pronounce 'Jenkins' after she said, I'm Zoë Jenkins.

I: Hmm, yeah, okay yeah. So what did she say those two dots are called, did she give it a name?

K: Sorry?

I: You know z, o, e, those two dots, did she give those two dots a name?

K: Hmm, she don't tell about.

I: Yeah because it's not common.

K: I knew it comes from Russia, Russia's letter.

I: Oh yeah?

K: Yeah.

I: I know they use it in German. But I don't think that's why she has it on her name. I don't think her name is German. [laughs] Ok, can you do one more thing for me? I wanted you to read just one page for me, aloud. This is from a book called *My Family and Other Animals.* It's Level 3 book so normal reader you would read, right? So what I would like you to do is read one page aloud for me just 'til here and every time you come to a name, either name of person or name of place, I want you to stop and tell me what does it refer to. I'll give you example, I'll give you example with this one. Ok (Reading) Introduction. Mme, Mrs., that must be title, Precious Momatsu, hmm, I don't know but maybe that's first name maybe that's family name, is a kind, large and warm-hearted African lady. African, adjective. She is also very unusual. She is the only lady private detective in Botswana. I know Botswana is country. And her agency, the No. 1 Ladies' Detective Agency, so that must be the name of her company, is the best. With the help of her secretary Mme Makoutsi, that is her secretary, and her good friend Mr. JLB Matakoni, that must be her friend, she solves a number of difficult and sometimes dangerous problems. Ok? So reading aloud, but every time you come to someone's name, or a name of a place, then stop and tell me who does it refer to.

K: Ahhh, okay.

I: These names are difficult for me too. [laughs] African names. Ok? So please start reading here.

K: (Reading) Larry's Idea. After July come [*sic*: came] the cold winds and the grey skies of August. My family had all their usual illnesses. My mother and I had had [*sic*:

bad] colds. My brother Lucille [*sic*: Leslie], uh, it's Larry's brother, had a problem with his ears. My sister Margo's, it's Larry's sister, spots were worse than ever. Only my oldest brother, huh? Larry, heh? my oldest (.) ah, I guess this 'my', my mother and I, and I is youngest brother, I guess maybe, was healthy, but he found the rest of us very difficult to live with. Why do we stay in England in this weather? he asked Mother. They're ill and you're looking older every day. I'm not, Mother replied. She was reading at the time. You are, Larry said. We need sunshine a country where we can grow. Yes dear, that's a good idea, Mother answered, not really listening. George (.) George is maybe lives in foreign country, says Corfu's wonderful. I don't know the word but it's maybe name of place. Do I need to say reason why? No? Why don't we go there? If you like, dear. It was important to keep Larry happy. When? asked Larry with surprise. Mother realized her mistake and put down her book. Perhaps you can go first and look at the place, she said cleverly. If it's nice, we can all follow. Larry looked at her. You said that last time. I waited in Spain, country name, for two months and you didn't come. No, if we're going to Greece, Greece, country name, let's go together. But I've only just brought [*sic*: bought] this house! Mother answered. Sell it again then! That's stupid, dear, said Mother. I can't do that. So we sold the house and ran from the English, heh? from the English summer. Heh, country? We travelled by train with our clothes and our most important belongings: Mother's cook books, Lucille's [*sic*: Leslie's], my, I's brother, guns, something for Margo's spots, Margo is (.) sister, Larry's books, eldest brother, my favourite insects and Roger, my dog, ah main character's dog. From Italy, country's name, we caught a boat. We slept when the boat left and then, very early the next morning, we watched for Corfu, ah place's name. The sea turned blue, then purple, and suddenly there was the sleeping island in front of us. We sailed nearer and, above the noise of the ship, we could hear the high, clear sounds of the insects.

I: Very good. So you said Corfu was name of a place. And then you said, Should I say why? So why did you think that, that Corfu was name of a place?

K: Ah, because I insists that I want to go other place. And George says Corfu is wonderful. So now the place is worse and Corfu is wonderful I guess.

I: Yeah, no, you're right. Corfu is a place. So when you come across names of places when reading, do you also go to Google sometimes to see?

K: Ah, yes. If (.) Corfu, I understand this time. But sometimes I can't understand so then I go Internet. But if I understand from the context, I don't do that because if I continue to read it maybe I can find out where is Corfu or what it's about.

I: Do you know from this page you just read, do you know where Corfu is?

K: Hmm I can't. But I can only understand Corfu is somewhere.

I: No that's great. That's perfect. Thanks very much. So was there anything else that you'd like to say, that you'd like to add? I have no more questions but maybe I should have asked you something that I didn't. Is there anything you want to tell me about reading and names?

K: Reading and names? Just a moment please (.) It's important for us Japanese to know the geographic because if the person will study about geographic, person will understand many countries name. Then it will be hint for reading. But I'm a little poor at geographic so sometimes it will be weak point.

I: So if you know place names, it can actually help you when you are reading. So names could be helpful, to help you understand.

K: So then I found that if I want to learn English I just have to study other subjects like geographic or (.) hmm or use it more cleverly. [laughs]

I: That's good. That's very interesting. Thank you very much. Anything else, if you think of anything else you want to tell me that has to do with names, you can send me email or you can tell me some other time.

K: Ok, sure.

I: 'Cause this was very helpful. So I'll stop.

K: Ok.

Interview #4 with Student W
December 11, 2014 4.25 – 4.45 (18 min. 53 sec.)
Interviewer's office at university

I: Ok I'm sitting in my office with W, and it's December 11 and it's 4.25, just after class. And we're sitting in my office. So thank you W for coming. As I told you, the purpose of my research is to look at how language learners, like you, look at names when they are reading something, like names of people or places. So your first language is.

W: Japanese.

I: And how long have you been studying English?

W: English (.) over 10 years.

I: Over 10 years? When did you (.) because most students say six, so you started earlier?

W: Yes, maybe five years old.

I: Oh yeah? OK. Yeah good. So about ten years, good. First of all, when you are reading something in Japanese, like a newspaper or something on the Internet, does it ever happen to you that you come across a Kanji, like for a name, that you don't know?

W: Hmm sometimes.

I: Ok. And what do you do when that happens?

W: I don't check the meaning. Just through (.) expect, guess.

