The Impact of L1 Morphology on Morphological Awareness and Word Reading in English as an L2

Janani Balaji

MSc. Applied Linguistics and Second Language Acquisition, 2024

Note that some graphs/tables/images may be removed in order to comply with copyright restrictions.

DECLARATION BY THE CANDIDATE AS AUTHOR OF THE DISSERTATION

1. I understand that I am the owner of this dissertation and that the copyright rests with me unless I specifically transfer it to another person.
2. I allow the Department to deposit on my behalf a copy of this dissertation in the Oxford University Research Archive ('ORA') where it shall be freely available online for use in accordance with ORA's Terms and Conditions of Use [https://ora.ox.ac.uk/terms_of_use].
3. I understand that this dissertation should not contain material that can be used to personally identify individuals or specific groups of individuals (unless permission has been obtained from the individuals) and that such material should be removed before this dissertation is deposited in ORA.
4. I agree to be bound by the terms of the ORA Grant of Non-exclusive Licence [https://ora.ox.ac.uk/deposit_agreements] and I warrant that to the best of my knowledge, making my thesis available on the internet will not infringe copyright or any other rights of any other person or party, nor contain defamatory material.
5. I agree that my dissertation shall be available for download in ORA in accordance with paragraphs 2, 3 and 4 above.

Signed [an electronic signature is sufficient]:	Janani Balaji
Date:	08/08/2024

University of Oxford

Department of Education

MSc. Applied Linguistics and Second Language Acquisition Dissertation

'The Impact of L1 Morphology on Morphological Awareness and Word Reading in English as an L2'

Janani Balaji

Supervisor: Dr Anna-Maria Ramezanzadeh

For my parents. For all their support and their love, through the many sleepless nights and hour-long calls

For my friends, who got me through the hard days and the many cups of coffee they brought with them.

For my supervisor, who managed to guide me through all my confusion and brought me to the finish line.

Thank you

ABSTRACT

Morphological awareness can be understood as the ability to reflect upon and manipulate morphemes and employ word formation rules in one's language. Morphological awareness has been identified as an important factor in word reading and reading comprehension, as it allows individuals to decompose words into the smallest possible units of meaning, thereby allowing them to access words faster. Literature in the field has also suggested the possibility of crosslinguistic transfer of morphological awareness from an individual's L1 to their L1 and the impact of this transfer on reading comprehension and word reading in the L1. However, literature on this subject has been limited to languages within the Germanic and Romance language families (with the exception of Korean and Chinese). This study examines the possibility of a crosslinguistic transfer of morphological awareness from a morphologically complex L1 to English and the possible impact L1 morphology would have on word reading in English. This study is focused on two under-explored languages: Tamil and Hindi, with a sample of 99 participants, who are university students currently enrolled in various Indian universities. This study employs a mixed methods approach, and qualitative and quantitative data is triangulated wherever necessary in order to provide a picture of the impact of L1 morphology on morphological awareness and word reading in English.

The first part of the research consists of two proficiency tests: in English and in the respective L1s (Tamil or Hindi). This was followed by a series of morphological awareness tests and word reading tests (with real and pseudowords). Finally, a stimulated recall was conducted with interested participants to understand the strategies used by participants to complete the tasks. The data from the proficiency tests were used to ensure that participants were equally distributed in both groups. Data from the morphological awareness and word reading tests were then analyzed for potential differences in group performance. Finally, data from the stimulated recalls were analyzed to understand whether there were significant differences in the strategies used by each group. Qualitative data was also used to supplement the quantitative findings whenever necessary.

Table of Contents

CHAPTER 1: INTRODUCTION

Morphological awareness can be defined as "morphological awareness can be understood as the ability to reflect upon and manipulate morphemes and employ word formation rules in one's language." (Kuo & Anderson, 2006, p. 161). Morphological awareness has been shown to have an impact on word reading and reading comprehension. Previous studies on this topic have largely been conducted on alphabetic languages such as Spanish and Turkish, except for Chinese, which is a logographic language. Within the field of morphological awareness (and applied linguistics in general), languages such as Tamil and Hindi are under-explored and under-studied.

Tamil and Hindi are two languages that belong to the diverse linguistic environment of India. Tamil is an alpha-syllabary language that belongs to the Dravidian language family. It is spoken by 53,006,368 people in India, 3.35 million people in Sri Lanka and over 2 million people in Singapore, Malaysia, South Africa, Mauritius, Fiji and Burma (Krishnamurti, 2003). Tamil has a number of dialects: Central Tamil, Kongu Tamil, Madras Bashai, Madurai Tamil, Nellai Tamil, Kumari Tamil, Batticaloa Tamil dialect, Jaffna Tamil dialect and Malaysian Tamil (Krishnamurti, 2003).

Hindi belongs to the Indo-Aryan language family, and it is also an alpha-syllabary language. Hindi is generally considered to be anywhere between the third and fifth most widely spoken language in the word and is natively spoken by upwards of 300 million people – it is, along with English – one of the two officially recognized national languages of India (Cardona & Jain, 2003). Hindi is mainly spoken in the Indian states of Uttar Pradesh, Uttaranchal, Madhya Pradesh, Chhattisgarh, Bihar, Jharkhand, Haryana, Rajasthan, Himachal Pradesh and Delhi – these states are also referred to as the Hindi belt. Hindi has five western dialects – Braj, Bundeli, Hariyanvi, Kanauji and Kauravi and three eastern dialects – Avadhi, Bagheli and Chhattisghari (Cardona & Jain, 2003).

The purpose of this study is to understand the effect of first language morphological complexity of morphological awareness in English. This study commences with a review of literature on morphological awareness, its impact on word reading and reading comprehension and the crosslinguistic transfer of morphological awareness among speakers of multiple languages. This study tests the performance of two groups of university students (Tamil and Hindi speakers) and investigates possible differences in the performances of these groups. A mixed methods approach is used, and qualitative and quantitative data is triangulated in order to explore patters in the performance of the participants. Following this,

a chapter integrating the results of the study along with a discussion of the findings is presented. Finally, this study concludes with a chapter presenting an overview of the findings, the limitations of the study and implication for English as a Second Language (ESL) pedagogy and future research.

CHAPTER 2: LITERATURE REVIEW

2.1. OVERVIEW

This section of the study outlines the previous research conducted on the topic of morphological awareness, word reading and reading comprehension. First, this section will outline the language learning context of India, followed by the features of the languages looked at in this study (Tamil, Hindi and English). Next, this section will outline the factors that impact word reading and will finally provide examples of previous research on morphological awareness, word reading and reading comprehension.

2.2. THE LANGUAGE LEARNING CONTEXT IN INDIA

According to the 2011 Indian Census, the Indian population speaks a total of 270 identifiable languages. Of these languages, 22 languages have been constitutionally recognized as official languages (Joshi, 2022). However, it is not just the presence of a large number of languages that makes up the multilingual ethos of India, but rather, the dynamics of the relationships between the languages and their users (Mohanty, 2006). Past research conducted on the dynamics of multilingualism in India has favoured the concept of domains of language use (Parasher, 1980; Sahgal, 1991) and has found that language use in the domain of family and everyday transactions, Hindi and other regional languages were most commonly used, while English dominated the domains of friendship, education and government.

The growing demand for English is not a new one - English was first introduced into the Indian subcontinent through the colonial process, however, it has been co-opted as an official language for a variety of reasons. The acceptance of English over Hindi as a national language has been seen as a strategy for keeping Hindi from being imposed upon non-Hindi speaking states (Mohanty, 2006). Additionally, English has been projected as a global language, as it dominates global commerce, technology and science. This dominance of English across the globe and within India has led to a rising demand for English-medium of instruction (hereby EMI) as parents and students view EMI as a road to success and social mobility (Mohanty, 2006). This growing demand for EMI has led to the weakening of the positions of regional and minority languages.

In 1950, a Three Language Formula (hereby TLF) was proposed by the Indian government, which provided a loose set of guidelines for each of the states to implement in terms of languages to be taught at the school level. The original guidelines provided by the TLF were as follows:

Grades	Media
1–4	Only one language is recommended which should be the mother tongue. In case of a non-standardized mother tongue it should be the regional standard language.
5–7	The study of two languages is obligatory. The second language may be either the official language of the Union (Hindi) or the co-official language (English).
8–10	The study of three languages is obligatory. These can be the regional standard, Hindi and English in the non-Hindi-speaking states. In the Hindi-speaking states these usually are Hindi, English and either a modern Indian language (MIL) or a classical Indian language like Sanskrit or Arabic.
9–12	No language is obligatory. This decision is left to the schools.

(Hornberger & Vaish, 2009, p. 310)

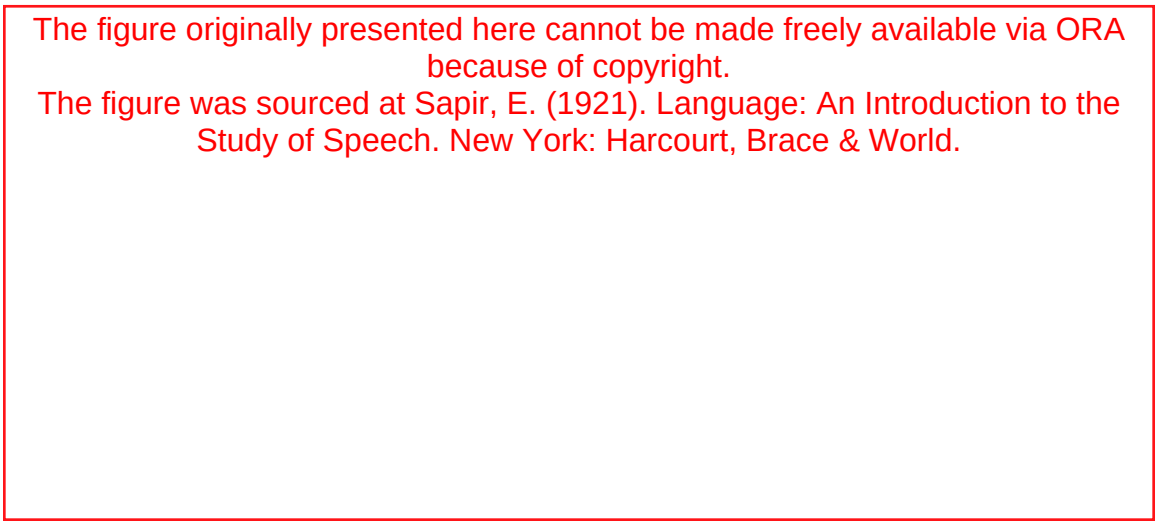
However, since 1991 when the Indian government ushered in globalization, the TLF has undergone a fundamental change. In the initial TLF, English was only offered as a subject of study in grade 5 or beyond. Since globalization has necessitated knowledge of English, schools have begun to offer English as early as grade 1 (Hornberger & Vaish, 2009). The TLF is applicable to government-sponsored education only - private education systems were free to introduce their own systems, which led to the proliferation of English medium schools. However, the spread of EMI schools has led to a new basis of stratification of Indian society between those educated in the expensive privately-run EMI schools and the less privileged who choose to go for the government-sponsored regional-language-medium schools (Hornberger & Vaish, 2009).

2.3. A GLIMPSE INTO THE LANGUAGES IN THIS STUDY

2.3.1. Tamil

Tamil belongs to the Dravidian language family along with other languages such as Telugu, Malayalam and Kannada. In terms of its orthography, Tamil uses the Brahmi script and is an alphasyllabary language (Nag & Narayanan, 2019). In terms of orthographic breadth, Tamil has a relatively constrained orthography, with 216 consonant-vowel combinations as well as a few more borrowed symbols - referred to as *grantha* and *ayudhaezhutthu*. Tamil's orthography is partially opaque as there are some cases of non-linearity in the visuo-spatial arrangement of the orthographic symbols (Nag & Narayanan, 2019). However, for the most part, Tamil has a relatively shallow orthography.

In terms of its morphology, it might be useful to classify Tamil's morphology using multiple systems of classification. According to Sapir (1921), languages can be classified into five types on the basis of synthesis (number of morphemes per word): polysynthetic, agglutinating,



fusional, analytic and isolating, with polysythentic being the most complex and isolating being the least complex. A flowchart highlighting each morphological type is presented below.

(Sapir, 1921)

According to this classification, Tamil is an agglutinating language. Another classification that might be useful to analyze Tamil morphology was proposed by Comrie (1989). Comrie suggests that it may be useful to place languages on two intersecting scales: index of synthesis and index of fusion (the number of meanings that are packed into a single morpheme). On these intersecting scales, Tamil can be understood as being highly synthetic as it is morphologically complex in terms of the number of affixes. However, Tamil has a relatively transparent morphology - where a single morpheme carries a single meaning.

Tamil is a highly inflected language. Nouns in Tamil are inflected for gender, case and number (marked by bound morphemes that are suffixed to a root noun). Verbs in Tamil are inflected for tense, person, number, gender and aspect and are again marked by bound morphemes suffixed to a root noun (Nag & Narayanan, 2019). On the other hand, derivation is rarer in Tamil. Nouns, adjectives and adverbs can be derived from nouns and verbs in Tamil through suffixation. An example has been provided below:

படி + ப்பு = படிப்பு

Padi (study, verb) + -ppu = padippu (studies, noun)

Prefixation is not as common in modern Tamil except in Sanskrit loan words which are not commonly used (Mallikadevi, 2023).

2.3.2. Hindi

Hindi belongs to the Indo-Aryan language family along with languages such as Bengali, Gujarati, Punjabi and Marathi. Hindi uses a Devanagari script and is an alphasyllabary language. Hindi has a broader orthography, with 400 symbols representing individual sounds, contributing to it being an orthographically shallow language (Bajre & Khan, 2019; Rimzhim et al., 2021).

According to Sapir's (1921) classification, Hindi is a fusional language. According to Comrie's (1989) intersecting scales, Hindi would fall on the lower end of the index of synthesis axis as it is relatively morphologically simpler than Tamil (as it has a smaller number of affixes) however, Hindi would fall on the higher end of the index of fusion. So, while Hindi is not as morphologically complex (at least in comparison to Tamil), Hindi's morphology is relatively more opaque than Tamil as one morpheme can convey multiple meanings. Hindi has many inflections. Nouns in Hindi are inflected for number and case. Gender is an inherent property of Hindi nouns; however, it is not morphologically marked but is realized via agreement (Singh & Sarma M., 2010). Hindi verbs express mood distinctions, however, no distinction is made between indicative and interrogative verbs (Kachru & Bhatia, 2018). Compared to Tamil, derivation is more common process in Hindi - the most common form of derivation in Hindi is nominal derivation. Nouns in Hindi can be derived from adjectives, verbs and other nouns through prefixation and suffixation. Additionally, adjectives and adverbs can be derived from nouns and verbs respectively through the process of suffixation (Rastogi & Khanna, 2014). An example has been provided below.

बालक + पन = बालकपन

Baalak (Boy, noun) Baalakpan (boyishness, adjective)

2.3.3. English

English belongs to the Germanic language family along with languages such as German and Dutch and is an alphabetic language. In terms of orthographic depth, English has an extremely deep orthography (grapheme-phoneme correspondences are less consistent and reliable), which makes decoding English words much more difficult. Therefore, in order to read English words, readers rely less on individual letters and instead use groups of letters, morphemes and lexical information (Miller, 2019).

In terms of morphology, English is a typologically heterogenous language (Bauer et al., 2013). The morphology of English has gone through somewhat of a transition from Old English to Contemporary English. In terms of Sapir's (1921) classification, Old English would have fallen under the category of isolating morphology, even though it had some clear fusional inflection. However, contemporary English has lost most of its fusional inflection and has become an analytic language. In terms of Comrie's (1989) index, English would rank relatively lower on the index of fusion and would fall somewhere in the middle of the index of synthesis. In English, inflection and derivation is very largely word-based. Inflectional affixes in English are attached to morphological entities that can stand alone (Bauer et al., 2013), unlike Tamil or Hindi which have stem-based as well as word-based affixation. English is a derivationally rich language - however, while English is rich in the means to derive agents and instruments, it is poor in the means of deriving patient nouns.

To summarize, table 2.1 provides a brief overview of the features of each language in this study.

Features	Tamil	Hindi	English
Morphological Type	Agglutinative	Fusional	Analytic
Inflection vs. Derivation	Inflectionally rich	Derivationally rich	
Morphological transparency	Morphologically transparant	Morphologically opaque	
Orthographic opacity	Opaque	Transparent	Opaque

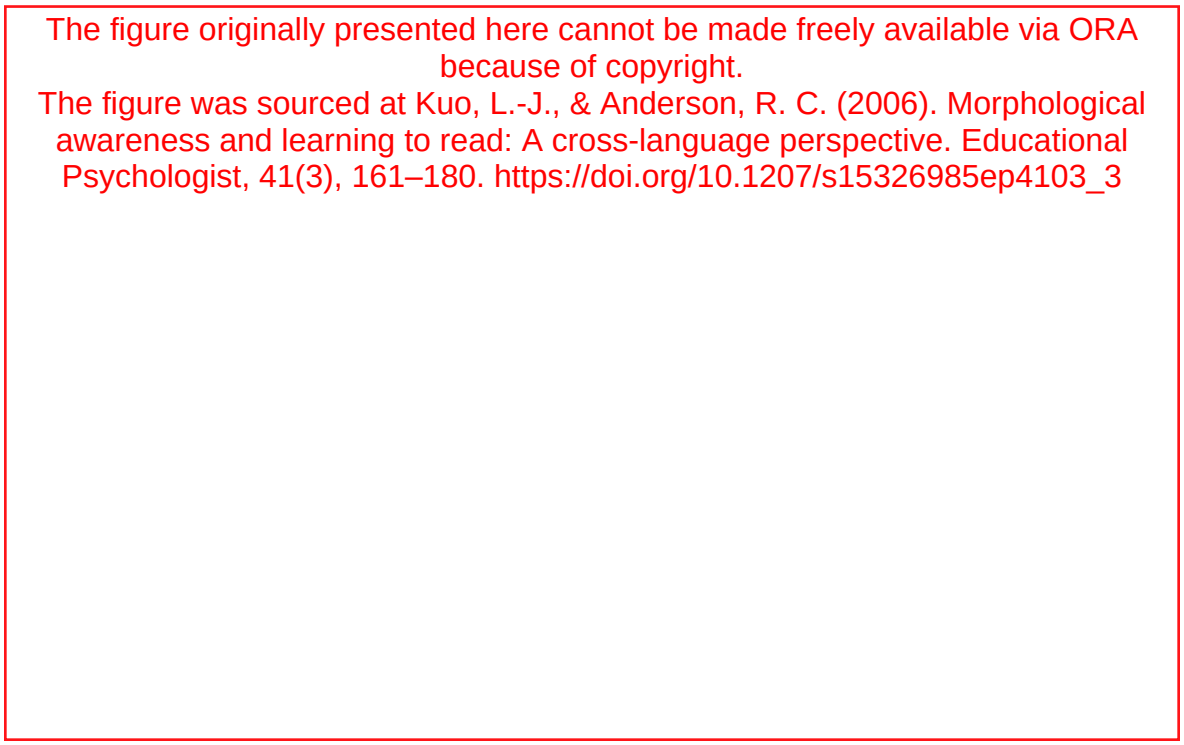
2.4. FACTORS THAT IMPACT WORD READING

Studies exploring the factors that impact development of word reading in L1 English have predominantly been conducted with children (Fumero & Tibi, 2020; Roman et al., 2009), but have yielded fairly consistent results. Five major factors have been identified to have a significant impact on the development of word reading and other literacy skills: phonetic

awareness, orthographic knowledge, word memory, vocabulary knowledge and morphological awareness (Fumero & Tibi, 2020). Phonological awareness can be defined as a "set of linguistic and metalinguistic skills concerned with children's sensitivity to the sound structure of spoken words and their ability to detect and manipulate sounds" (Fumero & Tibi, 2020, p. 573). Phonological awareness has been identified as a key factor in the development of early reading skills as children must be able to perform tasks such as counting syllables, rhyming and segmenting words into individual phonemes (Fumero & Tibi, 2020). Orthographic knowledge refers to "children's ability to learn the relationships between alphabetic symbols and the corresponding phonemes" (Fumero & Tibi, 2020, p. 573). Word memory refers to an individual's ability to store any oral or visual material that is presented to them - difficulties in word memory can impact their ability to access stored phonological or orthographic information, therefore hindering their reading speed or accuracy (Fumero & Tibi, 2020). Finally, both vocabulary knowledge and morphological awareness are interrelated in their impact on word reading - vocabulary knowledge enhances the knowledge of word segments (such as roots and affixes) which would positively impact word formation skills, while morphological awareness aids in rapid word recognition and written word pronunciation (Fumero & Tibi, 2020).

To understand how these skills develop in children as they develop and whether the skills drawn upon during reading change across grades, Roman et al. (2009) conducted a study with 92 students in grades 4, 6, and 8 in Nova Scotia, Canada. The researchers based their study on two major models: Ehri's (1995) Phase model which states that children go through different phases as they learn to read, and Share's (1995) phonological and self-teaching hypothesis, which states that while skills do develop in a particular order, phonological awareness and orthographic knowledge both develop first, followed by the other skills as they are core skills. In Roman et al's (2009) study, the participants were made to complete a series of tasks testing phonological awareness, orthographic knowledge, morphological awareness and rapid automatic naming ability. They found that phonological awareness, orthographic knowledge and morphological awareness each made a significant independent contribution to real word and pseudoword reading beyond the controls of age. While morphological awareness contributed to both real and pseudoword reading equally, phonological awareness was more important for pseudowords, and orthographic knowledge was more important for real words.

In their systematic review of the factors that impact reading development, Kuo and Anderson (2006) outlined the different aspects that were included in the umbrella construct of metalinguistic awareness:



The three large circles refer to the dominant aspects of metalinguistic awareness - phonological awareness, orthographic awareness and semantic awareness. Here, semantic awareness refers to the knowledge of how meanings are organized in a language (Kuo & Anderson, 2006). The researchers also presented a few more aspects that arise from the intersection of the dominant circles - grapho-phonological awareness (knowledge of the grapheme-phoneme and phoneme-grapheme conversion rules of a language), grapho-semantic awareness (knowledge of how semantic information is encoded in orthography), grapho-morphological awareness (ability to co-ordinate orthographic, phonological, and semantic information during reading) and morphological awareness (Kuo & Anderson, 2006).

2.5. WHAT IS MORPHOLOGICAL AWARENESS?

At its very core, morphological awareness is defined as the ability to reflect upon and manipulate morphemes in one's language (Kuo & Anderson, 2006), Specifically, it "comprises of primarily knowledge about the pairing of sound and meaning in a language and the word formation rules that guide the possible combination of morphemes." (Kuo & Anderson, 2006, p. 161).

Morphological awareness can be broken down into three sub-concepts: relational awareness (the knowledge of morphologically related words with the same base), syntactic awareness (the knowledge of the syntactic properties of derivational morphemes) and distributional awareness (sensitivity to any constraints on intraword morpheme concatenation) (Koda et al., 1998).

2.6. THE IMPACT OF MORPHOLOGICAL AWARENESS ON READING

Research on morphological awareness and its impact on reading has focused on both individual word reading (Havas et al., 2015; Ramirez et al., 2010; Schiff & Calif, 2007) and reading ability in general (Carlisle, 2000; De Freitas et al., 2018; Deacon & Kirby, 2004; Kieffer & Lesaux, 2012; Kirby et al., 2012; Marinova-Todd et al., 2013). These studies have been conducted mainly with young learners from various language backgrounds, although a few studies were conducted with adult learners, with somewhat consistent results.

There may be three main reasons as to why there would be a strong relationship between morphological awareness and learning to read. Firstly, since morphemes have semantic, phonological and syntactic properties (Mahony et al., 2000), morphological awareness would be integrally related to other aspects of language knowledge and may provide a more general index of metalinguistic capability than phonological or syntactic awareness considered on their own (Carlisle, 1995). Secondly, psycholinguistics studies with adult learners have shown that morphological information is utilized when processing complex words (Clahsen et al., 2003). Therefore, the fact that the adult mental lexicon is morphologically organized therefore suggests that morphological awareness may serve as a framework to effectively store words (Sandra, 1994). Kuo and Anderson (2006) suggest that therefore children with more developed morphological knowledge may have an advantage in acquiring and retaining complex vocabulary in other languages. Since vocabulary is a strong indicator of reading performance (Anderson & Freebody, 1981), morphological awareness should play a substantial role in reading development. Finally, morphological awareness may provide readers with additional insight into the writing system of a language (Nagy et al., 2003) since many writing systems encode both phonological and morphological information into their writing systems, thereby, an increased morphological awareness may lead to a benefit in reading comprehension.

2.6.1. Impact of morphological awareness on reading and comprehension in L2

A majority of the studies in this section have focused on monolingual learners (Carlisle, 2000; De Freitas et al., 2018; Deacon & Kirby, 2004; Kirby et al., 2012). All these studies showed consistent results - morphological awareness indeed had an impact on reading ability.

In her study of the relationship between awareness of word structure and reading comprehension, Carlisle (2000) administered word reading and reading comprehension tests along with a test of morphological structure (testing for derivation and decomposition) and a test of English vocabulary knowledge among thirty-four third graders and twenty-four fifth graders in the USA. She found that while the morphological awareness measures accounted for 41% of the variance in vocabulary and 43% of variance in reading comprehension, the measure accounted for greater variance at grade 5, with 53% in vocabulary and 55% in reading comprehension. The results indicated that not only did morphological awareness have a positive impact on reading comprehension, but also that its impact increased with age. However, one major limitation with this study was that students' interactions with English outside of the school environment (such as reading habits) were not accounted for, therefore, variance could also be attributed to differential exposure to print and English orthography.

Deacon and Kirby (2004) explored the impact of morphological awareness on reading development (beyond the impact of phonological awareness and intelligence among different age groups). They conducted a longitudinal study with 143 students in grade 3 (the number of participants fell to 103 by the time the children reached grade 5). Over the course of two years, the researchers administered a series of tasks measuring phonological awareness, morphological awareness, word reading, reading comprehension as well as verbal and nonverbal intelligence. They found that morphological awareness made a small but significant contribution to reading development, surviving the controls of phonological awareness and intelligence. They also found that morphological awareness contributed greatly to reading comprehension and pseudoword reading in comparison to real word reading. The impact of morphological awareness on reading also increased as the participants grew older.

De Freitas et al. (2018) wanted to understand whether morphological awareness helped enhance reading comprehension and word reading in Portuguese. Their participants were 132 fourth graders from Brazil. According to Ehri's (2005) Phase model, children go through a 'Consolidated Phase' where they become familiar with letter patterns than occur/recur in complex or longer words (Ehri, 2005). Based on this theory, the researchers suggested that

letter patterns such as the affix morphemes could be identified as units to facilitate accessible meanings. Therefore, they hypothesized that morphological analysis would benefit reading comprehension more than any other reading ability. The researchers administered a series of tasks measuring word reading, reading comprehension, nonverbal ability, phonological awareness and morphological awareness. They found that morphological awareness did make an independent contribution to several aspects of word reading and reading comprehension in Portuguese and the effects were larger for reading comprehension (8 to 10%) than word reading (3 to 4%). The researchers suggested that morphological awareness would help the children decode new and unfamiliar words, thereby improving speed and accuracy in word reading.

Kirby et. al (2012) conducted a longitudinal study with 103 Grade 3 students in Canada to understand the contribution of morphological awareness to reading ability. While previous research suggests that morphological awareness predicts pseudoword reading and text comprehension better than real word reading (Carlisle, 2000; Deacon & Kirby, 2004; Kuo & Anderson, 2006), the researchers suggest that since reading processes are independent, it makes sense that the benefits of morphological awareness will extend to other aspects of reading, rather than just reading comprehension (Kirby et al., 2012). The researchers administered tests of verbal ability, phonological awareness, morphological awareness, word reading and text reading. These tests were administered three times to the participants over the course of the study. The findings mainly showed that the effect of morphological awareness differed according to the grade in which it was assessed. In the first grade, morphological awareness did not contribute to word reading. In the second grade, the effect of morphological awareness increased to 1 to 4% and in grade three the effect of morphological awareness rose to 3 to 9%. Therefore, the researchers concluded that the relationship between morphological awareness and reading comprehension became stronger with the grade, which indicated an increase in reading for meaning. Finally, the researchers found that the effects of morphological awareness were stronger for the text-based tasks.

While the above mentioned research mainly focuses on monolingual child learners, some research has also been conducted with bilingual children (Kieffer & Lesaux, 2012; Marinova-Todd et al., 2013), yielding mixed results as to whether morphological awareness and reading comprehension are related and whether the impact of morphological awareness can be attributed to the L1 morphology rather than L2 exposure.

Kieffer and Lesaux (2012) investigated the effect of morphological awareness on reading comprehension in English - specifically they looked at whether morphological awareness had any direct or indirect contributions to reading comprehension. The participants in their study were 952 sixth grade students in southern California. Of these students, 323 students were native English speakers, 499 were Spanish speakers and 48 were Vietnamese speakers. The researchers administered tasks to test for morphological awareness, reading comprehension and reading ability. The study indicated that morphological awareness in participants' L1 indeed made a direct contribution to reading comprehension in English - the researchers attribute this contribution to the fact that morphological awareness would allow readers to access the meaning of complex words that cannot be decoded phonetically, therefore they would be able to comprehend the text better. Additionally, morphological awareness contributed to reading vocabulary which indirectly contributes to reading comprehension. However, there were a few limitations with this study. The main limitation was that the researchers did not include any measures of the L1, which meant that it could not be determined whether differences in the levels of morphological awareness could be attributed to the different L1 morphologies. Additionally, the study did not account for language use at home and the different opportunities that the students might have had to develop their written and oral skills.

Marinova-Todd et al. (2013) conducted a study among native English children as well as children from various language backgrounds to compare their morphological awareness and literacy skills in English. Their focus was on understanding the connections between morphological awareness, reading comprehension, real word and pseudoword reading. The participants in this study were 888 native English speakers while 244 students were English language learners (ELL) from various language backgrounds. The researchers tried to represent different morphologies in the study: Fusional (English, German, Afrikaans, Swedish, Norwegian, Dutch, Serbian, Croatian, Polish, Slovak, Bulgarian, Russian, Spanish, Italian and French), and Isolating (Cantonese and Mandarin and Agglutinative (Korean, Pilipino, Tagalog, and Persian (to an extent)). The participants were asked to complete a series of tasks measuring morphological awareness, syntactic awareness, phonological awareness, word reading, pseudoword reading, word reading fluency, reading comprehension, real word spelling and pseudoword spelling in English. The researchers found that there were significant differences only in the measures of syntactic awareness, reading comprehension and one word reading task. In almost all the tasks, the students from the fusional language groups scored higher than students from the agglutinative or isolating

languages. After conducting a series of multiple regression analyses, the researchers found that morphological awareness made an independent contribution to all the reading and spelling tasks (beyond phonological awareness). Though previous research has shown that the existence of a complex L1 morphology would present an advantage for morphological awareness (Deacon & Kirby, 2004; Kieffer & Lesaux, 2012), the students with a morphologically complex L1 (Korean) scored lower on the tasks. One explanation for this occurrence could be that the measures used by the researchers mainly tested derivational occurrences which are far more common in the fusional languages rather than the agglutinative languages.

2.6.2. Impact of morphological awareness in L1 on word reading in L2

While most studies on morphological awareness have focused on word reading in relation to reading comprehension, there have been fewer studies focused only on word reading. From the perspective of psycholinguistic, Portin et al. (2008) suggests that morphological awareness may lead to easier comprehension of morphologically complex words by way of word recognition. Researchers suggest that there are two ways to recognize morphologically complex words: either by decomposition (retrieving the word through the mental representations of their constituent morphemes) or through the full-form mode, where complex words could be accessed through mental representations corresponding to the whole word form (Portin et al., 2008). Of the two methods, comprehending words through decomposition would be more efficient, therefore word reading would be faster (Portin et al., 2008).

Ramirez et al. (2010) wanted to assess the relationship between morphological awareness and word reading in Spanish-speaking English language learners in grades 4 and 7. The participants in this study were 97 Spanish-speaking English language learners (44 fourth graders and 53 7th graders). The participants undertook two measures of derivational morphology to evaluate morphological awareness in Spanish and English as well as word reading. The researchers found that while there was a cross-linguistic transfer of morphological awareness from Spanish to English, there was no transference the other way around. However, they could not conclude that Spanish's morphological system allowed students to read English easily. This finding hints to the possibility of crosslinguistic transfer from a morphologically complex L1 to a morphologically simpler L2.

Schiff and Calif (2007) also conducted a study to assess whether the L1 Hebrew experience would impact word reading in English (the L2). They administered a series of tasks aimed at testing orthographic-phonological task awareness and morphological awareness. The findings showed that morphological awareness in Hebrew contributed to morphological awareness in English. The researchers attributed this transference to the morphological similarity between Hebrew and English.

Havas et al. (2015) conducted a very interesting study to understand the impact of L1 morphological structure on the acquisition of a bound morpheme in a novel language. The researchers conducted this study with university students from two different language groups: Finnish and Spanish. The bound morpheme focused on in this study was a bound suffix (-ro/-so and -ga/-sa) indicating gender marking. Both Spanish and Finnish are morphologically rich languages, however Spanish has overt gender marking while Finnish does not. However, Finnish has a very rich nominal morphology compared to Spanish. The participants were presented with word-picture-pairs, where the names were pronounceable novel strings and contained the gender marking in the form of a suffix. The researchers found that the Finnish participants showed more sensitivity to morphological structure - they also were able to generalize the morphological rules with more accuracy. The findings showed that participants who had more experience with morphological decomposition showed a greater advantage in acquiring the general rule - even though there was no such rule in the L1 (in this case, Finnish).

2.7. EVIDENCE OF CROSSLINGUISTIC TRANSFER ON MORPHOLOGICAL AWARENESS

While the previous studies have provided some insight into the possibility of crosslinguistic transference in morphological awareness, this section will look into it in greater depth. According to Fumero & Tibi (2020), L1 can facilitate the learning of an L2, however, the transfer between languages can be positive or negative - in order for positive transfer to occur, the languages need to share common features, while negative transfer occurs when orthographic structures between the languages are very difference. However, as it was seen in the case of Havas et al. (2015), although Spanish shared the same feature with the novel language, the Finnish participants performed better on the tasks. Their superior performance was attributed to the morphologically rich nature of their L1, suggesting that their familiarity with morphological decomposition allowed them to generalize the features of the new language with utmost ease. The following studies all explore the possibility of transference of morphological awareness from a morphologically complex one to a simpler one.

Wu and Juffs (2022) attempted to study morphological awareness in advanced English among learners with morphologically complex first languages. The researchers compared two groups of adult English language learners with two different L1 - Turkish and Chinese. Turkish has a regular and productive morphology - it is agglutinative and is rich in inflection and derivation. On the other hand, Chinese has almost no inflectional morphology - it is isolating. The target language studied here is English, an analytic language which is rather constrained in terms of its morphology. The main goal of the study was to understand whether the effects of L1 morphology persists in adult L2 learners or whether the experience of the L2 outweighed the morphological awareness brought forward by the L1. The participants in this study were 42 native Turkish speakers, 50 native Chinese speakers and 50 native English speakers. The participants completed a vocabulary knowledge task to control for vocabulary along with five morphological awareness tasks. The study found that even after controlling for LexTALE scores, the Turkish group outperformed the Chinese group on the derivation, morphological relatedness and suffix-ordering tasks. While the English group outperformed the Turkish group on the derivation, affix-choice and suffix ordering tasks, the differences became insignificant after accounting for vocabulary knowledge. Finally, the Turkish group outperformed both the English and the Chinese group on the morphological relatedness task (even without accounting for vocabulary knowledge) which indicates that the Turkish group showed a higher knowledge of relational morphology. This study showed that the participants' L1 did have an impact on morphological awareness in the L2. Since Turkish was the only language in this study that made significant use of derivational morphology, the Turkish participants may have had an advantage in terms of generalizing morphological rules in English.

Koda (2000) conducted another study to explore the impact of the L1 processing experience on morphological awareness in the L2. The participants were 22 Chinese speakers and 20 Korean speakers. The participants were made to complete a separability judgement task and a semantic inconsistency detection task. Participants were also completed two subsections of TOEFL - reading comprehension and vocabulary. The research found that both Chinese and Korean learners were virtually equal in intraword analysis efficiency with pseudowords, however Korean learners were considerably more efficient in intraword analysis. Hence the researchers attributed the efficiency of intraword analysis to the rich morphology of Korean and its transference onto English. However, there were no measures to test L1 proficiency - Koda relied on high school education as a marker of L1 language use.

Ramirez et al. (2011) conducted a study to analyze the impact of L1 structure on the development of derivational and compound awareness in English. The two languages in this study were Chinese (an isolating language with very little morphology) and Spanish (a fusional language with more derivation than Chinese). The participants were 244 fourth and seventh graders (89 Spanish speakers, 77 Chinese speakers and 78 native English speakers). The participants completed a series of tasks measuring vocabulary knowledge, nonverbal intelligence, morphological awareness, compound awareness and word reading. The findings suggested that Spanish speakers performed similarly to the native speakers on derivational awareness and outperformed the Chinese group, while the Chinese speakers performed similarly to the monolingual group on the compound awareness task and outperformed the Spanish group. The findings suggest that previous experience with a phenomenon in the L1 would allow for a transference into the L2.

2.7.1. Impact of the L1 on morphosyntactic processing in the L2

The concept of crosslinguistic transfer in morphological awareness from the L1 to the L2 has also been studied in the field of psycholinguistics. Portin et al. (2008) wanted to understand whether adult L2 learners employed the same morphological processing mechanisms in the L1 and the L2, especially if the languages are typologically different. The participants were Hungarian and Chinese speakers learning Swedish as an L2 after the age of 15. The participants were made to complete a visual lexical decision task. Portin found that there was a significant difference in Swedish L2 processing between the Hungarian and Chinese group. The Hungarian group showed a higher processing cost (a longer reaction time) for the low and medium frequency inflected nouns (which suggested that they decomposed these words) and used full-form processing for high frequency and monomorphemic words. In the Chinese group, reaction between the inflected and monomorphemic words was similar at all frequency levels, which suggests that they conducted full-form processing at all levels. The findings of this study reflect that there may be a possible strategy transfer between L1 and L2 with regards to morphological processing.

2.8. IDENTIFYING THE RESEARCH GAP

This study attempts to address the gaps that have not been addressed by previous research, while also attempting to account for factors that previous studies did not account for. This study will involve advanced learners of English - specifically, university students enrolled in Indian universities. Advanced learners have not been studied in past research, except for Wu and Juffs (2022). Most of the research in studies of morphological awareness have been

focused on child learners - These findings from young learners may not be generalizable to adults learning additional languages, hence the need for more studies involving adults.

Most of the studies conducted to explore morphological awareness have focused on languages that belong to the Romance and Germanic language families, except for Korean and Chinese. In terms of orthography, the languages studied in the past research have been alphabetic languages, except for Chinese, which is a logographic language. This study will be looking at native speakers of Tamil and Hindi - two alphasyllabary languages spoken in the Indian subcontinent. These languages have been under-studied in previous research in applied linguistics. This study would contribute greatly to shedding light onto these previously underexplored languages. Studying these two languages would provide insight into whether crosslinguistic transfer between L1 and L2 is significant when both languages significantly differ in terms of orthography and phonology. Additionally, both Tamil and Hindi are spoken by large populations, making their exploration very important.

Finally, this study attempts to cover some of the factors that have not been previously considered by past research. One of the main limitations of the studies on morphological awareness is that they did not include any measures to account for L1 proficiency and exposure. In this study, a test of L1 proficiency will be designed and administered to participants to determine whether L1 proficiency would impact morphological awareness in English. Participant self-ratings of reading, writing and speaking in their L1 will also be collected to triangulate proficiency measures.

CHAPTER 3: METHODOLOGY

3.1. OVERVIEW

To address the gaps in the literature on morphological awareness and word reading, this study has been designed to be exploratory and descriptive in nature. This study starts by collecting background and test data from the recruited sample of 99 participants. Additionally, qualitative data from 2 participants (distributed evenly between both groups) was collected to further supplement the quantitative data. Performance scores on the proficiency tests, morphological awareness tasks and word reading tasks as well as the self-reported proficiency tests were then analyzed to understand the differences in the participants' performance. Data on participants' task experience was also collected and analyzed to explain whether the participants approached the tasks differently based on their linguistic backgrounds. Data on the variables (morphological awareness and word reading) was collected using both qualitative and quantitative methods and triangulated where appropriate, to provide a detailed view of crosslinguistic transfer of morphological awareness and its impact on word reading speed.

3.2. RESEARCH DESIGN

This study employs a mixed-methods approach, combining both qualitative and quantitative data collection techniques, to offer comprehensive insight into the demographics of the participants and offer possible insight into any relationships between the morphological complexity of the participants' L1 and their morphological awareness and word reading speed. A mixed methods approach was adopted mainly because it holds the advantage of providing a variety of perspectives on a phenomenon (Ritchie et al., 2014). This study utilizes questionnaires, tasks and interviews to collect various quantitative and qualitative data from the participants. The participants were split into two groups: those who spoke Tamil as their L1 and those who spoke Hindi as their L1. The quantitative data collected from the participants consisted of data from language history questionnaires and scores on a series of tests (LexTALE, L1 multiple-choice cloze tests, morphological awareness tasks and word-reading tests) which were administered through Gorilla Experiment Builder. Six participants also participated in a stimulated recall and qualitative data was collected from them. In the stimulated recall, participants were presented with two tasks that they had completed previously (word attack and affix non-word choice) and they were asked to describe how they approached the task and how they decided the answers for the tasks. This was done in order

to understand how individual participants approached the questions provided to them - specifically focusing on the decoding processes that the participants made use of.

It must be noted that such an exploratory study inevitably involves several confounding variables. Every attempt has been made to account for these variables or to document them in order to analyze their impact on the outcomes where possible. Since this study involves participants from a variety of backgrounds, their individual environments (in terms of school, university, and language use) may vary. To account for these differences, an attempt was made to restrict the sample to those participants who fit the inclusion criteria. Additionally, information was gathered (via questionnaires) on the participants' language history – their language usage at home, school and university. These variables were recorded to take account of them where possible and to account for their possible impact on the outcomes.

3.3. RESEARCH QUESTIONS

A total of five research questions (RQs) were formulated to explore patterns in the performance of Tamil-speaking and Hindi-speaking participants. The questions are as follows:

1. What is the impact of L1 morphological complexity on task performance in English?
    a. What is the impact of the morphological complexity of L1 on morphological awareness in L2 English?
    b. What is the impact of the morphological complexity of L1 on word reading speed in L2 English?
2. What is the impact of English proficiency on task performance in English?
    a. To what extent does English proficiency impact performance on morphological awareness tasks?
    b. To what extent does English proficiency impact performance on the word reading tasks?
3. What is the impact of L1 proficiency on task performance in English?
    a. To what extent does L1 proficiency impact performance on morphological awareness tasks?

b. To what extent does L1 proficiency impact performance on the word reading tasks?

4. To what extent does the L1 of the participant impact how they approach the provided tasks?

5. How do participants with different English proficiency levels approach the provided tasks?

3.4. PARTICIPANTS

3.4.1. Population

The population in this study consisted of 53 Tamil-speaking and 46 Hindi-speaking university students. These students were between the ages of 18 to 24 and were enrolled in universities around India.