I: Make a guess?

W: Yes.

I: And how about pronunciation, can you guess how to pronounce from the Kanji?

W: I know the pronounce, maybe. I can guess.

I: Ok, so you just keep reading, make a guess. Ok. So when you are reading something in English, do you ever see names that you don't know?

W: Yes, yes.

I: Yeah? Ok. What do you do when that happens?

W: Hmm (.) I can't read but try to read, read like *romanji* (Japanese words represented in Roman letters).

I: Uh huh. So ah, for example, one of the reasons I asked you for this interview is last semester, when we did a writing assignment and I was asking students about how they felt about reading. And in that essay, you wrote, you wrote something, you said when you are reading these kind of readers, these graded readers, you said it was difficult when there's many characters in the book.

W: Ah, yes.

I: Yeah? Can you tell me a little bit more about that? About when you are reading these books and there's lots of people.

W: Hmm I can't find out the name is woman or the man and (.) For me, the name is like similar so I forgot very fast. So always I checked before the reading, and who is this.

I: Oh, in the front of the book, you mean? It has.

W: No, the first time the person is (.)

I: Oh, oh. So you go back and check. Okay. So you said the names sometimes look similar. How do they look similar?

W: How, hmm (.) how?

I: Like the spelling is similar?

W: Yes, yes. I decide (.) so the first letter, like the first letter is S, I mistake to read, and so (.) hmm.

I: Ok, so let's say the first letter is S. So you'll just kind of focus on the S, don't worry about the rest of the letters.

W: Yes.

I: Does it bother you, like let's say the name starts with S, hmm, do you think it would easier to remember if you could pronounce it, like if you could say the name?

W: No.

I: No, okay. Just want to get the reading going, yeah okay. So how would you describe your level of knowledge of English names, would you say you know a lot of names, or enough names, or a few names?

W: One more please.

I: Um, how many names you do know in English, do you think you know a lot of names?

W: Ah no, not enough.

I: Not enough, okay. So you wish you knew more?

W: Yes.

I: Do you think names are helpful when you are reading some difficult texts?

W: Yes.

I: Like that text we did (.) yesterday, the scientific text. There were a few names in there. Did you think those names were helpful or?

W: Yes.

I: Yeah? Oh good. And you mentioned something about women's names and men's names. Ah how do you, do you try to guess which ones are?

W: Ah, yes, guess.

I: Yeah? How do you know which ones are for women?

W: Hmm, just a guess.

I: Yeah. Because that happens to me too in Japanese. I had a student, her name was Mizuki, and this year I have a student, *his* name is Mizuki. So I was confused. Does that ever happen…you think it's going to be a girl, and then it turns out to be a boy…yeah? Ok, ah, let's say you are reading a text, like in the textbook or this kind of book, and you see a new word, that you've never seen before, how do you know if it's a name or a regular word?

W: How do you know?

I: Yeah, how do you know if it's a name or if it's just a word?

W: The capital letter. Or the (.) structure.

I: Of the sentence, like the context. And then, so if you see a name that you've never seen before, do you have any strategy, for example do you check your dictionary?

W: Ah, name? I don't check.

I: Does your dictionary have names?

W: No.

I: How about online, do you ever check?

W: No. But the person is famous, I check.

I: Oh yeah, to find out more information, yeah. Yeah good, ok could you do one thing for me. I'd like you to read one page for me, just this and this, this is from a book called *My Family and Other Animals*. And it's Level 3 so I think a normal level, right? And every time you come to a name, every time you see a name of a person or a name of place I want you to stop and tell me, who does that refer to. Ok? I'll give you example. I'll use this one because this one has difficult names for me. This book is set in Africa so the names are kind of difficult. So so I'll give you example. (Reading) Mme, I don't know this one, Mama, Mrs., must be title, Precious Romatsu, so this must be first name and this is family name, is a kind, large and warm hearted African lady. Ok, so this must be a woman. She is also very unusual. She is the only private lady detective in Botswana. I know Botswana is a country in Africa. And her agency, The No. 1 Ladies' Detective Agency, so this must be name of her company, is the best. With the help of her secretary, I guess that's Mrs. Makoutsi, must be family name, and her good friend Mr. JLB Matakoni, that must be family name and it's a man, she solves a number of difficult and sometimes dangerous problems. Ok? Do you understand? So every time you see a name, person or place, just stop and tell me who it refers to. Ok, starting there.

W: (Reading) Larry's Idea. Larry's maybe a woman. After July came the cold winds and the grey skies of August. My family had all their usual illnesses. My mother and I had

bad colds. My brother Leslie, Leslie is a man, had a problem with his ears. My sister Margo's, Margo's is the girl, spots were worse than ever. Only my oldest brother, Larry, Larry is a man and the brother, was healthy, but he found the rest of us very difficult to live with. Why do we stay in England, England is a country, in this weather?' he asked Mother. They're ill and you're looking older every day. I'm not, Mother replied. She was reading at the time. You are, Larry said. Larry is a man. We need sunshine, a country where we can grow. Yes dear, that's a good idea, Mother answered, not really listening. George says, George will maybe, maybe a man, Corfu's wonderful. Corfu is place? Why don't we go there? If you like, dear. It was important to keep Larry happy. Larry is the man. When? asked Larry with surprise. Larry is a man. Mother realized her mistake and put down her book. Perhaps you can go first and look at the place, she said cleverly. If it's nice, we can all follow. Larry looked at her. Larry is the man and her is maybe mother. You said that last time. You, you, you means (.) the mother. I waited in Spain, Spain is the country, for two months and you didn't come. No, if we're going to Greece, Greece is the country, let's go together. But I've only just bought this house! Mother answered. Sell it again then! That's stupid, dear, said Mother. I can't do that. So we sold the house and ran from the English summer. English is a country. We travelled by train with our clothes and our most important belongings: Mother's cook books, Larry's [*sic*: Leslie] guns. Larry means people, something for Margo's spots, Margo maybe person, Larry's books, Larry the man, my favourite insects and Roger, Roger is the dog, my dog. From Italy we caught a boat. Italy is a European country. We slept when the boat left and then, very early the next morning, we watched for Corfu. Corfu is the city. The sea turned blue, then purple, and suddenly there was the sleeping island in front of us. We sailed nearer and, above the noise of the ship, we could hear the high, clear sounds of the insects.