3.4.2. Sampling

Participants were mainly recruited through three methods: through a contact database; through snowballing and through social media. Only those participants that fit the inclusion criteria outlined below were invited to participate in the study.

Inclusion criteria

- Enrolled in an Indian university
- Must be able to read and write in their respective L1 (Hindi or Tamil)
- Must have at least a B1 level of proficiency in English

3.4.3. Recruitment

Once ethical approval was obtained (Appendix A), a poster (Appendix B) was created to advertise participant recruitment. Two QR codes were presented on the poster, which linked to the participant information sheet (Appendix C) and to the experiment. The researcher's contact details were also provided on the poster, so that participants could contact the researcher in case of any queries. Alongside this poster, a message was shared (Appendix D), outlining the details of the study, along with links to the participation information sheet and the experiment. The consent form (Appendix E) was integrated into the experiment and

participants could progress with the experiment only when they had completed the consent form.

The poster and the accompanying message were circulated to the researcher's colleagues and acquaintances, who forwarded it to interested participants who fit the criteria outlined by the researcher. The poster and message were also shared on social media (Instagram and Facebook) and participants were urged to contact the researcher in case of any questions.

For the stimulated recall, participants indicated their interest in participating through a form presented at the end of the online experiment (Appendix F). The interested participants were rerouted to another information sheet (Appendix G), which detailed the interview procedure. After they had read through the information sheet, the participants were routed to the interview consent form (Appendix H). Once they had signed the consent form, the finish page was displayed. The participants that had indicated that they were not interested in participating in the interview were directly taken to the finish page.

3.4.4. Participant Profiles

A total of 99 participants were recruited, with 53 Tamil speakers and 46 Hindi speakers. Initially, there were 47 Hindi speakers, but one participant had to be excluded as they had left two of the tasks incomplete. The tables below show the participant distribution in terms of age (Table 3.1), education level (Table 3.2) and gender (Table 3.3).

Table 3.1.: Participants' age distribution

Age range	No. of participants
18 - 20	25
21 - 23	38
24 - 26	36

Table 3.2.: Participants' Education Level

Education level	No. of participants
Bachelor's	48
Master's	40
Doctoral	11

Table 3.3.: Participants' Gender Ratio

Gender	No. of participants
Male	41
Female	58

3.5. INSTRUMENTS

The first part of this study involved an experiment administered through Gorilla Experiment Builder, consisting of a language history questionnaire along with a series of word reading and sentence completion tasks. The second part of the study involved a stimulated recall conducted on Microsoft Teams with the interested participants.

3.5.1. The Experiment (Part 1)

The experiment was administered through Gorilla, a cloud-based research platform that allows researchers to effectively create and deploy online experiments. The experiment involved a language history questionnaire, an English Proficiency Test (LexTALE), a multiple-choice Cloze test (either in Hindi or Tamil), two word reading tasks and four morphological awareness tasks. Initially, there were five morphological awareness tasks, however, the suffix ordering task had to be dropped due to an error with the data collection process.

3.5.1.1. Language Background and Proficiency

I. Language History Questionnaire

The Language History Questionnaire (Appendix I) was adapted from Li et al (2014). This questionnaire provides a standardized measure of language background, proficiency, usage and dominance (Li et al., 2014). Additionally, the LHQ is modular – it contains four modules containing a series of questionnaire items pertaining to users' linguistic history, proficiency in multiple languages, context and habits of language use and dominance and cultural identity of languages acquired (Li et al., 2014) and the researcher can customize the questionnaire according to which aspect of language background they want to measure. The questions in this experiment were mainly concerned with identifying participants' first language, everyday language use (languages used in home, university and with friends) and participants rating of their language abilities.

II. English Proficiency Test

In this study, LexTALE (Lexical Test for Advanced Learners of English) was adopted as a measure of English proficiency (Appendix J). LexTALE is a standardized vocabulary test created by Lemhöfer and Broersma (2012) in which participants are required to identify whether a presented English word is an actual word by clicking 'Yes' or 'No' on the screen. The test consists of 60 trials with 40 words and 20 nonwords. Participant scores are calculated

using a measure called *% correct*$_{av}$ also known as average % correct (*Scoring*, n.d.). It is calculated using the following formula:

((number of words correct/40*100) + (number of nonwords correct/20*100)) / 2

The LexTALE was available for use online at LexTALE.com and was adapted by the researcher to be administered through Gorilla. This test was chosen due to its ease of administration as well as the shorter time taken to complete the test (3 to 5 minutes), as well as the fact that it was previously used in studies on morphological awareness (Wu & Juffs, 2022). Additionally, LexTALE has been shown to be a useful and valid measure of English vocabulary knowledge for medium-proficiency learners and can be used as a ballpark measure of English proficiency. According to Lemhöfer and Broersma (2012), "LexTALE scores below 59% suggest lower-intermediate (B1) or lower proficiency level, those between 60 and 80% suggest upper-intermediate (B2) proficiency level and those between 80 and 100% suggest advanced (C1 and C2) proficiency level" (Lemhöfer & Broersma, 2012, p.341).

Only the data of participants who scored higher than 40% on LexTALE was included in the study. This was done to ensure that lack of English proficiency would not hinder participant performance in the tasks.

III. L1 Reading Tests

Testing for reading ability is an important aspect in studies of morphological awareness as phonemes and morphemes are encoded in the orthography, meaning that sensitivity to these orthographic units could play a pivotal role in decoding and morphological awareness (D. Mahony et al., 2000; Singson et al., 2000). Since this study is focused on understanding the possibility of crosslinguistic transfer of morphological awareness between L1 and L2, it can be reasonably assumed that reading ability in the L1 would determine the level of morphological awareness in the L1 and consequently, would impact morphological awareness in English. Additionally, in the case of this study, L1 reading ability is being used as a measure of L1 proficiency, as previous studies exploring crosslinguistic transfer of morphological awareness between languages have not included measures to account for L1 proficiency. L1 proficiency as measured by these tests will be used as a control, to account for any variance in proficiency - participants whose score corresponds to CEFR A2 or lower will not be included in this study.

To test for reading ability in the L1, a researcher-developed multiple-choice cloze test was administered as there was no existing test for this purpose in either Tamil or Hindi. In a standard cloze test, a passage is provided, and a number of words are deleted and replaced with blanks of standard length. Students are required to complete the passage by filling in the

correct words or their equivalents (Farley & McCollam, 2004). However, in a multiple-choice cloze test, participants are asked to choose the option that they felt best completed the sentence. In the case of this study, a multiple-choice cloze test was chosen because non-alphabetic keyboards could not be integrated into Gorilla, therefore participants would face a difficulty in completing this section of the questionnaire. Multiple choice cloze tests have been shown to be a reliable test of reading comprehension especially in the context of English as a Second Language (Aitken, 1978; Ajideh & Mozaffarzadeh, 2012; Panackal & Heft, 1978). The multiple-choice cloze test would circumvent the productive aspects of the cloze test, which would aid the student in focusing on text comprehension, making it an effective test of reading ability (Porter, 1976). While the validity of the multiple-choice cloze test has been questioned due to the drawback that test takers can guess the right answer without fully understanding the reading material (Alderson, 2005), validity can be ensured, especially if distractor items are chosen carefully (Ajideh & Mozaffarzadeh, 2012).

Tamil

First, a Tamil word frequency list (containing 5000 words along with their frequency rankings) was obtained from 'Tamil Words Frequency Processing' - a repository on GitHub containing a Python script that allows an efficient filtering of words and their frequency (Parthiban, n.d.-b). This repository drew from a corpus referred to as the Kaggle dataset, which provides a comprehensive collection of Tamil language text data for Natural Language Processing. This dataset includes a variety of text sources such as news articles, books, blogs and social media (Parthiban, n.d.-a).

The word list was then broken down into frequency bands aligning with different levels of this CEFR. This division was done along the lines of the suggestion made by Milton (2010) and Milton & Alexiou (2009) who attempted to map X_Lex scores to the CEFR levels. The table below shows the range of X_Lex scores and how they map onto the CEFR.

The figure originally presented here cannot be made freely available via ORA because of copyright.
The figure was sourced at Milton, J., & Alexiou, T. (2009). Vocabulary Size and the Common European Framework of
Reference for Languages. In B. Richards, M. H. Daller, D. D. Malvern, P. Meara, J. Milton, & J. Treffers-Daller (Eds.), *Vocabulary Studies in First and Second Language Acquisition: The Interface Between Theory and Application* (pp. 194–211). Palgrave Macmillan UK. https://doi.org/10.1057/9780230242258_12

On the basis of this list, the word frequency list was broken down into bands of 1000 words each. Altogether, there were five bands available, each containing 1000 words. In the case of this study, only individuals with L1 proficiency of B1 and above were to be studied, therefore the researcher only included words from Band 2 and above.

Once the sub-lists were created, a random number generator was used to randomly choose 10 items from each level. Therefore, the final test (Appendix K) had 40 items. Once the test items were chosen, the accompanying sentences were determined by taking example sentences from the Cambridge English to Tamil Dictionary (n.d.). The test sentences were taken from this dictionary because of the simplicity of the example sentences, which meant that the other words in the sentence would not be lower in frequency in comparison to the test item.

In terms of selecting distractors, there have been two methods used in previous research: rational selection and empirical method. In the rational selection method, experts select distractors by considering any possible sources of examinee errors while the latter uses the errors derived via pretesting of a cloze test (Hale et al., 1988). In this study, the rational method was chosen as the empirical method was not practical owing to the limited time available to the researcher. In terms of effectiveness of measuring proficiency, the rational selection method was shown to be as effective as empirical selection (Scholz & Scholz, 1981).

The distractors were generated by the researcher based on their previous experiences with Tamil and Hindi. Since the participants in the study were experienced speakers of the languages, the distractors chosen by the researcher were similar to the deleted item in terms of orthography and were both lexically and grammatically varied in order to be challenging enough for the participants. Each item was individually presented onscreen, with the four options provided below it.

Hindi

Apart from the word list creation, the process of creating the multiple-choice cloze test in Hindi (Appendix L) was identical to Tamil. To access the word frequency list in Hindi, the researcher downloaded and set up a Python library called ‘wordfreq’ and then obtained the top 5000 words in the Hindi dictionary along with their frequencies using the following code: (top500Hi = top_n_list (“hi” , 5000)).

Wordfreq is a Python library used for looking up the frequencies of words in multiple languages, based on multiple data sources (Speer, n.d.). Rather than being based on one

corpus, this library draws upon multiple data sources. For Hindi, the data was drawn from Wikipedia, Webpages, Twitter and Reddit (Speer, n.d.). The rest of the steps were identical to the Tamil multiple-choice cloze test.

Scoring

In each of the tests, the participants were awarded one point if they provided the correct answer and zero points if they provided the wrong answer.

3.5.1.2. Word Reading

The word reading tasks used in this study were adapted from the *Woodcock Reading Mastery Tests, Third Edition* (2011) – specifically, the word identification and word attack subtests. While the word reading tasks were initially designed to be administered in a one-on-one in-person setting, this test has been adapted to be delivered online. While the original test required the participants to read the words out loud, in this study, word reading speed will be measured by capturing the participants' reaction times for each item in these tasks. Reaction times were collected through Gorilla, which automatically measures reaction time as the difference between the display of the new screen and the response. When a new screen is displayed, a timestamp is recorded using performance.now. Once a response is recorded (for example, a keypress or a mouse click), a second timestamp is recorded and reaction time is calculated as the difference between the two timestamps (Timing in Gorilla, n.d.). Previous research conducted to measure the accuracy of administering reaction-time sensitive experiments through different experiment platforms have shown that Gorilla is a very reliable platform to collect accurate participant reaction times (A. Anwyl-Irvine et al., 2021; A. L. Anwyl-Irvine et al., 2020).

I. Word Identification

The word identification subtest (Appendix M) is a measure of single word reading in which participants are required to read individual English words and move to the next item. The stimulus items for this task were taken from Set A of the WRMT-III and contains 17 items. The items chosen were meant to be administered to participants (from multiple language backgrounds) above the age of 17, therefore they were considered challenging enough for the current study. While this test was originally meant to be administered in-person, in this study, it has been adapted into an online format. The words were presented individually onscreen,

and participants were asked to read the word aloud once and then advance onto the next word. Reaction times were collected from Gorilla.

II. Word Attack

The word attack subtest (Appendix N) is a measure of pseudoword reading in which participants are required to read individual pseudowords and move to the next word. The stimulus items from this task were taken from Set A of the WRMT-III and contains 17 items. In this test too, the chosen items were meant to be administered to students aged 17 and above. This test was administered in the same manner as the word identification subtest.

3.5.1.3. Morphological Awareness

This portion of the study was a reduplication of Wu & Juffs's (2022) paper on the transfer of morphological awareness from L1 to L2. The tasks used by Wu & Juffs (2022) were adapted from other research in the field of morphological awareness and transfer. This series of morphological awareness tasks aimed to test different kinds of morphological awareness.

I. Derivation Task

This task (Appendix O) was initially created by Carlisle (2000). In this task, participants were given a base word in English and were required to produce the derived form of the word in order to complete a carrier sentence. An example of a test item has been provided below:

My uncle is a _____ [farm]

Answer: farmer

There were 28 items in total, with a maximum raw score of 28. Participants were given one point if they answered correctly and no points if the answer was incorrect or missing. Since this task required the participants to provide the correct answer that fit the required syntactic category, this task tapped the knowledge of syntactic morphology. According to Carlisle (2000), the internal consistency estimate of reliability of this task (as measured by Cronbach's alpha) was 0.64. M

II. Affix Word Choice Task

This task (Appendix P) was initially created by Mahony (1994). In this task, the participants were required to select one out of four real word options to complete a carrier sentence consisting of real words in order to make a correct sentence. An example of a test item has been provided below:

John wants to make a good _____ on his date. [impressive, impressionable, impression, impressively]

Answer: impression

There are 26 items in total, with a maximum raw score of 26. Participants were given one point if they answered correctly and no points if they selected the wrong option. Like the derivation task, this task also tests knowledge of syntactic morphology. According to Mahony (1994), the internal consistency estimate of reliability (as measured by Cronbach's alpha) was 0.71.

III. Affix Non-Word Choice Task

This task (Appendix Q) was also created by Mahony (1994) and resembles the affix word task. However, in this case, participants were asked to choose between four suffixed non-word options to complete a carrier sentence consisting of real words.

An example of a test item has been provided below:

They _____ the data back in the office [curfamic, curfamation, curfamate, curfamity]

Answer: curfamate

There are 26 items in total, with a maximum raw score of 26. Participants were given one point if they answered correctly and no points if they selected the wrong option. Since the focus is on pseudowords, it can be ensured that participants may not make use of their lexical memory in order to make a decision and therefore had to be sensitive to suffixes themselves (Wu & Juffs, 2022). This task also tests for syntactic awareness. The reliability of this task is 0.78 as measured by Cronbach's alpha (Mahony, 1994).

IV. Morphological Relatedness Task

This task (Appendix R) was first put forth by Mahony et al. (2000). In this task, participants were required to decide whether the words in each pair were related to each other (whether one word came from the other word).

For example: Person-Personal [Answer: Yes]

This task contains 40 items with a maximum raw score of 40. Participants were given one point if they answered correctly and no points if they answered incorrectly. Since this task required participants to identify whether words were related to each other, this task taps knowledge of relational morphology (Wu & Juffs, 2022). The reliability of this task is 0.77 as measured by Cronbach's alpha (Mahony et al., 2000).

V. Suffix Ordering Task

Due to an error in the datLa collection process, where numerous entries in this task went unrecorded, the data from this task (Appendix S) was omitted.

3.5.2. The Stimulated Recall (Part 2)

The second part of the study involves a stimulated recall conducted one-on-one with two participants on Microsoft Teams. The researcher in this study was interested in understanding how participants went about answering the questions presented to them during the experiment – whether participants made use of morphological decomposition when answering some of the morphological awareness and word reading tasks.

The participants were presented with the test questions and were asked to walk the researcher through how they went about answering some of the questions. The tests looked at during the stimulated recall were the affix non-word task (for morphological awareness) and the word attack task (for word reading). Both tasks made use of pseudoword items, which meant that participants would not be able to rely on lexical memory and would have to be sensitive to affixes to decode the words. In terms of choosing the specific questions for the participants to review, the researcher aimed to choose the questions that had the most morphologically complex test items, which would require the participants to decompose the items to accurately answer the question. In total, 10 items were chosen from each test. These items were presented one by one through a PowerPoint presentation during the interview and the participants were

asked to walk the interviewer through how they chose the answers that they thought were correct.

The questions asked by the researcher (Appendix T) were prompts designed to allow the participants to explore their own thinking processes. Therefore, while there were some questions that remained consistent across all interviews, some of the questions were asked based on the previous responses of the participant to glean more information from them. The data was recorded through Microsoft Teams and transcribed. Any possible identifiers were removed. Once the transcripts (Appendix U) were obtained, the video recordings were deleted.

Stimulated recalls have been understood to be useful tools to explore cognitive processes. Prior research has shown that there are a few limitations in incorporating SR procedure into research design. In past research, the SR process usually involved a videotaped lesson or interview, supplemented by a think-aloud process - both of which are meant to supplement the subject's recall process (Lyle, 2003). Within the context of these studies, it was found that issues such as subject anxiety, conscious censoring of the recall by the subject and the limitations of a cues (i.e. being from a different perspective from the subject) can all undermine the validity of SR (Lyle, 2003). However, as long as care is taken to establish a rapport with the subject, ensuring a space for free-flow of the participants' inner voice and ensuring that the materials used in the SR mimic the perspective of the participant, these issues can be circumvented.

3.6. PROCEDURE

The data collection was conducted in three parts:

1. Once the participants read the information sheet and any queries were discussed with the researcher, the participants accessed the experiment through the link that was provided to them.

2. First, the participants were required to complete a consent form. Once they completed it, the participants moved onto the tasks. Completing all the tasks roughly took between 15 to 20 minutes. Participants names were collected in the process of data collection; however, this data was anonymized before analysis and a participant ID (a random string of letters and numbers) assigned by Gorilla was used instead. At the end of the questionnaire, participants were presented with a screen which asked them

whether they would be interested in taking part in another task. If participants indicated any interest, an information sheet was provided, followed by a consent form. Once they provided their consent, they were rerouted to the finish page. The researcher then contacted them to conduct the stimulated recall later. If a participant opted out of the second task, they were directly sent to the finish page.

3. The researcher reviewed the participants who had consented to the second task and organized their personal interviews through Microsoft Teams. During the interview, participants were presented with the individual test items and were asked to talk the researcher through how they went about answering the questions. Additional questions were asked to clarify information or to make the participant elaborate further. The interviews were recorded and then transcribed, upon which the recordings were deleted.

3.7. PILOTING

All researcher-developed test items - the L1 proficiency tests, the Language History Questionnaire and the Word reading tests were piloted among 10 participants (6 Tamil and 4 Hindi speakers) to ensure that the test items, sentences and distractors were reliable. The speakers had differing levels of reading ability in their respective languages. The researcher found that the tests were able to adequately differentiate between the different levels of proficiency among the speakers and the test scores correlated with the self-rating of the participants' reading ability.

Any feedback given by the pilot participants in terms of readability, ease of completion and correctness of items were undertaken by the researcher and implemented. After individual subtests were piloted, the whole experiment was piloted with 3 more people to ensure that the test was running smoothly.

3.8 ETHICS

Ethical clearance for this study was obtained from the Central University Research Ethics Committee of the University of Oxford before data collection commenced. Following the receipt of approval, the poster and the accompanying message was circulated on social media. Participants read the participation information sheet and contacted the researcher with any doubts. Once all their queries were answered, the participants began the experiment. They could begin the experiment only after they provided their consent. Through the information sheet, the participants were informed of the confidentiality of the study and were informed of

their right to withdraw their participation from the study at any given point until the specified date without any repercussions.

Every effort was made to endure that the participants remained anonymous. Before the process of data analysis, participant names were removed and replaced with participant IDs (randomized strings of numbers and letters) that were generated by Gorilla. The video recordings of the stimulated recalls were deleted after the transcription and anonymization of the data had been completed. Only those participants who have given the researcher explicit consent that they were alright with direct quotations being used in the research are quoted in the study.

3.9. DATA ANALYSIS

This study employs a mixed-methods design – both the qualitative and quantitative data was analyzed individually and then triangulated wherever possible to answer the relevant RQs.

3.9.1. Quantitative Analysis

The data from the Gorilla experiment was collated in SPSS (Version 29). Participants were separated into two groups based on their L1 (Hindi or Tamil).

The data from word reading and morphological awareness tasks, L1 proficiency test scores, LexTALE scores and self-reported proficiency scores were tested for normality and homogeneity of variance. Since data did not meet the assumptions for an independent samples t-test, Mann-Whitney non-parametric tests were run to see whether there were significant participant differences in the tasks. A two-factor ANOVA was conducted to test for interactions between English proficiency and task performance, with L1 group and English proficiency as the independent variables and morphological awareness and word reading scores as dependent variables. Finally, a Spearman rank-order correlation test was conducted to understand if there was a significant correlation between L1 proficiency and task performance.

3.9.2. Qualitative Analysis

Data from the stimulated recalls was transcribed, analyzed and coded inductively using categories emerging from the data, following the Grounded Theory approach (Strauss & Corbin, 1998). The data from the stimulated recalls were supplemented with qualitative data from the Language History Questionnaire whenever applicable.

3.10. METHODOLOGICAL LIMITATIONS

There were several identifiable limitations involved in the instruments and the sample. Since the participants of this study come from different universities, several factors related to language use and educational background differed among participants. These external factors could contribute to patterns in the performance between Tamil and Hindi speakers.

There are also limitations in terms of the research instruments. Both the word reading tasks were based on reaction times, which was measured as the time between when an item was displayed onscreen and the button was pressed to move on to the next item. This measure of word reading might open itself up to a few limitations as participants may skip through the items due to a lack of motivation, which may cause unreliable outcomes for the measure of word reading speed.

Another limitation was identified in the L1 proficiency tests, especially with the Tamil proficiency test. A lemmatized word frequency list in which words were divided into word families and then organized according to frequency was not available for Tamil. Since Tamil is a highly inflected language, this meant that words belonging to the same word family were counted as individual words, which lowers the representativeness of the Tamil word frequency list. Hence, the creation of a lemmatized word-frequency list would improve the representativeness of the Tamil proficiency test.

Additionally, due to a data collection error with Gorilla led to a large amount of missing data from the suffix ordering task, which meant that this task had to be completely forfeited.

Another limitation was identified in the stimulated recall process. As noted by Lyle (2003) – for stimulated recalls to be effective, the time between the completion of the task and the stimulated recall must be kept short so that the task is still present in the minds of the participants. However, due to issues with feasibility of conducting the stimulated recall immediately after the first task, there was a delay in conducting the stimulated recall with the participants – participants completed the stimulated recall a week after they had completed the tasks. This may have reduced the effectiveness of the stimulated recall to an extent.

CHAPTER 4: FINDINGS AND DISCUSSION

4.1. OVERVIEW

This chapter consists of an integrated presentation of the findings of the data collection for each research question, along with a discussion of the findings. This approach was chosen to be able to fully utilize the available qualitative and quantitative data and weave them together whenever necessary.

This chapter will begin with some descriptive statistics outlining the characteristics of the participants in this study, dividing them into two groups based on their L1, describing their language use patterns (at home, school, with friends and with university) as well as the self-ratings of their L1 proficiency (in reading, writing and speaking), L1 proficiency (as per their performance on the L1 test) and their English proficiency (based on LexTALE scores). The chapter then progresses into a quantitative comparison of the performance scores of both groups on the morphological awareness and word reading tasks. This was followed by further statistical analysis to understand the impact of other independent factors such as L1 proficiency and English proficiency on performance in the morphological awareness and word reading tasks. Finally, qualitative data from the interviews is investigated to provide insights into whether the morphological complexity of the participants' L1 would determine how they approach the test items. The results pertaining to each of the research questions are discussed in relation to existing findings in the broader literature.

4.2. OVERVIEW OF PARTICIPANT CHARACTERISTICS

In total, data was collected from 99 participants from various cities and towns in India – although the participants were mostly based in Chennai, Bangalore, Mumbai and Delhi. Participants were asked to indicate their L1 and were divided into groups based on L1: Tamil (n = 53) and Hindi (n = 46). The age and gender distribution of participants are displayed in table 4.1 and 4.2. Two participants (one Hindi participant and one Tamil participant) took part in the stimulated recall.

4.1 Age distribution of participants		
Age range	**Frequency**	**Percentage**
18 – 20	25	25.25%
21 – 23	38	38.38%
24 – 26	36	36.36%

4.2 Gender distribution of participants		
Gender	**Frequency**	**Percentage**
Male	41	41%
Female	58	59%

4.2.1. Self-reported language use

In the language history questionnaire, the participants were asked to describe their language use habits in terms of the languages they use to communicate in different domains: at school, in university, at home and with their friends. The results are displayed in tables 4.3, 4.4, 4.5, 4.6 and 4.7.

4.3 Language use in school – Medium of education					
Tamil			**Hindi**		
Language	**Frequency**	**Percentage**	**Language**	**Frequency**	**Percentage**
Tamil only	19	36%	Hindi only	21	46%
English only	15	28%	English only	15	33%
Tamil & English	19	36%	Hindi & English	10	22%

The participant distribution according to the languages used by their teachers in their former schools is reflective of the role of language in education according to the Three Language Formula in post-globalization India, accompanied by the rise of the demand for English medium schooling (Hornberger & Vaish, 2009). Table 4.3 also shows that the distribution of participants in terms of languages used by their teachers at the school level is relatively even.

4.4 Language use in university – Medium of education					
Tamil			**Hindi**		
Language	**Frequency**	**Percentage**	**Language**	**Frequency**	**Percentage**
English only	41	77%	English only	45	98%
English & Tamil	12	23%	English & Hindi	1	2%

The participant distribution according to the medium of education at the university level is reflective of the university system in India. In India, a majority of the colleges and universities operate in the English medium, however, at both the undergrad and postgrad level is largely

defined by lecture delivery and notes dictation, which may sometimes occur in the regional language (Jayaram, 1993). While Table 4.4 was mainly focused on the medium of teaching that participants currently experience, another question in the Language History Questionnaire (LHQ) asked participants to describe the languages that they themselves used to communicate at university - this could include communicating with peers, asking questions etc. The results for this question are displayed in Table 4.5.

4.5 Language use in university – Language of community					
Tamil			**Hindi**		
Language	**Frequency**	**Percentage**	**Language**	**Frequency**	**Percentage**
English only	23	43%	English only	24	52%
English & Tamil	30	57%	English & Hindi	22	48%

This table clearly demonstrates the bilingual milieu of the university system - despite classes occurring in English, regional languages are commonly used in the classroom situation and within the university system in general.

The LHQ also contained a section on understanding language use in the personal domains of family and friends. The results of these two sections are displayed in Tables 4.6 and 4.7 respectively.

4.6 Language use in university – Language of community					
Tamil			**Hindi**		
Language	**Frequency**	**Percentage**	**Language**	**Frequency**	**Percentage**
Tamil only	29	55%	Hindi only	13	28%
Tamil & English	9	17%	Hindi & English	7	15%
Tamil & Tulu	3	6%	Hindi & Marathi	7	15%
Tamil & Malayalam	3	6%	Hindi & Bangla	5	11%
Tamil & Telugu	3	6%	Hindi & Marwari	5	11%
Tamil & Kannada	2	4%	Hindi & Punjabi	2	4%
Tamil & Sourashtra	2	4%	Hindi & Bhojpuri	2	4%
Tamil & Urdu	1	2%	Hindi & Konkani	2	4%
Tamil & Dhakni	1	2%	Hindi & Urdu	2	4%
			Hindi & Odia	1	2%

From the table above, it can be noted that English does not play a serious part in the family domain - instead, other regional languages are used, which are spoken in addition to participants' L1. However, most participants in both groups spoke only their L1 at home. This could be a result of the fact that the participants in this study resided in several different cities and were therefore exposed to multiple languages throughout their lives. The findings of the LHQ somewhat differ from the findings of Parasher (1980) and Sahgal (1991) in the sense that while the dominant language in the family domain is the L1, other regional languages have started to enter this domain. This could be indicative of the fact that many individuals in this study came from families which had emigrated to other parts of India and therefore learned the language of the city they had moved to and began incorporating it into their daily lives – however it must be kept in mind that participants were merely asked to indicate any other languages that they used with their family, not about how often they used it or how proficient they were in these languages.

4.7 Language use with friends					
Tamil			**Hindi**		
Language	**Frequency**	**Percentage**	**Language**	**Frequency**	**Percentage**
Tamil only	16	30%	Hindi only	4	9%
English only	2	4%	English only	1	2%
Tamil & English	26	49%	Hindi & English	35	76%
Tamil, English & Telugu	4	8%	English, Hindi & Tamil	2	4%
Tamil, Hindi & English	2	4%	English, Hindi & Bengali	2	4%
Tamil, English & Tulu	1	2%	English, Hindi & Marwari	2	4%
Tamil, English & Malayalam	1	2%			
Tamil, English & Arabic	1	2%			

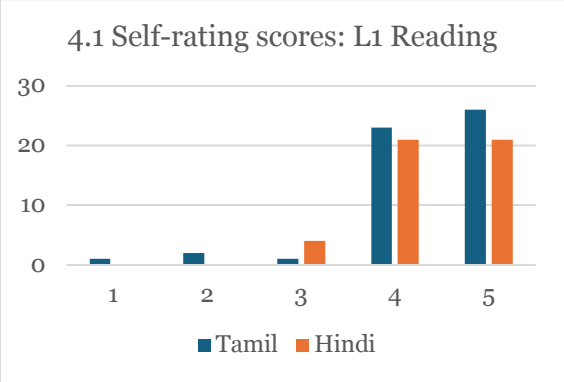
In the domain of friendship, the rise of English and its use among the participants is evident - however, unlike Parasher (1980) and Sahgal (1991), in whose studies English was seen as the dominant language in the friendship domain, here, English is integrated with the participants' L1 and the other regional languages. It must be noted that in the previous research, participants resided in elite neighborhoods where English remained the dominant language as almost all the participants were educated in English. However, in this study, the participants came from more diverse socioeconomic backgrounds, therefore the languages used in the domain of friendship are diverse. However, despite the use of Hindi and regional

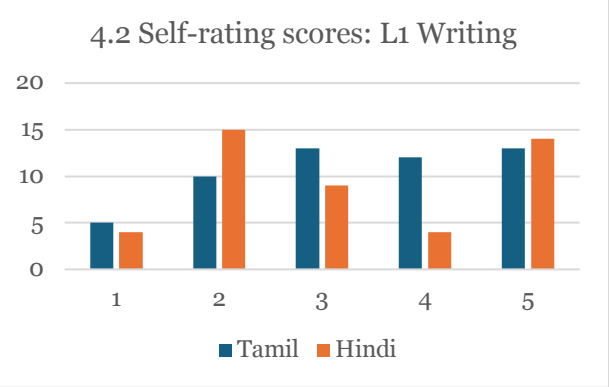
languages, most of the participants use English to some extent – one possible reason for the proliferation of English in this domain could be the rise of social media and the greater exposure to English that it brought with it.

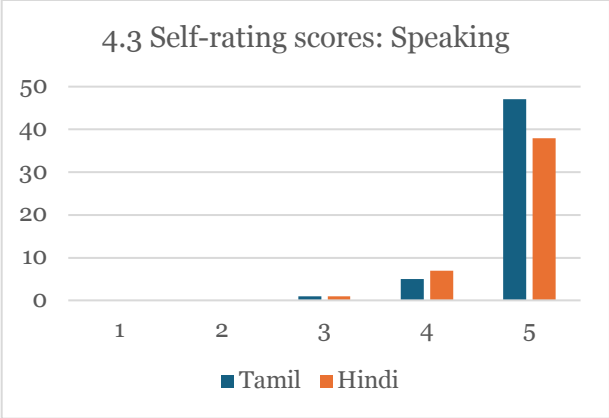
4.3 PARTICIPANT PROFICIENCY

4.3.1. Participant self-rating of L1 proficiency

In the LHQ, participants also rated their proficiency in L1 in three skills: reading, writing and speaking. Participants rated each skill on a Likert scale from 1 to 5, with 1 being 'not at all confident' and 5 being 'very confident'. A summary of participant self-ratings on each skill according to L1 is presented in graphs 4.1, 4.2 and 4.3. A detailed breakdown of raw data can be found in Appendix V.







To test if there were any significant differences in the participant self-ratings between groups, a test of between-group differences was run. The self-rating scores were tested for normality and homogeneity of variance, using the Kolmogorov-Smirnov and Levene's test and the data did not meet the standards required for an independent samples t-test, hence a non-parametric Mann-Whitney U test was applied. The results of the test are displayed in table 4.8.

4.8 Differences between Tamil and Hindi participants on different skills

	Reading	**Writing**	**Speaking**
Mann-Whitney U	1179.500	1102.000	1201.500
Z-Score	-0.309	-0.843	-0.203
Significance (2-tailed)	.757	.399	.839

As visible from the table 4.8, there were no significant differences between the self-reported L1 proficiency scores among both groups across any of the three skills. Participant scores in each of the skills have been presented in table 4.9. While Hindi participants rated themselves higher than Tamil participants in reading and speaking and Tamil participants rated themselves higher in the writing skill, the differences were insignificant.

4.9 Participant differences in L1 Self-Rating

	Tamil	**Hindi**
Skill	**Mean score**	**Mean score**
Reading	4.34	4.37
Writing	3.38	3.15
Speaking	4.83	4.85

4.3.2. L1 proficiency scores

In addition to the self-report in the LHQ, participants completed a multiple-choice cloze test in their L1 to measure their proficiency in their L1, using the skill of reading as a proxy for their proficiency. This test was designed by the researcher to correlate with the CEFR levels from B1 to C2 (see section 3.5.1.1 in the methodology). The distribution of the participants according to their test scores (converted into CEFR levels) is presented in table 4.10.

4.10 L1 proficiency test scores - distribution		
CEFR Level	**Tamil**	**Hindi**
B1	1	1
B2	4	3
C1	6	11
C2	42	31
Total	53	46

The raw scores from the L1 proficiency test are presented in table 4.11.

4.11 L1 proficiency test scores – raw score		
Group	**Mean score**	**Std. Dev.**
Tamil	33.64	6.604
Hindi	33.15	8.674

The test score data was tested for normality and homogeneity of variance using Kolmogorov-Smirnov and Levene's test for each L1 group. The data did not meet the parametric standards required to perform an independent samples t-test, therefore a Mann-Whitney U test was conducted to test for any significant difference in the performance between the two groups on the proficiency tests. Though the performance of the Hindi participants (table 4.11) was better than the performance of the Tamil participants (table 4.11) according to their raw scores, the Mann-Whitney U tests revealed no significant differences between the two groups (U = 980.000, *Z-score* = -1.685, p = .092).

4.3.3. English proficiency

Finally, another important aspect of this study was participants' English proficiency, which was measured through LexTALE. Participant raw scores on this test were then converted into CEFR grades on the basis of Lemhöfer & Broersma's (2012) research. The participant distribution of scores in terms of CEFR is presented in table 4.12.

4.12 English proficiency test scores - distribution		
CEFR Level	**Tamil**	**Hindi**
B1	4	3
B2	15	11
C1	14	16
C2	20	16
Total	53	46

The raw scores from LexTALE are presented in table 4.13.

4.13 LexTALE scores – raw score		
Group	**Mean score**	**Std. Dev.**
Tamil	81.037	13.744
Hindi	83.224	13.148

Kolmogorov-Smirnov and Levene's test indicated that the data did not meet the standards for running an independent samples T-test, so a Mann-Whitney U test was performed to test for any significant differences could be found between the groups. As the results show, the difference between the Hindi participants (table 4.13) and Tamil participants (table 4.13) is insignificant (U = 1104.000, *Z-score* = -.808, p = .422).

4.3.4. Rationale behind collecting participant proficiency data

The background information on participant self-ratings were collected to establish whether the two samples were balanced in terms of L1 proficiency, at least in terms of self-rating. It can be seen from table 4.8 that both groups of participants were balanced in terms of self-rated L1 proficiency.

Participant test scores collected from L1 proficiency tests and the LexTALE served as the means of participant selection. In the L1 proficiency test, participants who scored below A2

level were not included in the study. In the LexTALE, participants who scored below a B1 level were excluded from the study.

4.4. RESEARCH QUESTIONS

4.4.1. RQ1: What is the impact of L1 morphological complexity on performance in English tasks?

4.4.1.1. What is the impact of the morphological complexity of L1 on morphological awareness in L2 English?

To answer this research question, participants were asked to complete four morphological awareness tasks: derivation, affix word choice, affix non-word choice and morphological relatedness. Data from the morphological awareness tests were analysed for between-group differences. Kosmogorov-Smirnov and Levene's test revealed that data did not meet the assumptions for independent-samples t-tests, so Mann-Whitney U tests were conducted. The results of these tests are reported in table 4.14.

4.14 Results of Mann-Whitney test for morphological awareness tasks

	Mann-Whitney U	**Z-Score**	**Significance (2-tailed)**
Derivation	600.00	-4.407	<.001
Affix Word Choice	1154.5	-0.462	.644
Affix Non-word Choice	381.00	-5.912	<.001
Morphological Relatedness	1217.50	-0.011	.992

According to table 4.14, there were significant differences in the performance of the two groups in only two of the tasks: derivation and affix non-word choice.

In the derivation task, the Hindi group (Mean = 26.37) significantly outperformed the Tamil group (Mean rank = 23.91), with a reported effect size r = - 0.4429, which can be classified as a medium effect size (Field, 2019). The derivation task was a productive task, which required participants to derive the correct form of a given word to complete the given sentence. The results of this study align with findings from previous research by Marinova-Todd et al. (2013), Ramirez et al. (2011) and Wu and Juffs (2022). Marinova Todd et al. (2013) and Ramirez et al. (2011), found that participants who spoke fusional languages tended to outperform participants with other language backgrounds in tasks that tested for derivational morphology, as derivation is a more prevalent process in these languages. In Wu and Juffs (2022), the participants who spoke Turkish (an agglutinative language with rich derivational

morphology) outperformed the other participants in this task. Therefore, these results would suggest that having previous experience with derivational morphology through the L1 would contribute to greater performance in productive tasks that measure knowledge of derivational morphology.

There was also a significant difference in the performances of the two groups was the affix non-word choice task. In this task, the participants were presented with an incomplete sentence, along with four nonsense words. The participants were required to choose the option which they felt would complete the sentence best. It was a receptive task, as participants did not need to produce the words to complete the sentence. In this task, the Tamil group (Mean = 23.23) significantly outperformed the Hindi group (Mean = 16.83), with a reported effect size of r = - 0.594, which can be categorized as a large effect size (Field, 2019). This result is somewhat reflective of the findings of Havas et al. (2015) in which L1 Finnish (agglutinative language with no gender marking) participants and L1 Spanish (fusional language with gender marking) participants were tested on their ability to pick up patterns of gender marking inflection in a novel language. The researchers found that the Finnish group picked up the gender inflection faster despite the lack of gender marking in their own language. The researchers attributed this result to the fact that Finnish had more inflectional marking in comparison to Spanish, which allowed their participants to understand and generalize this finding faster in the novel language. This could also be applied to the Tamil participants in this study. Despite the lack of derivational morphology in Tamil, the presence of a complicated morphology in Tamil may have enabled participants to decompose the nonsense words into their constituent morphemes to a greater extent, which may have allowed them to answer the questions with more accuracy. This finding can also be supported by data from the interviews with the participants, in which the Tamil participant (T1) reported making use of suffixation knowledge to a greater extent than the Hindi participant (H1). Quotes from the interview detailing this process have been presented in table 4.15.

4.15 Quotations from interview, detailing usage of morphological decomposition to answer questions in affix non-word choice task	
T1	**H1**
Test item: The meeting was highly _____ and invigorating. [Options: loquarify, loquarial, loguarialize, loquarialism]	
“Yeah, I'm gonna go with B because when I said the meeting was highly -lism -lize -fy - -rial made the most sense.”	“I would say option B. So it's like loquarial …if you're describing someone, let's say that this person was highly. mercurial and invigorating or was high, was highly mercurial and entertaining”

While both participants answered the question correctly, two different strategies were used - the Tamil participant made use of morphological decomposition, while the Hindi participant seemed to rely on previous lexical knowledge - they first identified a previously known word which they felt would fit the sentence best, then applied that suffix onto the options given to find their answer.

In the affix word choice task, the participant was presented with an incomplete sentence, along with four real word options and participants were required to choose the option which they felt would complete the sentence the best. In this task, the Hindi participants (Mean = 18.48) outperformed the Tamil participants (Mean = 18.13), though this difference was insignificant. In Wu and Juffs (2022), the affix word choice task was the only task in which the differences between the groups became insignificant after accounting for LexTALE scores. This would suggest that the affix word choice task is more dependent on English proficiency rather than morphological awareness. Hence, since there was no significant difference between the English proficiency of Hindi and Tamil participants, there were no significant differences in the scores of these two groups on the affix word choice task.

In the morphological relatedness tasks, participants were asked to indicate whether a pair of given words were related or not. In this task, the Tamil participants (Mean = 32.81) scored higher than the Hindi participants (Mean = 32.78), however this difference was insignificant. Interestingly, this result goes against the findings of Wu and Juffs (2020), where the Turkish group significantly outperformed both the Chinese and English group in the morphological relatedness task (after accounting for LexTALE scores), suggesting that an L1 with complex morphology would contribute to a greater awareness of morphological relations between L2 words. However, in this study, there was no significant difference in the performance of both groups, which would suggest that rather than relying on morphological awareness, participants may have relied on lexical knowledge to answer these questions. This could especially be true in the case of this study as most of the participants were at CEFR level C1 and above.

4.4.1.2. What is the impact of the morphological complexity of L1 on word reading speed in English?

To answer this question, the data from the word identification and word attack tasks were analysed for between-group differences. The data for this task consisted of reaction times (measured in milliseconds) collected through Gorilla. Kolmogorov-Smirnov and Levene's test

revealed that data did not meet the assumptions for independent-samples t-tests, so Mann-Whitney U tests were conducted instead. The results of these tests are reported in table 4.16.

4.16 Results of Mann-Whitney test for word reading tasks			
	Mann-Whitney U	**Z-Score**	**Significance (2-tailed)**
Word Identification	203.00	-7.128	<.001
Word Attack	119.00	-7.717	<.001

In both tasks, there were significant differences between both groups. In the word identification task, participants were presented with real words on screen and were asked to read them out loud. In this task, the Tamil participants (Mean = 1490.675) took less time to read the words in comparison to the Hindi participants (Mean = 3220.402). The reported effect size was r = - 0.7163, which can be categorized as a large effect (Field, 2019). In the word attack task, participants were presented with pseudowords onscreen and were asked to read them out loud. In this task too, the Tamil participants (Mean = 1199.262) took less time to read the words in comparison to the Hindi participants (Mean = 3415.957). The reported effect size was r = - 0.7755, which can be categorized as a large effect (Field, 2019).