I: Good thank you, you're a very good reader. So, this is a good example of story with lots of people in it. I think there's Larry, Leslie, Margo, the mother, Roger the dog, so there's at least five characters all on just one page. So if you were reading this, would you do any strategies to try to remember who is who? Or would you just try to enjoy the story?

W: I just read but many times the same name are coming so the second time, I back to first sentence. But the letter I remembered.

I: Ok, for example, this story has two people with the letter L, Larry and Leslie, so how could you distinguish between those two because almost same length, Larry and Leslie. How could you distinguish between those two?

W: Hmm (.) the atmosphere.

I: Ok, from the character?

W: Yes.

I: OK, from what you read, what do you know about Larry?

W: He is the man.

I: Yes. What do you know about Leslie?

W: Leslie is the sister.

I: Yeah?

W: Hmm (.) Oh, Leslie is the brother.

I: Yes, he is also brother. So they are both men and they both start with L. So what strategy would you use then?

W: [Laughs]

I: I guess you have no problem pronouncing these ones, so that helps right? Ok. How about place names, are those ever difficult for you? This one had several place names, England, Spain, Italy, Corfu. Can you guess where is Corfu, from what you read?

W: Hmm. No.

I: No, ok. Do you ever check the meaning of place names? No? Ok good. Thank you, you're a very good reader. That's good. Just one more question, is there anything else you can tell me about names that you want to say?

W: I can't decide the first name or the last name.

I: Hmm, that's difficult. Because it's opposite of Japanese right?

W: Yeah. In Japanese, I can find this is the last name or the first name. But in English, I can't imagine that so, for example, in citation, I can't which is one should put down the statement.

I: Hmm, yeah, that's confusing. Because usually we write it first name then family name.

W: But sometimes capital letter or something changed.

I: Yeah so for citation you need to write last name, first.

W: Yeah, but I can't which is first name.

I: So in Japanese it is easy for you because you are familiar. But in English you aren't familiar so (.) So do you think it would be helpful to know more names? Do you think it would help your reading comprehension, like if you're reading this kind of story?

W: Yes.

I: Yeah, it's confusing because Japanese is opposite. This was very helpful. Thank you very much. So I'm going to stop the recording.

Appendix 2: How L2 readers approach proper names

2.1 Three texts

Easy Text: A History of Cars

The car was not invented by one inventor. Many people from all over the world had a hand in its development. Leonardo da Vinci was one these people. He had an idea for something like a car in the 15th century. In 1769, an engineer in France named Nicolas Cugnot invented the first carriage with wheels that could run by itself. Powered by a steam engine, it was large and even slower than horses. The French army used it to move heavy things. Many people call Cugnot the inventor of the first car.

Because steam worked well for moving boats and trains, people kept trying to use it for carriages. In 1801, Richard Trevithick in England built a steam carriage for use on roads. Steam carriages, however, were just not as efficient as steam-powered trains and boats.

While some inventors kept working with steam power, others were trying to use electricity to power carriages. In 1873, an inventor in the United States developed an electric car that got power from metal tracks in the streets. Around the same time, an inventor in France named Gaston Plante tried making a different kind of car. Like steam, though, electricity was not efficient.

Three years after electric cars were built in the United States, a German named Nicolaus Otto invented a gas motor engine. Ten years later, Karl Benz built the first car powered by a gas engine. It was a three-wheeled car. The following year, Gottlieb Daimler and Wilhelm Maybach made the first four-wheeled car by putting a gas engine on a horse carriage. Because their inventions used gas engines, as cars do today, many people say that either Benz or Daimler and Maybach invented the first car.

By 1900, gas-powered cars started to become popular. Companies in France were the first to make and sell complete cars. Companies in the United States and Germany soon followed. In 1901, an American car called the Oldsmobile was the first to be made in large numbers. In 1913, a car builder in the United States named Henry Ford developed a way to make many cars very quickly. Workers stood in a line, and each worker made the same part of the car again and again. This way, workers could make a whole car in only 93 minutes. By 1927, the Ford Company was the world's largest car builder.

These days, inventors are once again trying to find ways to make electric cars efficient. This is because of concerns about the effect of gas-powered cars on the environment and global warming. If an affordable, efficient electric car could be made, then no one would need to use gas-powered cars. Tesla Motors is one company working on an efficient electric car. The company become famous after the creation of the first fully electric sports car. In addition to electric cars, the company also makes stations where people can charge their electric Tesla car. They have charging stations in North America, Asia and Europe.

The world's best selling electric car is the Nissan Leaf. This all-electric car was introduced in Japan and the United States in 2010. It won several awards. By 2015, it

was the best selling electric car around the world for use on highways. These cars are very quiet, so they also reduce noise pollution.

Words: 551
Types: 231
Proper name types: 34
Proper names tokens: 50 (9.1%)
Lexical profile 98.2% at K3 (proper names ignored)

Moderate Text: Typhoid Mary

Mary Mallon, now known as Typhoid Mary, seemed a healthy woman when a health inspector knocked on her door in 1907, yet she was the cause of several typhoid epidemics. In fact, 47 illnesses and three deaths were attributed to her. She was the first "healthy carrier" of typhoid fever in the United States. She was forced to live in relative seclusion upon North Brother Island off New York. Who was Mallon, and how did she spread typhoid fever?

For the summer of 1906, banker Charles Warren wanted to take his family on vacation. They rented a summer home from George Thompson and his wife in Oyster Bay, Long Island. Also for the summer, the Warrens hired Mallon to be their cook.

On August 27, one of Warren's daughters became ill with typhoid fever. Soon, the wife and two maids became ill, followed by the gardener and another daughter. In total, six of the eleven people in the house came down with typhoid. Since the common way typhoid was spread was through water or food sources, the owners of the home feared they would not be able to rent the property again without first discovering the source of the epidemic. The Thompsons first hired investigators to find the cause, but they were unsuccessful.

Then the Thompsons hired George Soper, a civil engineer with experience in typhoid fever epidemics. It was Soper who believed the recently hired cook, Mallon, was the cause. She had left the Warren's approximately three weeks after the epidemic. Soper began to research her employment history for more clues.