The findings outlined above partially coincide with previous research by Carlisle (2000), De Freitas et al. (2018), Deacon & Kirby (2004) and Marinova-Todd et al. (2013). A common finding of all these studies is that to a certain extent, morphological awareness and L1 morphological complexity contribute to faster word reading. However, the impact was greater for pseudoword reading and reading comprehension rather than real word reading. The main reasoning for this impact was that greater morphological awareness would facilitate the occurrence of word decomposition, which would then lead to the ability to access words faster, thereby leading to faster word reading (Portin et al., 2008). It can be assumed that Tamil speakers made more use of decomposition, because Tamil has a larger inventory of morphological affixes compared to Hindi, which would allow them to read both the real words and pseudowords faster.

This was also illustrated by the participants in the stimulated recalls. When asked about how they read the test items, the Tamil participant (T1) highlighted the fact that when they were presented with the word onscreen, they split it into multiple parts to facilitate reading. Meanwhile, the Hindi participant (H1) drew upon previous lexical knowledge, using the familiar words as the basis to split words into constituent parts, to facilitate word reading.

4.17 Quotations from interview, detailing usage of morphological decomposition to assist in word reading	
T1	**H1**
Test item: Straced	
"First, what I think the word that came to me was starraised because I saw it as sta and then my mind came up with star…So I was like star-raised. So I saw it as two different words"	"It's S being silent and just traced or just raised - Just traced"

Table 4.17 suggests that, to some extent, the Tamil participants made use of decomposition, which allows for faster word reading, in contrast to whole-word retrieval, which was one of the main strategies used by the Hindi participants.

4.4.2. RQ2: What is the impact of English proficiency on performance on tasks in English?

To understand whether English proficiency had any impact on the four morphological awareness and two word reading tasks, a two-way ANOVA was run, with the L1 group and English CEFR level as the independent variables and the respective morphological awareness and word reading task scores as the dependent variable. A two-way ANOVA was chosen because, despite the data not meeting assumptions for homogeneity and normality, the ANOVA is considered fairly robust in terms of error rate when the sample sizes are relatively equal (Field, 2019).

4.4.2.1. To what extent does English proficiency impact performance on morphological awareness tasks?

The results of the ANOVA (with morphological awareness task scores as the dependent variable, and L1 and CEFR levels as the independent variables) are reported in table 4.18.

4.18 Results of Two Way ANOVA on Morphological Awareness : First Language and LexTALE					
Tasks	**Variables and Interactions**	**df**	**Mean Score**	**F**	**Significance**
Derivation task	Group	1	175.82	35.726	<.001
	English CEFR level	3	31.672	6.430	<.001
	Group*English CEFR level	3	27.92	5.672	<.001
Affix Word Choice	Group	1	11.357	0.966	.328
	English CEFR level	3	73.790	6.276	.171
	Group*English CEFR level	3	16.470	1.401	.044
Affix Non-word Choice	Group	1	40.633	31.855	<.001
	English CEFR level	3	8.122	1.899	.135
	Group*English CEFR level	3	21.398	0.380	.768
Morphological Relatedness	Group	1	80.245	0.27	.871
	English CEFR level	3	91.275	3.250	.025
	Group*English CEFR level	3	24.689	3.697	.015

In the derivation task, there was a significant main effect of English proficiency on performance in the derivation task, as evidenced in Table 4.18, with an effect size of $\eta^2 = 0.282$, which can be categorized as a large effect size (Cohen, 1969). Bonferroni *post hoc* tests revealed that the effect was significant between levels B1 and C2 (p = .018), B2 and C1 (p = .003), B2 and C2 (p = <.001). However, there was no significant effect between levels B1 and B2 (1.000) and C1 and C2 (1.000). A significant main effect was also observed for L1 groups (as seen in table 4.18), with an effect size of $\eta^2 = 0.175$, which can be understood as a larger effect size (Cohen, 1969). Simple effects analysis revealed that the effect of English proficiency was greater for the Tamil group (p = <.001) than the Hindi group (p = .461). Finally, as evidenced in table 4.18, there was a significant interaction between English proficiency and L1 group on derivation task score (p < .001) with an effect size of $\eta^2 = 0.158$, which can be interpreted as a large effect size (Cohen, 1969). Pairwise comparisons revealed that the effect of the L1 group was significantly different at CEFR levels B1 (p = .001), B2 (p = .001) and C1 (p = <.001). These results indicate that at higher levels of English proficiency, performance on the derivation task was better. However, this difference was not as prominent at the highest level of proficiency (C2). The impact of English proficiency was also more significant for the Tamil group, which would indicate that the Tamil participants rely more on lexical knowledge rather than the Hindi population. A possible explanation for this would be the lack of derivational morphology in Tamil, which would lead to a greater reliance on lexical knowledge rather than morphological awareness.

In the affix word choice task, there was a significant interaction between English proficiency and L1 group on task performance (table 4.18), with an effect size of $\eta^2 = 0.044$, which can be interpreted as a small effect size (Cohen, 1969). Pairwise comparisons showed that there was an insignificant effect across groups and across different proficiency levels. Similarly, the main effect of the L1 group was insignificant ($p = .328$). However, a significant main effect of English proficiency on task performance was found ($p = <.001$), with an effect size of $\eta^2 = 0.171$, which is a large effect size (Cohen, 1969). Bonferroni *post hoc* test showed that regardless of L1, there was a significant effect between levels B1 and C1 ($p = .004$), B1 and C2 ($p = .006$), B2 and C1 ($p = .013$) and B2 and C2 ($p = .024$). This would mean that regardless of L1 group, English proficiency had a significant impact on task performance in the affix word choice task. Again, this is reflective of the findings of Wu and Juffs (2022), who found that in this task, the impact of English proficiency was far larger than the impact of L1 morphology.

In the affix non-word choice task, there was no significant interaction between English proficiency and L1 group (table 4.18). The main effect of English proficiency on task performance was also insignificant ($p = .135$). However, there was a significant main effect of the L1 group on task performance: $p <.001$, $\eta^2 = 0.259$ (large effect size). Simple effect analysis showed that the impact of L1 group was greater for Tamil ($p = .264$) than Hindi ($p = .411$), however neither reached significance. This finding shows that rather than English proficiency, it was L1 morphological complexity which had a significant impact on task performance. This could be reflective of the fact that the affix non-word choice task is the only morphological awareness task that does not rely on whole word knowledge – it did not require the knowledge of the stem, only the knowledge of English suffixation (Wu & Juffs, 2022). Hence, participants with L1s that have a complex morphological system would have more familiarity with the suffixation process and would therefore perform better on this task.

In the morphological relatedness task, there was a significant interaction between English proficiency and L1 on task performance (table 4.18), with an effect size of $\eta^2 = 0.109$, which can be interpreted as a large effect size (Cohen, 1969). Pairwise comparisons showed that the effect of the L1 group was significantly different at levels B2 ($p = .048$) and C2 ($p = .012$). A significant main effect was identified for English proficiency (table 4.18), with an effect size of $\eta^2 = 0.097$, which is a medium effect size (Cohen, 1969). Bonferroni *post hoc* tests revealed that there was a significant difference in the task scores between B1 and C2 ($p = .013$). Finally, an insignificant main effect was found for L1 group (table 4.18). This reveals that English

proficiency had a significant impact on task performance, but this impact was evident only at the B2 and C2 levels.

The results of this section show that, to some extent, English proficiency does impact morphological awareness task performance across groups. In three of the tasks, there was a significant interaction between L1 group and English proficiency on task performance.

Past research in the field of morphological awareness has not necessarily focused on understanding the role of English proficiency in developing morphological awareness - instead, English proficiency has been a background variable to be controlled. The results of this study reveal some interesting patterns about the role of English proficiency in the tasks of morphological awareness. In the derivation, affix word choice and morphological tasks, English proficiency had a significant impact on performance. This could be attributed to the fact that these tasks are more lexically demanding. Interestingly, in the derivation task, the impact of English proficiency was greater for the Tamil group, which could mean that instead of relying on the process of derivation, the Tamil participants could have been relying on whole word knowledge to answer the test items, as seen in Wu and Juffs (2022). The only task in which English proficiency did not have an impact was the affix non-word choice, which could be attributed to the non-lexical nature of the task itself.

4.4.2.2. To what extent does English proficiency impact performance on word reading tasks?

A two-way ANOVA was conducted to understand the impact of English proficiency and Language group on the performance in the two word reading tasks, The results are reported in table 4.19.

4.19 Results of Two Way ANOVA on Word Reading: First Language and LexTALE

Tasks	Variables and Interactions	df	Mean Score	F	Significance
Word Identification	Group	1	43710651.4	38.178	<.001
	English CEFR level	3	827462.702	0.723	.541
	Group*English CEFR level	3	1790254.451	1.564	.204
Word Attack	Group	1	68629881.1	23.909	<.001
	English CEFR level	3	2212625.767	0.771	.513
	Group*English CEFR level	3	2020080.647	0.704	.552

In both the word identification and word attack tasks, there was no significant interaction between English proficiency and L1 group on task performance. In both tasks, only the group showed a significant main effect on the task performance (p = <.001). In the word identification task, regardless of English proficiency, there was a significant effect of the L1 group on task performance (B1: p = .048, B2: p = .019, C1 and C2: p = <.001), with an effect size of $\eta^2 = 0.296$, which is a large effect size (Cohen, 1969). In the word attack task, there was a significant effect of the L1 group across all CEFR levels (B2: p = .18, C1: p = .001, C2: p = <.001) except for the B1 level (p = 0.182). The effect size was $\eta^2 = 0.208$, which is a large effect size (Cohen, 1969). As evidenced by the results of this section, English proficiency did not have a significant impact on word reading in English. This could be since neither the word identification nor the word attack task actively tested lexical knowledge – rather, they just requiring passive reading, which meant that participants may have relied on processes such as morphological breakdown, rather than other lexical processes.

4.4.3. RQ 3: What is the impact of L1 proficiency on performance on tasks in English?

To understand whether L1 proficiency had an impact on performance in the four morphological awareness and two word reading tasks, a correlation test was conducted. Kolmogorov-Smirnov and Levene's test revealed that data did not meet the assumptions for Pearson's correlation test, so Spearman's rank-order correlation was conducted instead.

4.4.3.1. To what extent does L1 proficiency impact performance on morphological awareness tasks?

To determine if there was any correlation between L1 proficiency and performance on the four morphological awareness tasks, a Spearman's correlation test was run, with L1 proficiency test scores and morphological awareness task scores as the variables. The results of this test are displayed in table 4.20.

4.20 Correlation between L1 proficiency and Morphological Awareness

Task	Spearman's rho (r_s)	Significance
Derivation	-.172	.089
Affix word choice	-.221	.028
Affix non-word choice	-.131	.197
Morphological relatedness	-.028	.782

From table 4.20, it is evident that there was no significant correlation between L1 proficiency and task performance in all the tasks except the affix word choice task. In the affix word choice task r_s = -.221, which indicates a moderate negative correlation between L1 proficiency and task performance (Rea & Parker, 2014). The results indicated the higher the L1 proficiency, the lower the participants scored on the affix word choice task. This could indicate that since affix word choice is dependent on English proficiency (as evidenced in the previous section), L1 could be an interfering factor in performance on this task – L1 knowledge could interfere in the process of whole word retrieval employed by the participants to answer the test items.

L1 proficiency has not been included as an independent variable in previous studies on morphological awareness – in fact, a limitation in most studies on morphological awareness in bilingual populations was that L1 proficiency remained unaccounted for. However, as seen in table 4.20, there was no significant correlation between L1 proficiency and task performance. One significant reason this could have occurred because most of the participants in this study had scored above C1 level, which would have led to a clustering of the results around one end of the spectrum, thereby leading to insignificant results.

4.4.3.2. To what extent does L1 proficiency impact performance on word reading tasks?

To determine if there was any correlation between L1 proficiency and performance on the two word reading tasks, a Spearman's correlation test was run, with L1 proficiency test scores and word reading reaction times as the variables. The results of this test are displayed in table 4.21.

4.21 Correlation between L1 proficiency and word reading

Task	Spearman's rho (r_s)	Significance
Word identification	.063	.533
Word attack	.129	.204

As evidenced by table 4.21, there were no significant correlations between L1 proficiency and word reading performance. This could again be attributed to the fact that most of the participants in this study had above a C1 level of proficiency, which meant that the data would have been clustered around one end of the spectrum. Therefore, no significant correlations would have been able to be detected.

4.4.4. RQ 4: To what extent does the L1 of the participant impact how they approach the provided tasks?

The stimulated recalls provided the participants with the opportunity to expand upon their experience with the affix non-word task and word attack task. These responses were then coded, and three major strategies were identified (common to both tasks), which provide some insights into the performance of different groups of learners. The strategies have been outlined in table 4.22.

4.22 Strategies used by participants to address test items

Strategies	Quote
Using morphological decomposition	Word Attack: "Baf-mot-bem – Ok this I split into 3." Affix non-word choice: "Yeah, I'm gonna go with B because when I said the meeting was highly -lism -lize -fy, -rial made the most sense."
Using prior lexical knowledge in English	Word Attack: "straced - the stressed word here- straced, like it's S being silent and just traced or just raised" Affix non-word choice: "I would say option B. So it's like loquarial, … if you're describing someone, let's say that this person was highly. mercurial and invigorating or was high, was highly mercurial and entertaining"
Checking to see if a word feels right	Word attack: Not applicable Affix non-word choice: "'Im going with what like my gut says - OK, I think this is the tense that I am choosing"

Once the strategies were identified, an effort was made to count the number of times each strategy was used by each participant. Since there were some cases where participants reported using two strategies, each time a strategy was used, it was measured (regardless of whether multiple strategies were used for the same item). The quantised results are displayed in table 4.23.

4.23 Strategy use among stimulated recall participants

Strategy	Tamil (n=1)	Hindi (n=1)

Using morphological decomposition	19	4
Using prior lexical knowledge in English	10	17
Checking to see if a word feels right	3	1
No particular strategy	0	3

As evidenced by the table above, the Tamil participant (T1) made use of morphological decomposition to a greater extent compared to the Hindi participant (H1), both in the morphological awareness and the word reading tasks. However, for some of the words in the word attack task (especially the shorter ones), T1 used two strategies together to decode the word. An example of this is the test item 'Zirdn't' – the strategies are outlined in table 4.24.

4.24: Strategies used by T1 in word attack task	
Test item: Zirdn't	
Strategy	**Quote**
Using morphological decomposition	"See this again, same splitting method…ZI and then the apostrophe thing"
Using prior lexical knowledge	"It reminds me of the word 'didn't' or whenever we use the apostrophe"

T1 also used this mix of strategies in some items in the affix non-word choice task as well. An example has been highlighted in table 4.25.

4.25: Strategies used by T1 in affix non-word choice task	
Test item: Please try to be as totally _____ as possible Options: progenalism, progenalize, progenious, progenify,	
Strategy	**Quote**
Using morphological decomposition	"I saw the OUS sound again"
Using prior lexical knowledge	"I saw it like a statement someone was selling to a particular crowd…if someone was instructing me something and they told please try to be as totally _____ as possible, then I'll be like, OK, they're saying this only"

In the word attack task, H1 reported that there was nothing that went through their mind during three of the items in the task. Apart from these, they exclusively made use of prior lexical knowledge to guide them to what they thought the right way to read the word was. An example has been provided in table 4.26.

4.26: Strategies used by H1 in word attack task	
Test item: Zirdn't	
Strategy	**Quote**
Using prior lexical knowledge	"Probably is very common word. We use it. Didn't, I mean very abbreviated version of did not."

In the affix non-word choice task, for four items, H1 did pay attention to suffixes, however this was supplemented by their prior lexical knowledge. They first drew upon their lexical knowledge, then used that knowledge to guide them in selecting the suffix that they thought fit the sentence well. An example has been provided in table 4.27.

4.27: Strategies used by H1 in affix non-word choice task	
Test item: Please try to be as totally _____ as possible Options: progenalism, progenalize, progenious, progenify,	
Strategy	**Quote**
Using prior lexical knowledge	"I would still say Option C…just that one sentence…try to keep it as continuous as possible' would sound right because that actually sounds grammatically correct - You can add like justifiable or reasonable."
Morphological decomposition	"words with -OUS … something with that would actually fit in well out of the four"

From the data above, it's clear that there indeed was an impact of L1 on how individuals approach the task items. The Tamil participants made use of morphological decomposition at a greater rate than the Hindi participants, who relied more on prior lexical knowledge in English. These findings are reflective of Carlisle (2000), De Freitas et al. (2018), Deacon & Kirby (2004), Havas et al., (2015) and Marinova-Todd et al. (2013). The findings of these studies showed that greater L1 morphological complexity led to heightened awareness of the morphemes in the L2, which led to heightened word decomposition. This would suggest that participants whose L1s were morphologically complex (especially in terms of morphological inventory size) were more inclined to perform morphological decomposition, especially in the case of complex words.

4.4.5. RQ 5: How do participants with different English proficiency levels approach the provided tasks?

Table 4.28 describes the proficiency level and self-ratings of the participants who took part in the stimulated recall.

4.28 Participant details - stimulated recall		
Item	**T1**	**H1**
LexTALE score	110	135
CEFR	B2	C1

As evidenced by the table 4.28, H1 had a greater English proficiency compared to T1. This difference in English proficiency could provide an explanation as to why H1 relied more on lexical knowledge to answer the items in the affix non-word choice. H1 was able to provide unique sentences to answer the test items. Some examples have been provided in table 4.29.

4.29 Examples of Lexical recall by H1 in affix non-word choice task	
Test item	**Quote**
The meeting was highly _____ and invigorating' (options: loquarify, loquarial loguarialize, loquarialism)	"it's like loquarial …. if you're describing someone, let's say that this person was highly. mercurial and invigorating"
All those models are strictly _____ and outdated as well (options:	"I'll phrase another sentence out for example…It's like a football tournament…just say the sentence 'all these matches were highly competitive and very interesting as well.'

Additionally, while both participants answered all the items correctly, when asked about the task difficulty, H1 found the task items easier compared to T1. Additionally, H1 was able read through three of the word attack items without the need to decompose them or draw upon previous knowledge. Therefore, the data suggests that English proficiency may have influenced task performance. The usage of whole word retrieval in the affix non-word choice task by H1 contrasts with the findings of Wu and Juffs (2022) – H1's strategy of using whole word retrieval is more reflective of their findings in the derivation and affix word choice task, which are two tasks that rely the most on lexical knowledge. However, H1 has modified this strategy to enable them to answer the affix non-word choice task. This could be reflective of the impact of the higher level of English proficiency on task strategy. However, greater qualitative data is required to establish a significant correlation as there were only two participants in this study, thereby the results of these stimulated recalls may not be representative of the whole population.

4.5. SUMMARY OF FINDINGS

To summarize, there were a number of trends that emerged from the data in this study, some of which fit into the findings of previous research in the field, while other findings emerged

that were previously unexplored in the research. In terms of the impact of L1 morphological complexity on morphological awareness in English, significant differences were identified only for the derivation task and the affix non-word choice task. In the derivation task, the Hindi group significantly outperformed the Tamil group, which could be attributed to the presence of productive derivation in Hindi in comparison to Tamil, in which derivation was a rarer process. Meanwhile, in the affix non-word choice, the Tamil group outperformed the Hindi group, which could be attributed to the large inventory size of affixes in Tamil in comparison to Hindi, which means that Tamil participants would have possibly had a greater sensitivity to morphemes, allowing them to choose the right answer for the task items - this is supplemented by the information from the stimulated recall, which showed that the Tamil participant mainly focused on the suffixes in order to decide the correct answer for the questions. In both word reading tasks, it was found that Tamil participants outperformed the Hindi participants, which could be attributed to the fact that Tamil participants had a greater sensitivity to intraword morphemes, which would have led to the decomposition of the morphologically complex words, thereby allowing for faster recall. It was also found that English proficiency had a significant impact on performance in the derivation, affix word and morphological relatedness task - which could be attributed to the lexical nature of the tasks and their reliance on whole-word knowledge. Since the tasks involved real words in the English language, lexical knowledge of English could have been more important in comparison to morphological awareness to answer the question. The impact of L1 proficiency was only significant for the real word reading task (only above the C1 level), which suggests that for L1 to contribute to faster word reading, participants need to have a certain level of proficiency in their L1.

CHAPTER 5: CONCLUSION

5.1. SUMMARY OF FINDINGS

This research was designed to explore a previously unstudied population in the literature on morphological awareness, its impact of reading and cross linguistic transfer: L1 Tamil and Hindi speakers studying English. The results of this study offer insights into the impact of L1 morphological complexity on morphological awareness and word reading in English, from two under-explored languages.

The results of the quantitative data highlight a significant difference in the performance of Tamil and Hindi groups in the derivation task, the affix non-word choice task, word identification and word attack. Only the derivation task showed the impact of cross-linguistic transfer of similar features across different languages. The other tasks shed light on the role of general L1 morphological complexity in contribution to morphological decomposition in English. Further quantitative analysis revealed that English proficiency significantly contributed to performance on three tasks of morphological awareness (derivation, affix word choice and morphological relatedness) and L1 proficiency only contributed to real word reading at the level of CEFR C1 and above. Quantitative data further demonstrated the difference in how participants from different language groups approached these tasks and helped to elucidate a few reasons for the differing performance between the two groups.

5.2. LIMITATIONS

In addition to the methodological limitations highlighted in Chapter 3, there are other limitations that may impact the generalizability of the results of this study to a larger population. The first major limitation lies in the small sample size of the study, especially in the stimulated recall (SR). While the greatest attempt was made to ensure that the participants were equally distributed among the two language groups, since only two participants were interested in participating in the SR, participants could not be matched on English and L1 proficiency. While the participants were slightly mismatched in terms of their English proficiency, care has been taken during

qualitative analysis to account for the difference in the proficiency levels to not conflate differences between the two participants. In future research, efforts must be taken to recruit a larger population for a SR, to ensure that participants can be matched for proficiency (English and L1) and to provide richer, more robust data.