Mallon was born on September 23, 1869 in Cookstown, Ireland. According to what she told friends, she emigrated to the United States around the age of 15. Like most Irish immigrant women, she found a job as a domestic servant. Finding she had a talent for cooking, she became a cook, which paid better wages than many other domestic service positions.

Soper was able to trace her employment history back to 1900. He found that typhoid epidemics had followed her from job to job. From 1900 to 1907, he found that she had worked at seven jobs in which 22 people had become ill, including one young girl who died with typhoid fever shortly after she had come to work for them. He was satisfied that this was much more than coincidence; yet, he needed to scientifically prove she was the carrier.

In March 1907, he found her working as a cook in the home of Walter Bowen and his family. Soper handed his research and hypothesis over to the health department. She was taken to the Willard Parker Hospital, where samples were taken and examined; typhoid germs were found. The health department transferred her to an isolated cottage, part of the Riverside Hospital (in the East River near the Bronx).

She was taken by force and against her will and was held without trial. She had not broken any laws. She believed she was being unfairly persecuted. Wasn't she healthy? She couldn't understand how she could spread disease and caused death when she, herself, seemed healthy. She might have had such a weak case of typhoid fever that she experienced only flu-like symptoms. In 1909, after having been isolated for two years, she sued the health department.

Words: 550
Types: 259
Proper name types: 34
Proper name tokens 51 (9.3%)
Lexical Profile: 95.2% at K3 (proper names ignored)

Difficult Text: An Invisible Enemy

An ancient and deadly enemy for almost 10,000 years, smallpox killed millions of people all over the world. During epidemics, from 50 to 60% of the population would contract the disease, and 20 to 30% of its victims would die.

In the 16th century, Spanish conquerors brought smallpox to the Americas, decimating most of the native populations because they had never been exposed to the disease before and had no immunity. More Native Americans died of smallpox than died in battle with white settlers. In the 18th century, the British deliberately infected Native Americans with smallpox during the French and Indian War. In London, 80% of the children under five years old who caught the disease died; in Berlin, 98%. One-third of those who survived smallpox went blind. Spread by contact through the air, the disease ravaged every class of society. It caused a rash and blisters on the skin that left its survivors scarred for life. Queen Elizabeth I, Mozart, and George Washington all knew the suffering it brought.

Over the centuries, human beings gained knowledge of the "speckled monster." For one thing, they learned that survivors never caught the disease again. From this observation of natural immunity came the practice of variolation. Variolation probably began in China and India. In the early 18th century, Lady Mary Wortley Montagu observed its use in Turkey, where her husband was ambassador. Lady Mary brought the practice to England in 1721. Through a small cut, the pus from a smallpox blister is placed under the skin. If the disease does not develop any further, the person will get well and be immune ever after. It was a dangerous procedure because the virus was not always weak enough and those who were variolated could start epidemics.

Edward Jenner, an English country doctor, heard stories about farm workers who had contracted cowpox (a mild infection of animals) and then never got smallpox. Jenner became known not because he discovered the cowpox/smallpox connection – many farmers knew this from observation – but because he investigated it scientifically and proved it, showing the world how to benefit from it. Jenner set up experiments with various children. Jenner inoculated them with cowpox and then proved that they no longer reacted when variolated with smallpox. The children had acquired cross-immunity. Jenner called his prevention method a "vaccine" from the Latin for "cow". He sent his findings to the British Medical Association in 1797, but they refused to publish Jenner's revolutionary views without more proof. When other doctors and scientists repeated his experiments and found that he was right, Jenner became a hero. In France, Jenner was so revered that on his request the French emperor Napoleon Bonaparte freed some British prisoners of war, saying: "Ah, Jenner. I can refuse him nothing."

In 1853 universal vaccination became compulsory in England and Wales. More than a century later, however, 15 million people in the world were still getting smallpox every year and 2 million were dying from it. In 1967, the United Nations' World Health Organization began a global campaign to eradicate smallpox with free vaccines. The last known natural case of smallpox was in Somalia in 1977. In 1980, thanks to Edward Jenner and to all the microbe hunters fighting humanity's invisible enemies, the World Health Assembly declared the world free from smallpox.

Words: 552
Types: 284
Proper name types: 34
Proper name tokens: 51 (9.2%)
Lexical profile 89.6% at 3K (proper names ignored)

2.2 Consent form

Consent Form for Participating in Research

I (Kimberly Klassen) would like to conduct research on reading skills of Japanese learners of English. If you agree to participate in this research, your performance will not affect your grades for this course in anyway. Also, your decision to participate in this research or not will not affect your grade in this course.

Your name will NOT appear in published results of this research. Likewise, your class number will NOT appear in published results. In sum, there will be no way to identify you personally as a participant.

Taking part in the project will involve reading a text and completing a short task afterwards, and this will take approximately 15 minutes. You can withdraw from this study afterwards without giving a reason. You are free to ask questions at any time. The data collected from this research will be held confidentially.

学習者の英語のリーディング・スキルに関する調査のご協力をお願いします。この調査での内容が授業の成績に影響することは一切ありません。また、この調査に参加するか否かが成績に影響することもありません。

協力して頂いた方の名前や学籍番号は研究結果の中に記載致しませんので、あなた個人がこの調査の参加者として特定されることはありません。

この調査では、テキストを読んで短いタスクに答えて頂きます。時間はおよそ15分程度です。理由なしにこの研究への参加をやめることができます。質問があればいつでもお尋ねください。研究から得られたデータは内密に扱われます。

If you agree to take part in this research, please sign the consent below:

Student Name: _____ Student No. _____

Date: _____ Signature: _____

2.3 Debriefing form

I want to thank you for participating in my research about the effects of vocabulary on reading comprehension of Japanese learners of English. In this research, I'm looking at how names of people and places can affect students' understanding of a reading text. You have helped me with this by reading the text and completing the task.

Please be assured that the data I collected from you will be held confidentially, and your name and class number will not appear in any published article. You can ask me questions about this research at any time. You can still withdraw from this study if you wish, without giving a reason. If you want to withdraw, I will delete your responses to the task from my data set.

If you have any questions, you can contact me at kklassen@kansaigaidai.ac.jp. You can also contact my supervisor at Cardiff University, Tess Fitzpatrick at FitzpatrickT@cardiff.ac.uk.

2.4 Practice passage

Practice

Task 1: Read this text slowly and carefully. As you read, circle any words you do not know (any words you have not seen before or do not know the meaning of). Note any words you do not circle may appear on a quiz next week.

タスク１：ゆっくり、そして注意深くこのテキストに目を通してください。読み進めている際、分からない語彙があれば、それを丸で囲んでください。目にしたことのないもの、意味の分からないもの）丸がつけられなかった語彙は来週のボキャブラリークイズに出題される可能性があります。

What is it that tells us this animal is a “dog” and that one is a “wolf”? Modern wolves and dogs can be easily identified by their appearance. The most important difference is in the snout. In almost all dogs, the snout is shorter and wider than wolf snouts. Another crucial difference is the animal’s manner and attitude toward humans. Dogs are genetically predisposed to want human attention and approval and to accept human leadership. Wolves are not.

Because early dogs looked more like wolves than dogs do today, it can be difficult to distinguish between wolf and dog skeletons from the far past. But recently, a team led by palaeontologist Mietje Germonpre of the Royal Belgian Institute of Natural Sciences reported a stunning new finding in the February 2009 issue of Journal of Archaeological Science.

****

Task 2: Choosing from the words you circled, put the words in order that you would look them up in a dictionary or on the Internet. Put the word that you would look up first in space 1, the word you would look up next in space 2, and so on.

タスク２：丸で囲んだ語彙のなかで、あなたが辞書やインターネットを使ってその意味を調べる際に、どのような順番で調べていくのかを記してください。あなたが初めに調べるであろう語彙をスペース１に記してください。２番目に調べるであろう語彙はスペース２に記してください。このように３番目、４番目と続けて記してください。

1. _____

2. _____

3. _____

Appendix 3: Cultural familiarity with proper names (replication study)

3.1 Examples of adjustment in the text

Original short story	Adjusted short story
"First let's go see a football game. A professional football game," Frances said, because she knew Michael loved to watch them. "The Giants are playing. And it'll be nice to be outside all day today and get hungry and later we'll go down to Cavanagh's and get a steak as big as a blacksmith's apron, with a bottle of wine, and after that, there's a new French picture at the Filmarte that everybody says... Say, are you listening to me?"	"First let's go see a baseball game. A professional baseball game," Misaki said, because she knew Takuya loved to watch them. "The Hanshin Tigers are playing. And it'll be nice to be outside all day today and get hungry and during the game we'll go down to Koshien's Cafeteria and get a big Koshien Curry with a beer, and after that, there's a new Yoshimoto Shingigeki comedy play at Namba Kagetsu that everybody says... Say, are you listening to me?"
They joined hands consciously and walked without talking among the baby carriages and the old Italian men in their Sunday clothes and the young women with Scotties in Washington Square Park.	They joined hands consciously and walked without talking among the baby carriages and the old men holding their red pens and sports newspapers heading to Wins and the young women with Shiba-inu dogs on Namba Parks.
"I like the girls in the offices. Neat, with their eyeglasses, smart, chipper, knowing what everything is about, taking care of themselves all the time." He kept his eye on the people going slowly past outside the window. "I like the girls on Forty-fourth Street at lunchtime, the actresses, all dressed up on nothing a week, talking to the good-looking boys, wearing themselves out being young and vivacious outside Sardi's, waiting for producers to look at them. I like the salesgirls in Macy's, paying attention to you first because you're a man, leaving lady customers waiting, flirting with you over socks and books and phonograph needles. I got all this stuff accumulated in me because I've been thinking about it for ten years and now you've asked for it and here it is."	"I like the girls in the offices. Neat, with their eyeglasses, smart, chipper, knowing what everything is about, taking care of themselves all the time." He kept his eye on the people going slowly past outside the window. "I like the girls on Amerika mura at lunchtime, the actresses, all dressed up on nothing a week, talking to the good-looking boys, wearing themselves out being young and vivacious outside on Hikakebashi, waiting for producers to look at them. I like the salesgirls in Takashimaya, paying attention to you first because you're a man, leaving lady customers waiting, flirting with you over socks and books and jewelry. I got all this stuff accumulated in me because I've been thinking about it for ten years and now you've asked for it and here it is."

3.2 Posttest questions for the original story

Recall Test

Instructions: You have 15 minutes to complete the test. You may not look back at the reading. You may not use a dictionary.

True/False/Not Given

Write (T) if the given statement is true; write (F) if it is false; and write (NG) if it is not mentioned in the story.

1. _____ Michael and Frances have known each other for two years.

2. _____ Michael never makes love with other women.

3. _____ The Stevensons know what Michael feels for other women.

4. _____ Frances does not want to see anybody all day because she wants to have a rest.

5. _____ Michael looks at other women only in the streets.

6. _____ Frances feels good all day when she has breakfast with Michael.

7. _____ The waiter was very kind to them in the bar.

8. _____ They didn't go to the cinema to see a French picture.

9. _____ Michael asked Frances to have a drink when they were on the way to Cavanagh's.

10. _____ The Stevensons will come to the bar to pick them up.

Put the Following Events into Order

Below are eight statements from the short story you have just read. Put them into the correct order of happening. Write the number in the brackets.

(_____) They walked to a bar on Eighth Street.

(_____) Frances began to cry, silently, into her handkerchief.

(_____) Frances and Michael had slept late and had a good breakfast.

(_____) They decided to call the Stevensons.

(_____) They started to walk from the Brevoort toward Washington Square.

(_____) Frances got up from the table and walked across the room.

(_____) Frances planned a day of activities that Michael will enjoy.

(_____) Frances asked Michael to stop talking about women and to keep it to himself.