An additional limitation arises when the simplicity of some of the test items is considered. While previous research (e.g. Wu & Juffs, 2022) has suggested that the tests of morphological awareness adopted in this study are suitable for advanced learners of English, this research suggests that a few of these tests may be too simplistic to accurately measure differences between groups of advanced learners. This is especially true for the affix word choice and the morphological relatedness tasks, where participants would have been able to recognize the correct answer by sight. Future research would benefit from tests of greater complexity to accurately measure the contribution of L1 morphological complexity on task performance. Usage of pseudowords as test items has allowed for the testing of morphological awareness rather than lexical knowledge and creating a test of morphological relatedness using only pseudowords (or infrequent English words) would allow us to effectively test for knowledge of relational morphology without the confounding effect of English lexical knowledge.

Finally, in this study, LexTALE was used as a measure of English proficiency owing to its ease of administration. However, as the study shows, most of the participants scored at a C1 level and above. There is a possibility that LexTALE was not sensitive enough to accurately measure the proficiency level of advanced learners of English - future research might benefit from using a test of proficiency that would be able to highlight subtle differences among the different levels of proficiency, thereby avoiding the possibility of overestimating participants' English proficiency.

5.3. IMPLICATIONS FOR RESEARCH AND PEDAGOGY

Despite the above-mentioned limitations, this study sheds light onto previously under-researched languages in applied linguistics, with significant implications for pedagogy and research, especially in the field of English language learning.

This research was conducted with an under-represented group in literature on morphological awareness - adult, advanced learners of English, with alphasyllabary L1s, belonging to a language context where knowledge of English is a vital commodity. This research has important implications for future research and pedagogy in the Indian language learning context. Existing literature has shown that morphological awareness in English is an important skill for reading comprehension and word reading. This study shows that the morphological complexity of an L1 has an impact on morphological awareness in English and subsequently, on word reading as well. The findings of this study could potentially be generalized to a larger number of languages such as Telugu, Malayalam, Marathi, Bengali etc. owing to their similarity to the languages in this study, though further empirical exploration of this is needed.

The findings of this study suggest that since morphological awareness contributes to word reading in English, explicit morphological instruction would be an effective method of English language teaching, especially for older English language learners. This fits into other studies which explore the impact of explicit morphological instruction on word reading and vocabulary learning (Amirjalili & Jabbari, 2018; Anwar & Rosa, 2020; Bowers et al., 2010; Brady & Mason, 2023; Jiang & Kuo, 2019).

Finally, this study also shed some light on the current scenario of English language learning in India. Although language use in India was briefly touched upon, the findings of this study slightly differed from the existing literature on the subject, shedding light on the outdated research on the topic. Previous studies (Parasher, 1980; Sahgal, 1991) overestimated the dominance of English at the school level in India – however, the participants in their studies belonged to the elite sectors of Indian society. In this study, participants came from various walks of life, which meant that results varied, showing the prevalence of regional language, even in the spheres of education (except for university). Further research is required to fully capture the existing language landscape existing in India today, to benefit other research on language education in India.

REFERENCES

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388–407. https://doi.org/10.3758/s13428-019-01237-x

Anwyl-Irvine, A., Dalmaijer, E. S., Hodges, N., & Evershed, J. K. (2021). Realistic precision and accuracy of online experiment platforms, web browsers, and devices. *Behavior Research Methods*, *53*(4), 1407–1425. https://doi.org/10.3758/s13428-020-01501-5

Aitken, K. G. (1978). Constructing a Multiple-Choice Cloze Test of E.S.L. *RELC Journal*, *9*(2), 105–108. https://doi.org/10.1177/003368827800900208

Ajideh, P., & Mozaffarzadeh, S. (2012). C-test vs. Multiple-choice Cloze Test as Tests of Reading Comprehension in Iranian EFL Context: Learners' Perspective. *English Language Teaching*, *5*(11), Article 11. https://doi.org/10.5539/elt.v5n11p143

Alderson, J. C. (2005). *Assessing reading*. Cambridge University Press.

Amirjalili, F., & Jabbari, A. A. (2018). The impact of morphological instruction on morphological awareness and reading comprehension of EFL learners. *Cogent Education*, *5*(1), 1–30. https://doi.org/10.1080/2331186X.2018.1523975

Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. Guthrie (Ed.), *Comprehension and Teaching: Research Reviews* (pp. 77–117). International Reading Association.

Anwar, I., & Rosa, R. N. (2020). The Role Of Morphological Awareness And Explicit Morphological Instructions In ELT. *Language Literacy: Journal of Linguistics, Literature, and Language Teaching*, *4*(1), 28–37. https://doi.org/10.30743/ll.v4i1.1825

Bajre, P., & Khan, A. (2019). Developmental dyslexia in Hindi readers: Is consistent sound-symbol mapping an asset in reading? Evidence from phonological and visuospatial working memory. *Dyslexia*, *25*(4), 390–410. https://doi.org/10.1002/dys.1632

Bauer, L., Lieber, R., & Plag, I. (2013). English morphology in a typological perspective. In L. Bauer, R. Lieber, & I. Plag (Eds.), *The Oxford Reference Guide to English Morphology* (p. 0). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198747062.003.0028

Bowers, P. N., Kirby, J. R., & Deacon, S. H. (2010). The Effects of Morphological Instruction on Literacy Skills: A Systematic Review of the Literature. *Review of Educational Research*, *80*(2), 144–179. https://doi.org/10.3102/0034654309359353

Brady, S., & Mason, L. H. (2023). A Literature Review of Morphological Awareness Interventions and the Effects on Literacy Outcomes. *Learning Disability Quarterly*, 1–14. https://doi.org/10.1177/07319487231171388

Cardona, G., & Jain, D. (2003). *The Indo-Aryan languages* (1st ed.). Routledge. https://doi.org/10.4324/9780203945315

Carlisle, J. F. (1995). Morphological Awareness and Early Reading Achievement. In L. B. Feldman (Ed.), *Morphological Aspects of Language Processing* (pp. 189–209). Erlbaum.

lCarlisle, J. F. (2000). Awareness of the structure and meaning of morphologically complex words: Impact on reading. *Reading and Writing*, *12*, 169–190.

Clahsen, H., Sonnenstuhl, I., & Blevins, J. P. (2003). Derivational morphology in the German mental lexicon: A dual mechanism account. In R. H. Baayen & R. Schreuder (Eds.), *Morphological structure in language processing* (pp. 125–156). De Gruyter Mouton. https://doi.org/10.1515/9783110910186.125

Cohen, J. (1969). *Statistical Power Analysis for the Behavioural Sciences*. Academic Press.

Comrie, B. (1989). *Language Universals and Linguistic Typology: Syntax and Morphology* (2nd ed.). University of Chicago Press. https://press.uchicago.edu/ucp/books/book/chicago/L/bo24426144.html

De Freitas, P. V., Mota, M. M. P. E. D., & Deacon, S. H. (2018). Morphological awareness, word reading, and reading comprehension in Portuguese. *Applied Psycholinguistics*, *39*(3), 507–525. https://doi.org/10.1017/S0142716417000479

Deacon, S. H., & Kirby, J. R. (2004). Morphological awareness: Just "more phonological"? The roles of morphological and phonological awareness in reading development. *Applied Psycholinguistics*, *25*(2), 223–238. https://doi.org/10.1017/S0142716404001110

Ehri, L. C. (1995). Phases of development in learning to read words by sight. *Journal of Research in Reading*, *18*(2), 116–125. https://doi.org/10.1111/j.1467-9817.1995.tb00077.x

Ehri, L. C. (2005). Learning to Read Words: Theory, Findings, and Issues. *Scientific Studies of Reading*, *9*(2), 167–188. https://doi.org/10.1207/s1532799xssr0902_4

Farley, A. P., & McCollam, K. (2004). Learner Readiness and L2 Production in Spanish: Processability Theory on Trial. *Estudios de Lingüística Aplicada*, *40*, 47–69.

Field, A. (2019). *Discovering Statistics using IBM SPSS Statistics* (4th ed.). SAGE Publications Ltd.

Fumero, K., & Tibi, S. (2020). The Importance of Morphological Awareness in Bilingual Language and Literacy Skills: Clinical Implications for Speech-Language Pathologists. *Language, Speech, and Hearing Services in Schools*, *51*(3), 572–588. https://doi.org/10.1044/2020_LSHSS-20-00027

Hale, G. A., Stansfield, C. W., Rock, D. A., Hicks, M. M., Butler, F. A., & Oller Jr., J. W. (1988). Multiple-Choice Cloze Items and the Test of English as a Foreign Language. *ETS Research Report Series*, *1988*(1), i–117. https://doi.org/10.1002/j.2330-8516.1988.tb00258.x

Havas, V., Waris, O., Vaquero, L., Rodríguez-Fornells, A., & Laine, M. (2015). Morphological learning in a novel language: A cross-language comparison. *Quarterly Journal of Experimental Psychology*, *68*(7), 1426–1441. https://doi.org/10.1080/17470218.2014.983531

Hornberger, N., & Vaish, V. (2009). Multilingual language policy and school linguistic practice: Globalization and English-language teaching in India, Singapore and South Africa. *Compare: A Journal of Comparative and International Education*, *39*(3), 305–320. https://doi.org/10.1080/03057920802469663

Jayaram, N. (1993). The language question in higher education: Trends and issues. *Higher Education*, *26*(1), 93–114. https://doi.org/10.1007/BF01575108

Jiang, Y.-L. B., & Kuo, L.-J. (2019). The development of vocabulary and morphological awareness: A longitudinal study with college EFL students. *Applied Psycholinguistics*, *40*(4), 877–903. https://doi.org/10.1017/S014271641900002X

Joshi, V. (2022). *Census of India: Language Atlas of India 2011*. Office of the Registrar General & Census Commissioner, India.

Kachru, Y., & Bhatia, T. K. (2018). Hindi-Urdu. In *The World's Major Languages* (3rd ed., pp. 409–426). Routledge.

Kieffer, M. J., & Lesaux, N. K. (2012). Direct and Indirect Roles of Morphological Awareness in the English Reading Comprehension of Native English, Spanish, Filipino, and Vietnamese Speakers. *Language Learning*, *62*(4), 1170–1204. https://doi.org/10.1111/j.1467-9922.2012.00722.x

Kirby, J. R., Deacon, S. H., Bowers, P. N., Izenberg, L., Wade-Woolley, L., & Parrila, R. (2012). Children's morphological awareness and reading ability. *Reading and Writing*, *25*(2), 389–410. https://doi.org/10.1007/s11145-010-9276-5

Koda, K. (2000). Cross-linguistic variations in L2 morphological awareness. *Applied Psycholinguistics*, *21*(3), 297–320. https://doi.org/10.1017/S0142716400003015

Koda, K., Takahashi, E., & Fender, M. (1998). Effects of L1 processing experience on L2 morphological awareness. *Ilha Do Desterro A Journal of English Language, Literatures in English and Cultural Studies*, *35*, 57–85.

Krishnamurti, B. (Ed.). (2003). Introduction. In *The Dravidian Languages* (pp. 1–47). Cambridge University Press. https://doi.org/10.1017/CBO9780511486876.003

Kuo, L.-J., & Anderson, R. C. (2006). Morphological Awareness and Learning to Read: A Cross-Language Perspective. *Educational Psychologist*, *41*(3), 161–180. https://doi.org/10.1207/s15326985ep4103_3

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, *44*(2), 325–343. https://doi.org/10.3758/s13428-011-0146-0

Li, P., Zhang, F., Tsai, E., & Puls, B. (2014). Language history questionnaire (LHQ 2.0): A new dynamic web-based research tool. *Bilingualism: Language and Cognition*, *17*(3), 673–680. https://doi.org/10.1017/S1366728913000606

Lyle, J. (2003). Stimulated Recall: A Report on Its Use in Naturalistic Research. *British Educational Research Journal*, *29*(6), 861–878.

Mahony, D. L. (1994). Using sensitivity to word structure to explain variance in high school and college level reading ability. *Reading and Writing*, *6*(1), 19–44. https://doi.org/10.1007/BF01027276

Mahony, D., Singson, M., & Mann, V. (2000). Reading ability and sensitivity to morphological relations. *Reading and Writing*, *12*(3), 191–218. https://doi.org/10.1023/A:1008136012492

Mallikadevi, N. (2023). The processes of nominal derivation in Tamil: A descriptive analysis. *Strad Research*, *10*(9), 275–287. https://doi.org/10.37896/sr10.9/030

Marinova-Todd, S. H., Siegel, L. S., & Mazabel, S. (2013). The Association Between Morphological Awareness and Literacy in English Language Learners From Different Language Backgrounds. *Topics in Language Disorders*, *33*(1), 93–107. https://doi.org/10.1097/TLD.0b013e318280f5d5

Miller, R. T. (2019). English Orthography and Reading. In *The TESOL Encyclopedia of English Language Teaching* (pp. 1–7). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118784235.eelt0461

Milton, J., & Alexiou, T. (2009). Vocabulary Size and the Common European Framework of Reference for Languages. In B. Richards, M. H. Daller, D. D. Malvern, P. Meara, J. Milton, & J. Treffers-Daller (Eds.), *Vocabulary Studies in First and Second Language Acquisition: The Interface Between Theory and Application* (pp. 194–211). Palgrave Macmillan UK. https://doi.org/10.1057/9780230242258_12

Milton, J. (2010). The development of vocabulary breadth across the CEFR levels. *EUROSLA MONOGRAPHS SERIES*, *1*, 211–232.

Mohanty, A. K. (2006). Multilingualism of the Unequals and Predicaments of Education in India: Mother Tongue or Other Tongue? In *Imagining Multilingual Schools* (pp. 262–283). Multilingual Matters. https://doi.org/10.21832/9781853598968-014

Nag, S., & Narayanan, B. (2019). Orthographic Knowledge, Reading and Spelling Development in Tamil: The First Three Years. In R. M. Joshi & C. McBride (Eds.), *Handbook of Literacy in Akshara Orthography* (Vol. 17, pp. 55–83). Springer International Publishing. https://doi.org/10.1007/978-3-030-05977-4_4

Nagy, W., & Townsend, D. (2012). Words as Tools: Learning Academic Vocabulary as Language Acquisition. *Reading Research Quarterly*, *47*(1), 91–108. https://doi.org/10.1002/RRQ.011

Panackal, A. A., & Heft, C. S. (1978). Cloze Technique and Multiple Choice Technique: Reliability and Validity. *Educational and Psychological Measurement*, *38*(4), 917–932. https://doi.org/10.1177/001316447803800411

Parasher, S. V. (1980). Mother Tongue-English Diglossia: A Case Study of Educated Indian Bilinguals' Language Use. *Anthropological Linguistics*, *22*(4), 151–162.

Parthiban, L. (n.d.). *Tamil Words Frequency Processing*. GitHub. https://github.com/aviiciii/tamil-word-frequency

Parthiban, L. (n.d.). *Tamil Words Frequency*. Kaggle.Com. https://www.kaggle.com/datasets/aviiciii/tamil-words-frequency

Porter, D. (1976). Modified Cloze Procedure: A More Valid Reading Comprehension Test. *English Language Teaching Journal*, *30*(2), 151–155.

Portin, M., Lehtonen, M., Harrer, G., Wande, E., Niemi, J., & Laine, M. (2008). L1 effects on the processing of inflected nouns in L2. *Acta Psychologica*, *128*(3), 452–465. https://doi.org/10.1016/j.actpsy.2007.07.003

Ramirez, G., Chen, X., Geva, E., & Kiefer, H. (2010). Morphological awareness in Spanish-speaking English language learners: Within and cross-language effects on word reading. *Reading and Writing*, *23*(3–4), 337–358. https://doi.org/10.1007/s11145-009-9203-9

Ramirez, G., Chen, X., Geva, E., & Luo, Y. (2011). Morphological awareness and word reading in English language learners: Evidence from Spanish- and Chinese-speaking children. *Applied Psycholinguistics*, *32*(3), 601–618. https://doi.org/10.1017/S0142716411000233

Rastogi, M., & Khanna, P. (2014). Development of Morphological Analyzer for Hindi. *International Journal of Computer Applications*, *95*(17), 1–5. https://doi.org/10.5120/16683-6799

Rea, L. M., & Parker, R. A. (2014). *Designing and Conducting Survey Research: A Comprehensive Guide* (4th ed.). Wiley Publications. https://www.wiley.com/en-

us/Designing+and+Conducting+Survey+Research%3A+A+Comprehensive+Guide%2C+4th+Edition-p-9781118767030

Rimzhim, A., Johri, A., Kelty-Stephen, D. G., & Fowler, C. A. (2021). Transposition Effects in an Aksharic Writing System: The Case of Hindi. *Language and Speech*, *64*(4), 804–838. https://doi.org/10.1177/0023830920971315

Roman, A. A., Kirby, J. R., Parrila, R. K., Wade-Woolley, L., & Deacon, S. H. (2009). Toward a comprehensive view of the skills involved in word reading in Grades 4, 6, and 8. *Journal of Experimental Child Psychology*, *102*(1), 96–113. https://doi.org/10.1016/j.jecp.2008.01.004

Sahgal, A. (1991). Patterns of language use in a bilingual setting in India. In J. Cheshire (Ed.), *English around the World: Sociolinguistic Perspectives* (pp. 299–307). Cambridge University Press. https://doi.org/10.1017/CBO9780511611889.021

Sandra, D. (1994). The morphology of the mental lexicon: Internal word structure viewed from a psycholinguistic perspective. *Language and Cognitive Processes*, *9*(3), 227–269. https://doi.org/10.1080/01690969408402119

Sapir, E. (1921). *Language: An Introduction to the Study of Speech*. Harcourt Brace.

Schiff, R., & Calif, S. (2007). Role of Phonological and Morphological Awareness in L2 Oral Word Reading. *Language Learning*, *57*(2), 271–298. https://doi.org/10.1111/j.1467-9922.2007.00409.x

Scholz, G. E., & Scholz, C. M. (1981). *Multiple-Choice Cloze Tests of EST Discourse: An Exploration*. 15th Annual TESOL Conference, Detroit, Michigan. https://files.eric.ed.gov/fulltext/ED208656.pdf

Scoring. (n.d.). LexTALE. Retrieved 25 June 2024, from https://www.lextale.com/scoring.html

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, *55*(2), 151–218; discussion 219-226. https://doi.org/10.1016/0010-0277(94)00645-2

Singh, S., & Sarma M., V. (2010). Hindi Noun Inflection and Distributed Morphology. In S. Müller (Ed.), *Proceedings of the HPSG10 Conference* (pp. 307–321). CSLI Publications.

Singson, M., Mahony, D., & Mann, V. (2000). The relation between reading ability and morphological skills: Evidence from derivational suffixes. *Reading and Writing*, *12*(3), 219–252. https://doi.org/10.1023/A:1008196330239

Speer, R. (n.d.). *Wordfreq*. GitHub. https://github.com/rspeer/wordfreq

Speer, R. (n.d.). *Exquisite-corpus*. GitHub. https://github.com/LuminosoInsight/exquisite-corpus

Strauss, A., & Corbin, J. (1998). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. Sage Publications, Inc.

Timing in Gorilla. (n.d.). Gorilla Support. https://support.gorilla.sc/support/troubleshooting-and-technical/timing#measuringreactiontimes

Wu, Z., & Juffs, A. (2022). Effects of L1 morphological type on L2 morphological awareness. *Second Language Research*, *38*(4), 787–812. https://doi.org/10.1177/0267658321996417

APPENDICES

Appendix A – CUREC Approval Application

UNIVERSITY OF OXFORD
DEPARTMENT OF EDUCATION
15 Norham Gardens, Oxford OX2 6PY
Tel:+44(0)1865 274024
general.enquiries@education.ox.ac.uk
www.education.ox.ac.uk
Director Professor Victoria Murphy

1st March 2024

Dear Janani,

RESEARCH ETHICS APPROVAL

Research title: The Impact of L1 Tamil Morphology on Morphological Awareness and Word Reading in English as an L2

Research ethics reference: EDUC_C1A_24_086

The above research has been considered on behalf of the Department of Education Departmental Research Ethics Committee (DREC) in accordance with the University's procedures for ethical approval of all research involving human participants.

I am pleased to confirm that, on the basis of the information provided to the DREC, ethics approval has now been granted for this research. Please note the following:

Reference: Please ensure that you **use this ethics reference number** on all your own consent processes with participants.

Personal data: It is the responsibility of the PI to ensure that all personal data collected during the project is managed in accordance with the University's guidance and legal requirements.

In-person activities: Any data collection involving in-person interactions with participants must have an up-to-date fieldwork risk assessment in place; further guidance is available from the Safety Office's website.

Amendments: Please notify the committee if you intend to make any amendments to the information in your ethics application as submitted at date of this approval, as all changes must receive ethical approval prior to implementation. The amendment form is available on the SSH IDREC webpage.

We welcome feedback on your experience of the ethical review process and suggestions for improvement. Please email any comments to staff.curec@education.ox.ac.uk / student.curec@education.ox.ac.uk or ethics@socsci.ox.ac.uk.