Short Answer Questions

1. Why does Frances want to take Michael to a football match?

2. What does Michael think of when he thinks of the city of New York?

3. Where did Frances & Michael meet for the first time? Describe Michael's feelings at that time.

4. What's the first thing Michael noticed when he first came to New York from Ohio?

5. Why does Frances feel good on that Sunday morning?

6. How has Michael physically changed since he moved from Ohio?

7. What are they going to do together with the Stevensons?

8. Why do the salesgirls in Macy's pay attention to Michael?

9. What is the favour that Frances asks Michael to do for her?

10. What does Michael feel about Frances when she gets up from the table?

Appendix 4: Three treatments of proper names

4.1 Three versions of two texts

Text 1: Internet Fraud* (version A, English proper names)

All the places for new students had been taken at a famous primary school in London, England. The school had a waiting list. Any places that opened up would be offered to children at the top of the list. One child (let's call her Wendy) was near, but not at, the top. Next to her on the list, one step higher, was another girl (let's call her Jane). Wendy's mother set up an email account in the name of Jane's mother and sent the school an email asking them to remove Jane from the waiting list. As a result, Wendy rose one step. The fraud* was discovered when Jane's real mother called the school to ask about the list. By the way, Wendy and Jane were only four years old.

Wendy's mother took part in a small-scale act of Internet fraud. This case of false identity was not very serious – except perhaps to Jane's mother. The school had no useful way of checking identification, probably because it did not prepare for cheating by desperate parents. It was especially unprotected for fraud on the Internet, since no face-to-face contact occurred. If Wendy's mother had actually had to go to the school to remove Jane, someone might have recognised her. Her plan almost worked because she successfully established a false identity, even if only for a short time. If she had been required to show an identification (ID) card, the fraud could probably not have been committed at all.

fraud: when someone lies to people by pretending to be someone they are not*

(version B, no proper names)

All the places for new students had been taken at a famous primary school. The school had a waiting list. Any places that opened up would be offered to children at the top of the list. One girl was near, but not at, the top. Next to her on the list, one step higher, was a boy. The girl's mother set up an email account in the name of boy's mother and sent the school an email asking them to remove the boy from the waiting list. As a result, the girl rose one step. The fraud* was discovered when the boy's real mother called the school to ask about the list. By the way, the girl and the boy were only four years old.

The girl's mother took part in a small-scale act of Internet fraud. This case of false identity was not very serious – except perhaps to the boy's mother. The school had no useful way of checking identification, probably because it did not prepare for cheating by desperate parents. It was especially unprotected for fraud on the Internet, since no face-to-face contact occurred. If the girl's mother had actually had to go to the school to remove the boy, someone might have recognised her. Her plan almost worked because she successfully established a false identity, even if only for a short time. If she had been required to show an identification (ID) card, the fraud could probably not have been committed at all.

(version C, Japanese proper names)

All the places for new students had been taken at a famous primary school in Tokyo, Japan. The school had a waiting list. Any places that opened up would be offered to children at the top of the list. One child (let's call her Haruna) was near, but not at, the top. Next to her on the list, one step higher, was another girl (let's call her Misaki). Haruna's mother set up an email account in the name of Misaki's mother and sent the school an email asking them to remove Misaki from the waiting list. As a result, Haruna rose one step. The fraud* was discovered when Misaki's real mother called the school to ask about the list. By the way, Haruna and Misaki were only four years old.

Haruna's mother took part in a small-scale act of Internet fraud. This case of false identity was not very serious – except perhaps to Misaki's mother. The school had no useful way of checking identification, probably because it did not prepare for cheating by desperate parents. It was especially unprotected for fraud on the Internet, since no face-to-face contact occurred. If Haruna's mother had actually had to go to the school to remove Misaki, someone might have recognised her. Her plan almost worked because she successfully established a false identity, even if only for a short time. If she had been required to show an identification (ID) card, the fraud could probably not have been committed at all.

Text 2: The Weather Goes to Court (version A, English proper names)

The witness, Mrs. Smith, said that she had heard the defendant[1], Mr. Brown, admit to stealing a car. She was sitting on a park bench, Mrs. Smith said, when Mr. Brown, speaking loudly and pointing forcefully toward the parking lot, told another man he had just "stolen that silver Toyota." Mrs. Smith said she could easily hear it because the Mr. Brown was standing only about 50 metres northeast of her. Mrs. Smith knew it was Mr. Brown because he was on a small hill where she could easily see him. The judge, Mr. Johnson, thanked her and she sat down.

Now it was Mr. Williams, the defense lawyer's turn. Mr. Williams' plan was to make Mr. Johnson doubt what he had just heard. Mr. Williams called a new witness, Mr. Miller, a meteorologist[2]. People throughout the courtroom[3] wondered: Why call a weather expert[4]?

Mr. Miller stated that it had been sunny with clear skies on the day in question. Weather records said so. Could Mrs. Smith have seen Mr. Brown talking? The weather would not have been a problem, Mr. Miller said. Could Mrs. Smith have heard what Mr. Brown said? "Well, the way Mrs. Smith described it, probably not. The wind was a bit strong that day, out of the southwest at about 25 kilometres per hour. Mr. Brown was northeast of Mrs. Smith and standing on a hill. Sound waves heading into wind get pushed up. By the time they had traveled 50 metres, they would have been too high to reach Mrs. Smith's ears."

[1]defendant**:** *the person in court who is accused of a crime*

[2]meteorologist**:** *someone who studies the weather, especially to say how it will be in the future*

[3]courtroom**:** *the room where a judge decides whether someone is guilty of a crime*

[4]expert**:** *someone who has a lot of skill and knowledge about something*

(version B, no proper names)

The witness said that she had heard the defendant[1] admit to stealing a car. She was sitting on a park bench, she said, when the defendant, speaking loudly and pointing forcefully toward the parking lot, told another man he had just "stolen that silver car." She said she could easily hear it because the defendant was standing only about 50 metres northeast of her. She knew it was him because he was on a small hill where she could easily see him. The judge thanked her and she sat down.

Now it was the defense lawyer's turn. His plan was to make the judge doubt what he had just heard. He called a new witness, a meteorologist[2]. People throughout the courtroom[3] wondered: Why call a weather expert[4]?

The expert stated that it had been sunny with clear skies on the day in question. Weather records said so. Could the previous witness have seen the defendant talking? The weather would not have been a problem, the expert said. Could she have heard what he said? "Well, the way she described it, probably not. The wind was a bit strong that day, out of the southwest at about 25 kilometres per hour. He was northeast of her and standing on a hill. Sound waves heading into wind get pushed up. By the time they had traveled 50 metres, they would have been too high to reach her ears."

(version C, Japanese proper names)

The witness, Mrs. Nakamura, said that she had heard the defendant[1], Mr. Tanaka, admit to stealing a car. She was sitting on a park bench, Mrs. Nakamura said, when Mr. Tanaka, speaking loudly and pointing forcefully toward the parking lot, told another man he had just "stolen that silver Toyota." Mrs. Nakamura said she could easily hear it because Mr. Tanaka was standing only about 50 metres northeast of her. Mrs. Nakamura knew it was Mr. Tanaka because he was on a small hill where she could easily see him. The judge, Mr. Yamada, thanked her and she sat down.

Now it was Mr. Okada, the defense lawyer's turn. Mr. Okada's plan was to make Mr. Yamada doubt what he had just heard. Mr. Okada called a new witness, Mr. Yoshida, a meteorologist[2]. People throughout the courtroom[3] wondered: Why call a weather expert[4]?

Mr. Yoshida stated that it had been sunny with clear skies on the day in question. Weather records said so. Could Mrs. Nakamura have seen Mr. Tanaka talking? The weather would not have been a problem, Mr. Yoshida said. Could Mrs. Nakamura have heard what Mr. Tanaka said? "Well, the way Mrs. Nakamura described it, probably not. The wind was a bit strong that day, out of the southwest at about 25 kilometres per hour. Mr. Tanaka was northeast of Mrs. Nakamura and standing on a hill. Sound waves heading into wind get pushed up. By the time they had travelled 50 metres, they would have been too high to reach Mrs. Nakamura's ears."

4.2 Posttest questions for version A of two texts

Internet Fraud: Comprehension questions (version A)

True, false or not given? Write T for true, F for false and NG for statements that cannot be inferred from the text.

_____1. Jane's mother pretended to be someone else.
_____2. Wendy was lower on the list than Jane.
_____3. Parents have to go to the school in person to remove their child's name from the waiting list.
_____4. The school in London was not very careful in determining parents' identities.
_____5. In the end, Jane was not able to attend the school.

Multiple Choice Circle the letter of the best answer for each question.

6. Why were Wendy and Jane on the waiting list at the school?
    a. They were only four years old.
    b. They were not famous.
    c. The school was not famous.
    d. The school was full.

7. Wendy's mother sent the school an email about:
    a. taking Jane off the list.
    b. taking Wendy off the list.
    c. putting Jane on the list.
    d. putting Wendy on the list.

8. When was the fraud discovered at the school?
    a. When Wendy's mother emailed.
    b. When Wendy's mother called.
    c. When Jane's mother emailed.
    d. When Jane's mother called.

9. Why did Jane's mother call the school?
    a. to ask about the list.
    b. to ask about the fraud.
    c. to take Jane off the list.
    d. to take Wendy off the list.

10. What kind of fraud did the mother take part in?
    a. face-to-face fraud
    b. letter fraud
    c. telephone fraud
    d. Internet fraud

The Weather Goes to Court: Comprehension questions (version A)

True, false or not given? Write T for true, F for false and NG for statements that cannot be inferred from the text.

_____1. Mrs. Smith saw Mr. Brown steal the car.
_____2. Mr. Miller is a meteorologist.
_____3. It was not possible that Mrs. Smith heard Mr. Brown speaking about stealing the car because of the wind.
_____4. Mr. Miller said Mrs. Smith could not have seen Mr. Brown clearly on that day.
_____5. Mr. Johnson decided that Mr. Brown had stolen the car, based on what Mr. Miller said.

Multiple Choice Circle the letter of the best answer for each question.

6. An expert in the weather is called a:
    a. defendant.
    b. judge.
    c. witness.
    d. meteorologist.

7. In the stolen car case, Mrs. Smith saw Mr. Brown speaking to another man:
    a. in a car.
    b. on a small hill.
    c. in a parking lot.
    d. on a farm.

8. Mr. Williams, the defense lawyer, works for:
    a. Mrs. Smith, the witness.
    b. Mr. Johnson, the judge.
    c. Mr. Brown, the defendant.
    d. Mr. Miller, the meteorologist.

9. The weather on the day Mr. Brown may have stolen the car had been:
    a. sunny and calm.
    b. windy and rainy.
    c. rainy and calm.
    d. sunny and windy.

10. Who did Mr. Williams, the defense lawyer, not want to believe what the first witness said?
    a. Mr. Johnson, the judge.
    b. Mrs. Smith, the witness.
    c. Mr. Brown, the defendant.
    d. Mr. Miller, the meteorologist.

4.3 Marking rubric for free recall summaries

Text 1: Internet

Five main points from the text, worth 2 points each for a total of 10 points. Marker may award 1 point for partial ideas.

1. Wendy/Haruna/the girl's mother pretended to be another student's mother.
2. Wendy/Haruna/the girl's mother sent an email to the school to remove the other student's name from a waiting list.
3. The fraud was discovered when the real mother called the school to ask about the list.
4. The fraud happened because parents do not have to go to the school in person to remove a name.
5. Internet fraud can happen because there is no face-to-face interaction OR no ID is required.

Clarifications made on rubric after marking included:

Point 1: Student must have 'mother' and some form of 'pretending' (including 'committed fraud' or 'lied') for two points.

Point 2: Student must have 'email' and some reference to 'removing name from list' (including 'change position on list') for two points.

Point 4: Student must have some reference to 'school' and 'not appearing in person or with ID' for two points. Conditional sentences are acceptable; for example, *If the parents had to go to the school, they would have recognised them*.

Point 5: Must have some reference to 'internet' or 'internet fraud' and either 'no face-to-face' or 'no ID' for two points.

Text 2: Meteorology

Five main points from the text, worth 2 points each for a total of 10 points. Marker may award 1 point for partial ideas.

1. The woman is a witness in a court case or courtroom.
2. The woman said she heard a man/the defendant say he had stolen a car.
3. The expert witness is a meteorologist or weather expert.
4. The expert said the woman could not have heard the man say he stole the car.
5. The wind was the reason she could not have heard the man.

Clarifications made on rubric after marking included:

Point 1: Student must mention that the woman is either a 'witness' or in a 'courtroom' or 'court case' or 'court' for two points.