Yours sincerely

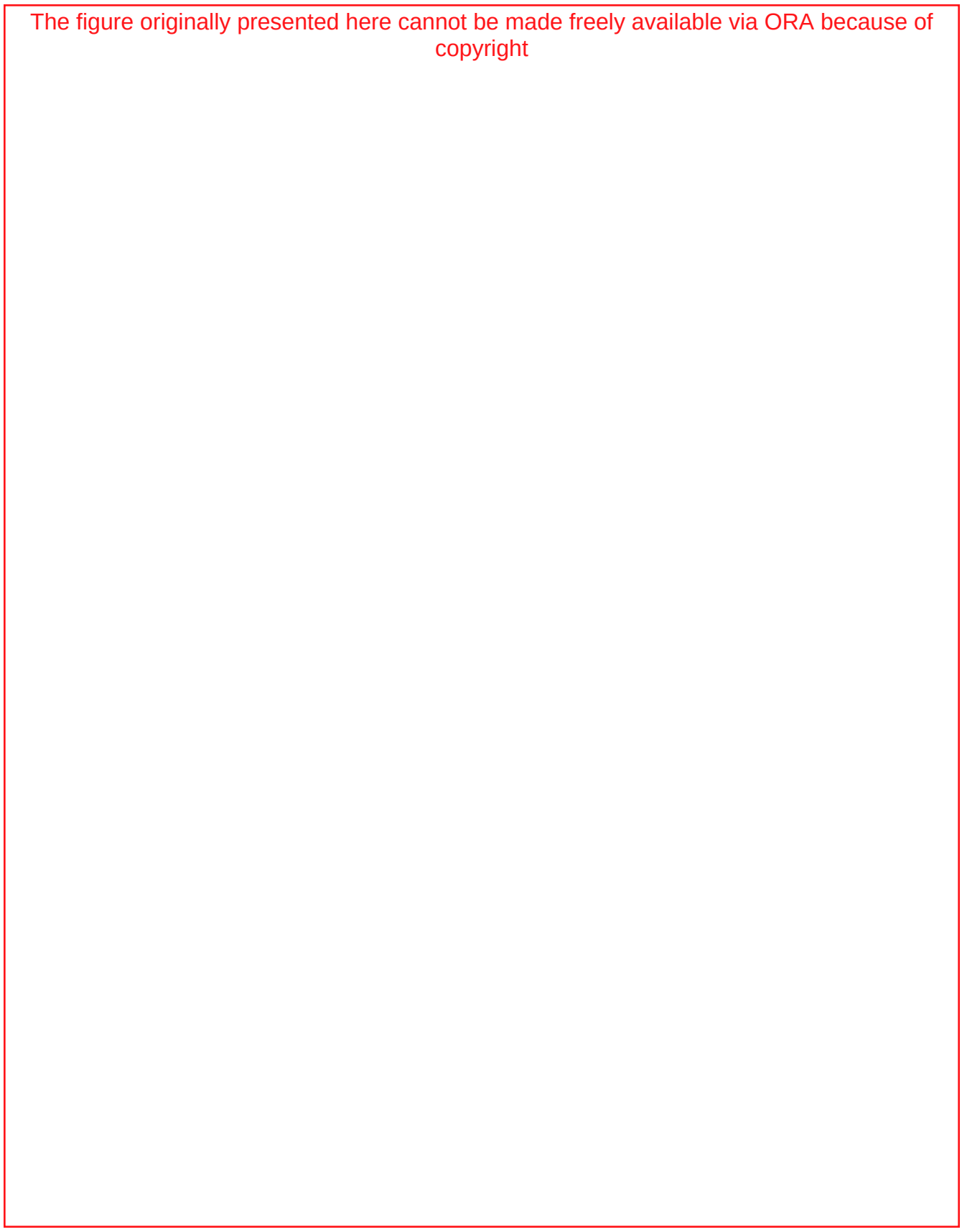
Dr Nigel Fancourt,

Associate Professor of Education and Values

DREC member

cc: student.curec@education.ox.ac.uk anna-maria.ramezanzadeh@education.ox.ac.uk

Appendix B – Recruitment Poster



Appendix C – Participant Information Sheet – Task 1

Janani Balaji
MSc. Student
Department of Education, University of Oxford,
Oxford, UK, OX2 6PY
Oxford University email address: janani.balaji@education.ox.ac.uk

The Impact of L1 on Morphological Awareness and Word Reading in English as an L2

PARTICIPANT INFORMATION SHEET

Central University Research Ethics Committee Approval Reference: EDUC_C1A_24_086

1. Introductory paragraph

You are being invited to take part in a research project. Before you decide to participate, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. If there is anything that is not clear or if you would like more information, please ask the researcher. Take time to decide whether you wish to take part.

2. Why is this research being conducted?

This research is being conducted as a part of my master's dissertation. The goal of this research is to understand how a person's first language impacts how they understand the structure of English words and also how fast they are able to read in English. I chose to focus on two Indian languages: Tamil and Hindi, because they are under-researched in the field of Applied Linguistics.

3. Why have I been invited to take part?

You have been invited to take part in this study because you are a university student studying in an Indian university, aged between 18 to 25 years, who is fluent in Tamil or Hindi (including reading and writing).

4. Do I have to take part?

No. It is up to you to decide whether to take part. You can withdraw yourself from the research without giving a reason by contacting me using the details provided. The deadline by which you can withdraw any information you have contributed to the research is 25th July 2024 (25/07/2024). After this time, the data will have been anonymised. If you decide to withdraw your participation before this date, any collected from you will be permanently deleted.

5. What will happen to me if I take part in the research?

This research will be conducted entirely online and it will be conducted in two parts. Prior to beginning the tasks, you will be presented with a consent form for you to go through. You should proceed with the study only after you have given consent.

The first part of the research involves an online questionnaire. You will be asked to answer questions about your language background – the questions will focus on language use in your everyday life. Then, you will complete a proficiency test in your first language and in English. Next, you will be asked to complete a few basic grammar tasks. These will be simple tasks which involve activities like reading individual words, sentence completion, deciding whether two words are related, and deciding if a given English word is an actual word.

If you choose to volunteer for the second task, you will be undertaking a short one-on-one interview with the interviewer (on Microsoft Teams), where you will be asked about how you went about performing the tasks in Part 1.

6. What are the possible disadvantages and risks in taking part?

There are no foreseeable disadvantages or risks associated with taking part in this study. The data obtained from your participation will be anonymised – identifiers like your name will be removed. You will not be directly identifiable from the data.

7. Are there any benefits in taking part?

While there are no immediate benefits for individual participants in this research, it is hoped that this research will lead to the development of effective methods for teaching English as a Second Language in India.

8. What information will be collected and why is the collection of this information relevant for achieving the research objectives?

I am interested in understanding whether your first language impacts how you understand the structure of English words. To answer this question, the following data will be collected from you: consent forms, language background details, test results and interview recordings.

The consent forms, test results, language background details and transcripts of interview recordings will be stored in a secure, encrypted OneDrive for Business account connected with the University of Oxford. The video recordings will be permanently deleted after transcription is complete. The collected data will be stored for 3 years after publication of the work following which it will be permanently deleted. Your data will not be shared with any third parties.

For the duration of the research, only my supervisor (Dr. Anna-Maria Ramezanzadeh) and I will have access to the data.

9. Will the research be published? Could I be identified from any publications or other research outputs?

The findings from this research will be written up as a part of a master's dissertation submitted to the University of Oxford. A copy of my dissertation will be deposited both in print and online in the Oxford University Research Archive, where it will be publicly available to facilitate its use in future research.

You will not be directly identifiable from the output of the research. I would also like your permission to use direct quotations without identifying you in any research outputs. You can also choose to opt out of this if you wish to do so.

10. Data Protection

The University of Oxford is the data controller with respect to your personal data, and as such will determine how your personal data is used in the research. The University will process your personal data for the purpose of the research outlined above. Research is a task that is performed in the public interest. Further information about your rights with respect to your personal data is available from the University's Information Compliance website at https://compliance.admin.ox.ac.uk/individual-rights.

11. Who has reviewed this research?

This research has received ethics approval from a subcommittee of the University of Oxford Central University Research Ethics Committee. (Ethics reference**: EDUC_C1A_24_086)**.

12. Who do I contact if I have a concern about the research or I wish to complain?

If you have a concern about any aspect of this research, please contact Janani Balaji at janani.balaji@education.ox.ac.uk or Dr Anna-Maria Ramezanzadeh at anna-maria.ramezanzadeh@education.ox.ac.uk, and we will do our best to answer your query. I will acknowledge your concern within 10 working days and give you an indication of how it will be dealt with. If you remain unhappy or wish to make a formal complaint, please contact the Chair of the Research Ethics Committee at the University of Oxford, who will seek to resolve the matter as soon as possible:

The Chair, Education Departmental Research Ethics Committee,
Email: student.curec@education.ox.ac.uk

13. Further Information and Contact Details

If you would like to discuss the research with someone beforehand (or if you have questions afterwards), please contact:

Janani Balaji

Department of Education

University of Oxford

15 Norham Gardens

Oxford, United Kingdom

OX2 6PY

Email: janani.balaji@education.ox.ac.uk

Appendix D – Recruitment Message

Hello! I'm Janani, and for my master's dissertation, I am conducting a study to explore the impact of an individual's first language on their understanding of the structure of English words.
I am looking for volunteers aged 18 to 25 years to partake in a series of online tasks focused on measuring word reading and understanding grammatical relationships between words. Participants must be currently enrolled in an Indian university and must have either Tamil or Hindi as their first language (must be fluent in speaking and reading/writing in their first language).
You will be required to complete an online experiment, which will take you through a few simple tasks. More information will be given in the information sheet that I have linked here. Please read through the information sheet for all the details.
Completing the experiment will take 20 minutes at most!
If you are interested and have any other queries, please contact me at janani.balajis@education.ox.ac.uk.
Thank you!

Links to the documents:
Participation information sheet
https://drive.google.com/file/d/156Aq409QYVWKo1KSym6nf5P9ZZASc818/view
Link to the experiment
https://research.sc/participant/login/dynamic/93E80816-7479-4AC3-B8F8-C548B8D70F1D

Appendix E – Task 1: Consent Form

Consent to take part in 'The Impact of L1 on Morphological Awareness and Word Reading in English as an L2'

Central University Research Ethics Committee (CUREC) approval reference:
EDUC_C1A_24_086

Purpose of Study: To identify whether an Individual's first language impacts how they understand the structure of English words

	Please initial each box if you agree with the statement
I confirm that I have read and understand the information sheet for the above research. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.	
I understand that my participation is voluntary and that I am free to withdraw any data collected by 25/07/2024 , without giving any reason.	
I understand who will have access to any personal data provided, how the data will be stored and what will happen to the data at the end of the project.	
I understand that I will not be identifiable from the research outputs in the published dissertation	
I give permission for you to contact me again to clarify information.	
I understand how to raise a concern or make a complaint.	
I agree to take part.	

Name of participant Date e-Signature

Appendix F – Task 2: Interest Form

PARTICIPATION IN TASK 2

The second task in this interview consists of a one-on-one interview with the interviewer, which will be conducted on Microsoft Teams. A detailed information sheet will be provided if you choose to participate in this task.

Participation in this task is voluntary.

Please indicate your interest to participate below.

I am interested to participate in the task

Yes
No

Appendix G: Participant Information Sheet – Task 2

The Impact of L1 on Morphological Awareness and Word Reading in English as an L2

PARTICIPATION INFORMATION SHEET

Central University Research Ethics Committee Approval Reference: EDUC_C1A_24_086

Introductory paragraph

You are being invited to take part in the second task of this research project. Before you decide to participate, it is important for you to understand what this task will involve. Please take time to read the following information carefully and discuss it with others if you wish. If there is anything that is not clear or if you would like more information, please ask the researcher. Take time to decide whether you wish to take part.

What is this task and why is it being conducted?

This is the second task of this study, and it is a one-on-one interview with the researcher. The goal of this task is to understand how participants went about answering the questions provided in Task 1 – the online questionnaire.

Why have I been invited to take part in this task?

You have been invited to take part in this task because you indicated in the questionnaire that you would be interested in partaking in a second task.

Do I have to take part in this task?

No. It is up to you to decide whether to take part. You can withdraw your data from the research without giving a reason by contacting me using the details provided. The deadline by which you can withdraw any information you have contributed to the research is [date to be set once ethical approval is received – approximately 1 month after end of data collection ends]. After this time, the data will have been anonymised. If you decide to withdraw your participation before this date, any collected from you will be permanently deleted.

What will happen in this task?

This is a one-on-one interview which will happen on Microsoft Teams. This interview will be recorded with your permission. The interviewer will present some of the questions you had answered in the previous task and will ask you some questions about how you found the answers to these questions. There are no right or wrong answers in this task.

Are there any risks associated with taking part in this task?

There are no foreseeable disadvantages or risks associated with taking part in this task. The data obtained from the interviews will be anonymised – identifiers like your name will be removed during the transcription process. You will not be directly identifiable from the data

.**What information will be collected and how will it be used?**

In this task, video recordings will be collected, with your consent. The recordings will be deleted after they have been transcribed. The transcripts of interview recordings will be stored in a secure, encrypted OneDrive for Business account connected with the University of Oxford. The collected data will be stored for 3 years after publication of the work following which it will be permanently deleted. Your data will not be shared with any third parties. For the duration of the research, only my supervisor (Dr. Anna-Maria Ramezanzadeh) and I will have access to the data.The data collected will be presented in a general manner, and quotations will be included only with your permission. You will not be identifiable in any of the research output. You can also opt out of this.

Will the research be published?

Could I be identified from any publications or other research outputs? The findings from this research will be written up as a part of a master's dissertation submitted to the University of Oxford. A copy of my dissertation will be deposited both in print and online in the Oxford University Research Archive, where it will be publicly available to facilitate its use in future research. You will not be directly identifiable from the output of the research.

Data Protection

The University of Oxford is the data controller with respect to your personal data, and as such will determine how your personal data is used in the research. The University will process your personal data for the purpose of the research outlined above. Research is a task that is performed in the public interest.

Further information about your rights with respect to your personal data is available from the University's Information Compliance website at https://compliance.admin.ox.ac.uk/individual-rights.

Who has reviewed this research?

This research has received ethics approval from a subcommittee of the University of Oxford Central University Research Ethics Committee. (Ethics reference: EDUC_C1A_24_086)

.**Who do I contact if I have a concern about the research or I wish to complain?**

If you have a concern about any aspect of this research, please contact Janani Balaji at janani.balaji@education.ox.ac.uk or Dr Anna-Maria Ramezanzadeh at anna-maria.ramezanzadeh@education.ox.ac.uk, and we will do our best to answer your query. I will acknowledge your concern within 10 working days and give you an indication of how it will be dealt with.If you remain unhappy or wish to make a formal complaint, please contact the Chair of the Research Ethics Committee at the University of Oxford, who will seek to resolve the matter as soon

as possible:The Chair, Education Departmental Research Ethics Committee, Email: student.curec@education.ox.ac.uk

Further Information and Contact Details

If you would like to discuss the research with someone beforehand (or if you have questions afterwards), please contact:

Janani Balaji
Department of Education
University of Oxford
15 Norham Gardens
Oxford, United Kingdom
OX2 6PY
Email: janani.balaji@education.ox.ac.uk

Appendix H – Task 2: Consent Form

Consent to take part in the Interview task

Central University Research Ethics Committee (CUREC) approval reference:
EDUC_C1A_24_086

Purpose of Study: To identify whether an Individual's first language impacts how they understand the structure of English words

	Please initial each box if you agree with the statement
I confirm that I have read and understand the information sheet for the above research. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.	☐
I understand that my participation is voluntary and that I am free to withdraw any data collected by 25/07/2024, without giving any reason.	☐
I understand who will have access to any personal data provided, how the data will be stored and what will happen to the data at the end of the project.	☐
I understand that I will not be identifiable from the research outputs in the published dissertation	☐
I give permission for you to contact me again to clarify information.	☐
I understand how to raise a concern or make a complaint.	☐
I agree to take part.	☐
I consent to being video recorded.	☐
I understand how the videos will be used in research outputs.	☐
Use of quotations: Please indicate your preference (select one option):	
(a) I do not wish to be quoted	☐
(b) I agree to the use of quotations in research outputs if I am not identifiable.	☐

Name of participant	Date	e-Signature

Appendix I – Language History Questionnaire

Background

Please answer the following questions about yourself.

Demographic Data

Name

Are you...

Male
Female
Non-binary
Prefer not to say
Other (please specify)

Age

Please select...
18 - 20
21 - 23
24 - 26
26 - 28

Country of Origin

Country of Residence

Education level (Indicate the level at which you currently are)

High school, Bachelor's, Master's, Doctoral

Indicate your native language

Tamil

Hindi

Indicate any other languages you know and the age at which you learned them

Indicate the language of instruction used by your teachers during your school years

Indicate the language of instruction used by your teachers during university

What language(s) do you use to communicate at Home?

What language(s) do you use to communicate at University?

What language(s) do you use to communicate with your friends?

Rate your current ability in terms of listening, speaking, reading, and writing in your native language

	1 (not at all confident)	**2**	**3**	**4**	**5 (very confident)**
Reading					
Writing					
Speaking					

Use the comment box below to provide any other information about your language background or usage (Leave unanswered if there are no additional details)

Do you also speak/use any dialects of the languages you know? Please indicate the name(s) of the dialect(s) (If none, leave question unanswered)

Appendix J - LexTALE

Instructions

In this task, you will see a string of letters onscreen. You will need to decide whether this is an existing English word or not.
If you think it is an existing English word, you click on "yes"
If you think it is not an existing English word, you click on "no".
If you are sure that the word exists, even though you don't know its exact meaning, you may still respond "yes".
If you are not sure if it is an existing word, you should respond "no".
In this experiment, we use British English rather than American English spelling.
For example:
"realise" instead of "realize"; "colour" instead of "color", and so on. Please don't let this confuse you.
You have as much time as you like for each decision.
There are 60 items in this task.

Test Items

1. Mensible
2. Scornful
3. Stoutly
4. Ablaze
5. Kermshaw
6. Moonlit
7. Lofty
8. Hurricane
9. Flaw
10. Alberation
11. Unkempt
12. Breeding
13. Festivity
14. Screech
15. Savoury
16. Plaudate
17. Shin
18. Fluid
19. Spaunch
20. Allied
21. Slain
22. Recipient
23. Exprate
24. Eloquence
25. Cleanliness
26. Dispatch
27. Rebondicate
28. Ingenious
29. Bewitch

30. Skave
31. Plaintively
32. Kilp
33. Interfate
34. Hasty
35. Lengthy
36. Fray
37. Crumper
38. Upkeep
39. Majestic
40. Magrity
41. Nourishment
42. Abergy
43. Proom
44. Turmoil
45. Carbohydrate
46. Scholar
47. Turtle
48. Fellick
49. Destription
50. Cylinder
51. Censorship
52. Celestial
53. Rascal
54. Purrage
55. Pulsh
56. Muddy
57. Quirty
58. Pudour
59. Listless
60. Wrought

Appendix K – Tamil Proficiency Test

Instructions

You will be presented with a Tamil sentence with a missing word, along with four options. You must choose one of the options that you feel completes the sentence best.
There are 40 items in this task.

Test Items

நான் அவரை _____ பார்க்கிறேன்.

A. அடிக்குது
B. அடிக்கும்
C. அடிக்கடி
D. அடிப்பது

ஒரு தொலைபேசி _____.

A. ஏன்
B. என்
C. என்னை
D. எண்

அவர் _____ வாசித்தார்.

A. குழல்
B. புயல்
C. முயல்
D. சூழல்

_____ அந்த பேனாவை கொடு.

A. என்னுடைய
B. அதுக்கு
C. எனக்கு
D. உனக்கு

_____ தேதி என்ன?

A. இன்று
B. ஒன்று
C. உன்று
D. என்று

ஆசியாவில் எத்தனை _____ உள்ளன?

A. நாடினால
B. பாடல்கள்
C. நாடுகள்
D. நாட்டுகள்

இங்க வேலை _____.

A. கொள்ளப்படும்
B. செய்யப்படும்
C. தள்ளப்படும்
D. சொல்லப்படும்

இந்த பொருளின் விலை _____.

A. அதிகப்படும்

B. அதிகாரிக்கும்
C. வடிக்கப்படும்
D. படிக்கப்படும்

அவர் அந்த பள்ளியில் ஒரு _____.

A. கடைக்காரர்
B. மனிதர்
C. அதிகாரி
D. ஆசிரியர்

மலைகளில் _____ உண்டு

A. பனியும்
B. பாணியும்
C. பனியில்
D. பனியின்

அவர்கள் _____ விளையாடினார்கள்.

A. திங்களாகவே
B. தாங்களாகவே
C. தேன்களாகவே
D. எங்களாகவே

அந்த நோய் _____.

A. பறந்தது
B. பரவியது
C. பறித்தது
D. பிறந்தது

ஒரு _____ தாய்

A. .பலமான
B. பழமான
C. பழிமான
D. பலமணி

பள்ளியில் _____ மாணவர்கள் உள்ளனர

A. ஏதாவது
B. ஏராளமான
C. ஆரோக்கியமான
D. அதிகமான

வாழ்க்கையில் அவருக்கு ஒரே _____ உண்டு

A. போக்கும்
B. பக்கம்
C. ஏக்கம்
D. நோக்கம்

அவன் இவனுடைய _____.

A. சிறுவன்
B. சிறுவகை
C. சிறுநகை
D. சிறுமண்

இங்கு பல பணிகள் _____.

A. ஏற்படுவது
B. ஏற்படுகின்றன
C. ஏற்றப்படும்
D. ஏற்படப்படுவோம்

அவர்கள் அவரை செந்தில் என்று _____.

A. அழைத்தனர்
B. அழைத்தநீர்கள்
C. அழைத்தார்களே
D. அழைப்பார்

கதையில் ஒரு பகுதி _____ பாடப்படும்.

A. இசைமகை
B. இசைக்காக
C. இசையுள்ள
D. இசையுடன்

பொங்கல் ஒரு பெரிய _____.

A. திருவிழா
B. தணிவிழா
C. பெருவிழா
D. சிருவிழா

இது ஒரு _____ இடம்.

A. புண்யமான
B. புனிதமான
C. புனிதமையை
D. புனிதமோட

அவர் _____ மாற்றினார்.

A. விஷயத்தை
B. விஷ்யத்துடன்
C. விஷ்யத்த்தோடு
D. விஷயம்

சிங்கம் காட்டின் _____ என்று அழைக்கப்படும்

A. ராஜை
B. ராஜ்யம்
C. ராஜயோகம்
D. ராஜா

உலகின் பல _____ அதிகம் மக்கள் உள்ளனர்.