Point 2: Student must include woman *heard* the man say something and that the car was stolen, but not necessarily that *the man* stole the car. (In the text, the use of past participle – *stolen* – along with the reporting of the action in quotation marks may have caused confusion for some students: some wrote that the man told the other man to steal the car, perhaps confusing it with the command verb).

Point 5: It was agreed that *weather* was an acceptable substitute for *wind*. So, mention of only the wind or the weather was worth one point. Mention of the wind or weather being *the reason they couldn't hear the voice* was worth two points.

4.4 Consent form

Consent Form for Participating in Research

I (Kimberly Klassen) would like to conduct research on reading skills of Japanese learners of English. If you agree to participate in this research, your performance will not affect your grades for this course in anyway. Also, your decision to participate in this research or not will not affect your grade in this course.

Your name will NOT appear in published results of this research. Likewise, your class number will NOT appear in published results. In sum, there will be no way to identify you personally as a participant.

Taking part in the project will involve reading a text and completing a short test afterwards, and this will take approximately 20 minutes, in three classes (for a total of 60 minutes). You can withdraw from this study afterwards without giving a reason. You are free to ask questions at any time. The data collected from this research will be held confidentially.

日本人学習者の英語のリーディング・スキルに関する調査のご協力をお願いします。この調査での内容が授業の成績に影響することは一切ありません。また、この調査に参加するか否かが成績に影響することもありません。

協力して頂いた方の名前や学籍番号は研究結果の中に記載致しませんので、あなた個人がこの調査の参加者として特定されることはありません。

この調査では、テキストを読んで短いテストに答えて頂きます。時間はおよそ６０分程度です。理由なしにこの研究への参加をやめることができます。質問があればいつでもお尋ねください。研究から得られたデータは内密に扱われます。

この調査にご協力頂ける方は、下記にご署名をお願いします。

If you agree to take part in this research, please sign the consent below:

Student Name: _____Student No. _____

Date: _____Signature: _____

Appendix 5: Using context to identify proper names

5.1 Tasks A and B, participant versions

TASK A Name: _____

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

インストラクション：次の文章を読んでください。大文字に変更する必要があると思う小文字を、大文字に変更してください。辞書を使用してもかまいません。例を見てください。

I *N* *Y*

Example: ~~i~~ arrived in ~~n~~ew ~~y~~ork 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.

TASK B Name: _____

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

インストラクション：次の文章を読んでください。大文字に変更する必要があると思う小文字を、大文字に変更してください。辞書を使用してもかまいません。例を見てください。

I *N* *Y*

Example: ~~i~~ arrived in ~~n~~ew ~~y~~ork 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?

5.2 Consent form

Consent form for participating in research

I [Kimberly Klassen] would like to conduct research on reading skills of Japanese learners of English. If you agree to participate in this research, your performance will not affect your grades for this course in anyway. Also, your decision to participate in this research or not will not affect your grade in this course.

Your name will NOT appear in published results of this research. Likewise, your class number will NOT appear in published results. In sum, there will be no way to identify you personally as a participant.

Taking part in the project will involve a short reading task. This will take approximately 15 minutes and will be done twice in different class periods. You can withdraw from this study afterwards without giving a reason. You are free to ask questions at any time. The data collected from this research will be held confidentially.

日本人学習者の英語のリーディング・スキルに関する調査のご協力をお願いします。この調査での内容が授業の成績に影響することは一切ありません。また、この調査に参加するか否かが成績に影響することもありません。

協力して頂いた方の名前や学籍番号は研究結果の中に記載致しませんので、あなた個人がこの調査の参加者として特定されることはありません。

この調査では、テキストを読んで短いタスクに答えて頂きます。時間はおよそ15分程度で、二度にわたり別々の授業時間内に行われます。理由なしにこの研究への参加をやめることができます。質問があればいつでもお尋ねください。研究から得られたデータは内密に扱われます。

この調査にご協力頂ける方は、下記にご署名をお願いします。

If you agree to take part in this research, please sign the consent below:

Student Name: _____ Student No. _____

Date: _____Signature:_____

5.3 Debriefing form

I want to thank you for participating in my research. In this project, I'm investigating how context (the other words in a sentence) can help a Japanese student of English identify names when reading. You have helped me with this by reading the sentences and completing the task.

Please be assured that the data I collected from you will be held confidentially, and your name and class number will not appear in any published article. You can ask me questions about this research at any time. You can still withdraw from this study if you wish, without giving a reason. If you want to withdraw, I will delete your responses to the task from my data set. You can withdraw by telling me in an email or in person.

If you have any questions, you can contact me at kklassen@swu.ac.jp. You can also contact my supervisor at Cardiff University, Tess Fitzpatrick at FitzpatrickT@cardiff.ac.uk.

5.4 Context rater sentences

Instructions: Please read each sentence and fill in the blank with one suitable word.

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

2. he suspects babur is a secret smoker because sometimes _____ can smell the tobacco on him.

3. she led her out of the nursery and then julie picked her up when _____ came out of the school.

4. it was agreed that the statement should come from the _____ house, with immediate confirmation in london.

5. bernard _____ played a maverick detective in last night's new drama telltale.

6. chesham secretary tony greeham said _____ resigned for personal reasons, and that there was no animosity.

7. after winning, john _____ patrolled the commons tea-room, soliciting opinions on the next debate.

8. next morning i continued to make excuses for _____, but as i now knew i was making them, they sounded false.

9. i asked warren beatty not to cast sean _____ in the role for dick tracy.

10. in magisterial style, _____ dealt peremptorily with the committee's inflexible attitude.

11. martin _____ had to run along the top of a train doubling for roger moore.

12. brian _____ was the strong, solid, ever-dependable central defender.

13. the most likely cause of the dark spots, _____ reasoned, was water, a common molecule that absorbed at the wavelengths detected by his camera.

14. during the _____ hearing, the sheriff, city officials, firemen and other witnesses testified about the death.

15. the fact that he could not drive the image from his mind, struck _____ as a proof that yvette's beauty had been her curse.

16. peter _____ revealed another deeply held secret that shocked and embarrassed everyone in the court.

17. sometimes she pictured _____ and thought about a future in which they had a child.

18. i don't get the feeling _____ is under a lot of pressure.

19. take those cucumbers _____ admired so much.

20. what if _____ suddenly appeared and caught me cuddling with his wife?