A. நகரங்களில்
B. நகரங்களோட
C. நகரங்களின்
D. நகரங்கள்

சில _____ கதவுக்கு வெளியே காத்திருந்தனர்.

A. குழந்தைகள்
B. குழந்தைகளுடன்

C. குழந்தை
D. குழந்தையுடன்

நான் என் _____ நேரத்தில் தூங்கினேன்.

A. வழக்கமோடு
B. வழக்கமுடன்
C. வழக்கம்படி
D. வழக்கமான

ஒரு மணி நேரத்தில் அறுபது _____ உள்ளன.

A. நிமிஷத்தில்
B. நிமிடங்கள்
C. நிமிடங்களாக
D. நிமிடம்

இன்று புத்தகக் _____ நிறைவு நாள்.

A. கண்காட்சி
B. கண்காட்சியில்
C. கண்காட்சியின்
D. கண்காட்சியோடு

எனக்கு ஐந்து _____ வேண்டும்.

A. ஆணி
B. ஆணிகள்
C. ஆணிகளை
D. ஆணிகளே

அவர்கள் அனைவரும் _____ விளையாடினர்.

A. கால்கள்கால்நடை
B. கால்வண்டி
C. கால்பந்து
D. கால்பந்து

நாங்கள் _____ விளையாடினோம்.

A. கடற்கரையான
B. கடற்கரையோடு
C. கடற்கரையுடன்
D. கடற்கரையில்

நீங்கள் என்ன சாப்பிட _____?

A. விரும்புறார்கள்
B. விரும்புகிறீர்கள்
C. விரும்புறோர்
D. விரும்பறை

என்னால் உன்னை _____ முடியவில்லை.

A. உணர்ந்துகொள்ள
B. உணர்ந்துட
C. உணர்ந்துபார்க்க
D. உணர்ந்துடைய

நம் _____ காத்திருக்கலாம்.

A. பேரூந்துநிலையம்
B. பேருந்துநிலையத்தில்
C. பேரூந்துநிலையத்துடன்
D. பெருந்துநிலத்த்தோடு

வயலில் நிறைய _____ உள்ளன.

A. கால்நடைகள்
B. கால்பந்து
C. கால்நடைக்கு
D. கால்பந்துகள்

எனக்கு மிகவும் _____ இருந்தது.

A. வருத்தம்
B. வருத்தமாக
C. வருத்தமிக
D. வருத்தகமுடன்

குழந்தைகளுக்கு _____ வளர்ச்சி தேவை.

A. ஆரோக்கியமுடன்
B. ஆரோக்கியம்
C. ஆரோக்கியத்துடன்
D. ஆரோக்கியமான

_____ ஒரு எழுத்து இருக்கணும்.

A. குறைந்தது
B. குறைந்தபட்சி
C. குறைந்தபட்சம்
D. குறைந்தபகுதி

அவர் ஒரு _____ மனிதர்.

A. சாதாரண
B. சாதாரணமாய்
C. சாதாரணமுடன்
D. சாதாரணமாக

இது மிகவும் மோசமான _____.பகுதம்

A. பகுதி
B. புத்தகம்
C. பக்கம்
D. பகுதி

Appendix L – Hindi Proficiency Test

Instructions

You will be presented with a Tamil sentence with a missing word, along with four options. You must choose one of the options that you feel completes the sentence best.
There are 40 items in this task.

Test Items

वह एक _____पढ़ रहा है |

A. किसान
B. किताब
C. विमान
D. किसका

मुझे उसे _____था |

A. रोकना
B. रोका
C. रुकी
D. रुका

वर्षा के दिनों में इन नदियों में बाढ़ एक बड़ी _____है |

A. समाया
B. तपस्या
C. समस्या
D. समास्या

मैं _____बात कर रहा था |

A. आपको
B. आपके
C. आपस
D. आपसे

आज मेरे _____का जन्मदिन है |

A. पिताजी
B. पिटाई
C. पतानहीं
D. पताजी

मुझे _____दे दो |

A. ख़तम
B. कलम
C. कमाल
D. कालम

आप _____सही हैं |

A. बिलिकुल
B. बाकी

C. बिल्कुल
D. बिक्कुल

यह _____घर है |

A. पूरा
B. पुरानी
C. पूरी
D. पुराना

आपको दी गई पुस्तकों का _____करें |

A. उपयोग
B. उपयरग
C. उपयाग
D. उपराग

किताबें _____ में समय लगता है |

A. लिखा
B. लिखने
C. लिखना
D. लिखी

सूरज की _____बहुत अच्छी है |

A. रोशन
B. राशन
C. रोशनी
D. रोसन

मेरे पास कोई और _____ है |

A. विकास
B. सकल्प
C. निकलप
D. विकल्प

क्या तुम _____?

A. समझे
B. समाज
C. समजो
D. समज

उनका _____ बहुत खराब था |

A. स्वयं
B. स्वभाव
C. स्वयम
D. स्वदेश

मेरी _____ देखो |

A. तरस
B. चारश
C. तरफ़
D. सराफ

आज हम कम _____ हुआ है |

A. बिका
B. बेकारी
C. बकरी
D. बिक्री

यह हमारी कंपनी का _____है |

A. निदेशक
B. निदेश
C. नदी
D. नदीसब

यह मेरी _____ से अधिक भारी है |

A. आपसक
B. अपेक्षा
C. आपेक्षक
D. आभास

यह परीक्षा _____स्तर के छात्रों के लिए है |

A. मधुमं
B. मढ़ा
C. मध्यम
D. मृद्यम

_____ पानी में गयी |

A. मच्चर
B. मगर
C. मटर
D. मछली

हम _____समझौते पर पहुंचे हैं |

A. आपसी
B. आपसे
C. आपस
D. आप

आज _____ हो रहा है |

A. समझाकर
B. साक्षात्कार
C. समस्कार

D. चमकतार

यह _____ हमारे लिए सबसे उपयुक्त रहेगा |

A. करीब
B. खराब
C. करार
D. क़रीर

_____ मांगना बंद करें |

A. बक
B. भावी
C. बिका
D. भीख

_____ मांगना गलत है |

A. दहेज़
B. दही
C. दहा
D. देहा

जब भारत स्वच्छ होगा, स्वस्थ होगा, तभी भारत _____ भी होगा |

A. समर
B. समृद्ध
C. समिद्ध
D. समर्थ

दिल्ली इस देश में एक _____है |

A. नागरम
B. निगारी
C. नगरी
D. निगारम

भारत में कई ____ नदियाँ है |

A. बॉडी
B. बेटी
C. बड़ी
D. बढ़िया

जंग के _____ का कारण साफ़ नहीं है |

A. अलग
B. एलम
C. आला
D. ऐलान

एक नया काम _____ |

A. बनाएं

B. बनेगा
C. बन
D. बनिया

उसके पास तीन _____उपाधियाँ हैं |

A. समीक्षा
B. शैक्षणिक
C. शिस्क्षा
D. शीषक

मैं अब एक _____ लगा सकता हूँ |

A. अँधा
B. अंडिया
C. अंदाजा
D. अंदाज़

इस बार _____बहुत ठंडी होगी |

A. सरदर्द
B. सरदार
C. सिर
D. सर्दी

मेरा परिवार _____ है |

A. धनी
B. घड़ी
C. घनी
D. दबी

यह बहुत _____ रखती है |

A. माया
B. मायने
C. मणि
D. मानी

कभी घोडे की _____ नहीं की क्या |

A. सागरी
B. सागारी
C. सवारी
D. सवरी

हम लगातार उन _____को देख रहे हैं |

A. खबरों
B. घबरा
C. घबराए
D. खबर

हम _____हैं |

A. वीजा
B. विजयी
C. भिजाई
D. विजय

मैंने उस _____ से बात की |

A. अजनब
B. अचानक
C. अजनबी
D. अच्छामि

हमने अच्छे _____देखे हैं |

A. नदियां
B. नतीजा
C. नदी
D. नतीजे

Appendix M – Word Identification Task

Instructions

A word will be presented to you on the screen.
Read the word one time.
Once you have read the word, click the 'Next' button to move to the next word.
There are 17 items in this task.

Test items

1. Epidemic
2. Proximity
3. Embassy
4. Judicious
5. Volatile
6. Naïve
7. Opulent
8. Philanthropist
9. Alkali
10. Quintessence
11. Tubercle
12. Expostulate
13. Epistrophe
14. Tableau
15. Scintillant
16. Zeitgeist
17. Oeuvre

Appendix N – Word Attack

Instructions

A made-up word will be presented to you on the screen.
Read the word one time.
Once you have read the word, click the 'Next' button to move to the next word.
There are 17 items in this task.

Test items

1. Bufty
2. Vunhip
3. Knaf
4. Twem
5. Adjex
6. Yeng
7. Laip
8. Zirdn't
9. Straced
10. Cedge
11. Wrey
12. Whumb
13. Knoink
14. Bafmotbem
15. Monglustamer
16. Pnir
17. Ceisminadolt

Appendix O – Derivation Task

Instructions

In this task, you will be given a sentence with a missing word, along with a a base word. You are required to produce the right form of the word to complete the sentence.

For example: My uncle is a _____ (Farm).

The answer is 'farmer'.

You will need to type out your answer in the space provided onscreen.

Please pay attention to the spelling and ensure that you are typing in lowercase letters.

There are 28 items in this task.

Test items

1. warm. He chose the jacket for its _____ .[warmth]
2. teach. He was a very good _____. [teacher]
3. permit. Father refused to give _____. [permission]
4. profit. Selling lemonade in summer is _____. [profitable]
5. appear. He cared about his _____. [appearance]
6. express. 'OK' is a common _____. [expression]
7. four. The cyclist came in _____. [fourth]
8. remark. The speed of the car was _____. [remarkable]
9. protect. She wore glasses for _____. [protection]
10. perform. Tonight is the last _____. [performance]
11. expand. The company planned an _____. [expansion]
12. revise. This paper is his second _____. [revision]
13. reason. Her argument was quite _____. [reasonable]
14. major. He won the vote by a _____. [majority]
15. deep. The lake was well known for its _____. [depth]
16. equal. Boys and girls are treated with _____. [equality]
17. long. They measured the ladder's _____. [length]
18. adventure. The trip sounded _____. [adventurous]
19. absorb. She chose the sponge for its _____. [absorption]
20. active. He tired after so much _____. [activity]
21. swim. She was a strong _____. [swimmer]
22. human. The kind man was known for his _____. [humanity]

23. wash. Put the laundry in the _____.[washer]
24. humor. The story was quite _____. [humorous]
25. assist. The teacher will give you _____. [assistance]
26. mystery. The dark glasses made the man look _____. [mysterious]
27. produce. The play was a grand _____. [production]
28. glory. The view from the hill top was _____. [glorious]

Appendix P – Affix Word Choice Task

Instructions

In this task, you will be given a sentence with a missing word. Choose one of the words given below to complete the sentence.

Example: John wants to make a good _____ on this date. [impressive, impressionable, impression, impressively]

Answer: impression

There are 26 items in this task.

Test items

1. Fortunately, age improved his _____ . (personality personal personify personalize)
2. My assistant will _____the new procedure. (demonstration demonstrate demonstrative demonstrable)
3. The secret police arrested the _____ before he could give his speech. (active activist activate activize)
4. They _____ those fields early in the spring. (fertilizer fertility fertilization fertilize)
5. John didn't anticipate the harshly _____ response to his work. (criticism criticize critical critically)
6. The committee was not persuaded by the arguments of the _____ (reductionist reduce reductive reductional)
7. Frank broke down under the highly _____ questioning. (intensive intensity intensify intensification)
8. The _____ of the geese was complete by Thanksgiving. (migration migratory migrate migrational)
9. The success of the entire _____ depends on Bob (operative operational operation operationalize)
10. All four studies produced nearly _____ results. (identity identical identify identification)
11. They _____ their own desires at the expense of the group. (gratification gratify gratuity grateful)
12. Three separate agencies _____ the traffic in that sector (regular regularity regulation regulate)
13. They hope to _____ their investments. (diversity diversion diversify diversionary)
14. It is impossible to _____ people's thoughts. (legislate legislative legislature legislation)
15. The _____ of their approach prevented many errors. (systematic systematicity systematize systematically)
16. The cost of _____ keeps going up. (electric electrify electrical electricity)

17. His consistently _____ behavior eventually destroyed his family. (adultery adulterate adulterous adulterousness)
18. They should _____ that room if they plan to grow orchids in there. (humidity humid humidifier humidify)
19. Only the most _____ males survived the winter (activity active activation activate)
20. You can't _____ results from studies done only on rats. (generalization generality generalize generalizable)
21. The new owners turned the failing business into a highly _____ operation (production produce productive productivity)
22. The _____ targeted the new administration. (satiric satirical satirist satirize)
23. They planned to _____ the entire southern coast. (colonist colonize colonial colonization)
24. Only the most _____ farmers showed any profit that year. (industrious industry industrialize industrialization)
25. Continued food shortages finally caused the _____to revolt (popular popularity popularize population)
26. It was an overwhelmingly _____ conclusion. (glorify glorification gloriousness glorious)
27. We all appreciate the tremendously _____ part you played in securing the grant (instrumental instrumentation instrumentality instrument)

Appendix Q – Affix Non-Word Choice

Instructions

In this task, you will be given a sentence with a missing word. Choose one of the nonsense words given below to complete the sentence.

Example: John wants to make a good _____ on this date. [impressive, impressionable, impression, impressively]

Answer: impression

There are 26 items in this task.

Test items

1. They _____ the data in the back office. (curfamic curfamation curfamate curfamity)
2. All those models are strictly _____ and outdated as well. (ambilemptify ambilemptivist ambilemptity ambilemptive)
3. In spite of his _____ he did an outstanding job. (dispribize dispribation dispribational dispribify)
4. Desert animals are not normally _____.(commalianization cornmalious commalianism commalianize)
5. He is so _____ that he offends almost everyone. (dictopithify dictopithification dictopithial dictopithity)
6. You can't even begin to _____without modern equipment. (equamanize equamanizable equamanity equamanive)
7. They presented the highly _____ evidence first. (credenthive credenthification credenthicism credenthify)
8. They hope to _____ the two sides together. (uniromosity uniromify uniromous uniromative)
9. He wants to _____while he still can (fidamoration fidamorian fidamorational fidamorate)
10. Please try to be as totally _____ as possible. (progenalism progenalize progenious progenify)
11. Please _____ these forms as soon as possible. (scribsumptist scribsumptious scribsumptian scribsumptize)
12. The story of the _____ was repeated every year. (vergalize vergalicious vergalify vergalist)
13. The most _____ samples were discarded. (birendal birendment birendalize birendify)
14. Dr. Jones, a well-known _____, is speaking tonight. (circumtarious circumtarist circumtarify circumtarize)
15. We should _____ that money by the end of the year. (relaptification relaptian relaptify relapmble)
16. His _____ is greatly admired. (superfilize superfilive superfilial superfilation)
17. The meeting was highly _____ and invigorating. (loquarify loquarial loguarialize loquarialism)

18. Too much _____ is bad for the economy. (malburuity malburuify malburnicious malburuable)
19. Their progress was stopped by an unexpected _____ . (postramify postramic postramity postramicize)
20. Their approach to the problem is deceptively _____. (torbatify torbative torbativize torbature)
21. The breeders _____ their stock every four generations. (genilify genility genilification geniliar)
22. She met her first _____ when she moved out west. (benedumptist benefumptify benedumptize benedumptuous)
23. Everyone resented the obvious _____ on the manager's part. (spectitious spectitionalize spectition spectitive)
24. You must _____ them quickly or you'll ruin the colors. (premanicism premanicize premanicity premanic)
25. All the suspiciously _____ specimens are kept in a separate tank. (tribacize tribacion tribacism tribacious)
26. The new equipment will _____ everything automatically. (transurbate transurbativity transurbatist transurbative)

Appendix R – Morphological Relatedness Task

Instructions

You will be shown a pair of word and will be asked to indicate whether you think the words presented onscreen are related to each other (whether one comes from the other).

If you think the words are related, click 'Yes'

If you think the words are not related, click 'No'

Example: Plenty, Plentiful

Answer: Yes

There are 40 items in this task.

Test items

Person - Personal

Allow - Allowance

Beauty - Beautiful

Ill - Illegal (foil)

Region - Regional

Bag - Baggage

Pity - Pitiful

Let - Letter (foil)

Parent - Parental

History - Historic

Atom - Atomic

General - Generosity (foil)

Superior - Superiority

Poet - Poetic

Participate - Participation

Humor - Humanity (foil)

Deep - Depth

Sincere - Sincerity

Supervise - Supervision

Major - Magic (foil)

Meter - Metric

Nature - Natural

Add - Addition

Ear - Earth (foil)

Associate - Association

Concentrate - Concentration

Divide - Division

Import - Impression (foil)

Relate - Relation

Decide - Decision

Reduce - Reduction

Insult - Insulation (foil)

Sign - Signature

Resign - Resignation

Bomb - Bombard

Numb - Numbers (foil)

Crumb - Crumble

Sign - Signal

Knowledge - Acknowledge

Comb - Combination (foil)

Appendix S – Suffix Ordering Task

In this task, words were created by combining a highly frequent base with two derivational suffixes of either correct or incorrect ordering and participants were asked to decide whether the given string of letters was an English word or not (An example: suititiable).

This task consists of 44 trials, with 22 correct words and 22 incorrect words, with a maximum raw score of 44. The focus of this task was to examine participants' knowledge of the selectional constraints of English suffixes

Instructions

You will be presented with a string of letters onscreen.

On a scale of 1 to 6, you will have to rate how sure you are that the presented word is definitely an English word.

The scale: 1 - Definitely not an English word to 6 - Definitely an English word

There are 60 items in this task.

Test items

Regularness, Neutralness, Ethnicness, Legalness, Rapidness, Jokeable, Smileable, Ar rivable, Departable, Leavable, Smartal, Largeal, Coldal, Darkal, Slowal, Hopenessful, Sinnessful, Thanknessful, Forcenessful, Wastenessful, Affordi tiablem Suititiable, Readitiable, Dependitiable, Repeatitiable, Correctaltion, Operatealtion, Opposealtion, Distributealtion, Relatealtion, Darkness, Awareness, Illness, Fitness, Goodness, Workable, Eatable, Reliable, Acceptable, Comfortable, Traditional, Functional, National, Personal, Optional, Playfulness, Truthfulness, Forgetfulness, Mindfulness, Cheerfulness, Usability, Respectability, Predictability, Applicability, Adaptability, Intentional, Additional, Educational, Institutional, Situational

Appendix T – Stimulated Recall: Questions and Test Items

General Introduction

Thank you for joining. Just to give you a recap, this is an informal interview, just a one-on-one chat walking me through how you went about answering some of the items from two of the tasks. I just wanted to remind you that this interview will be recorded. Once the interview has been transcribed, this video will be deleted. Any and all data that may identify you will be removed, so you will be completely anonymous. If at any point you feel that you want to withdraw from this interview, you're completely free to do so.

Introductory question

These were the tasks which contained made up words. All I want you to do is walk me through your thought process as you went about solving these tasks. I will present you with each individual task, just walk me through what exactly went through your mind as you answered the question. You are not being marked on your answer, but just keep yourself in the same frame of mind as you were when you took the test. Let's begin.

Word Attack – Questions

- Was there any particular aspect of this word that caught your eye?
- Did you find any of the items particularly difficult?
- What went through your mind as you read this word?

Test items

Vunhip, Adjex, Zirdn't, Straced, Cedge, Whumb, Knoink. Bafmotbem, Monglustamer, Ceisminadolt

Affix Non-Word Choice Task – Questions

- What made you choose your answer?
- Were there any items that you found difficult?

Test items

- All those models are strictly _____ and outdated as well. (ambilemptify ambilemptivist ambilemptity ambilemptive)
- Desert animals are not normally _____.(commalianization cornmalious commalianism commalianize)
- You can't even begin to _____without modern equipment. (equamanize equamanizable equamanity equamanive)

- Please try to be as totally _____ as possible. (progenalism progenalize progenious progenify)
- The most _____ samples were discarded. (birendal birendment birendalize birendify)
- The meeting was highly _____ and invigorating. (loquarify loquarial loguarialize loquarialism)
- Too much _____ is bad for the economy. (malburuity malburuify malburnicious malburuable)
- She met her first _____ when she moved out west. (benedumptist benefumptify benedumptize benedumptuous)
- You must _____ them quickly or you'll ruin the colors. (premanicism premanicize premanicity premanic)
- The new equipment will _____ everything automatically. (transurbate transurbativity transurbatist transurbative)

Appendix U – Sample of Transcript

Transcript
Task 2: Interview with Tamil Participant

Researcher:
OK. I'll just start it.
Alright, so just to give you a recap, basically this whole task is this very informal - It's an informal task, so like no worries, but you and the other participant I'm interviewing both of you have completed the task that I gave you in the first round, right? And in that there were these couple of tasks that might have been a bit strange? Yeah, like a little bit like, you know, the ones with pseudo words and the meet up words.

T1:
Hmm.

Researcher:
I am conducting this interview just to find out how you guys came to the conclusions that you guys came to in that task, so I will be presenting you with the item like this is not all the items that will take way too long. I've chosen like 10 items from each task and I'm going to just show you the item on the screen and I just want you to sort of transport yourself back to the first time you saw that and just give me a brief rundown of how exactly you think you approach the task.
It may not be like one word, one sentence. It could give me, like, even if you wanna talk for like, "oh, I think I did this. And then I did that. And then I did that. And then I did that." You could do that. OK.
Yeah. So it's basically just, it's this is something called a stimulated recall. It's basically just me trying to find out exactly the thought processes that were running behind your head when you first started doing the task. Cheers?

T1:
Ok.

Researcher:
I'm going to present my screen. I just want you to tell me if it's visible. OK.

T1:
OK.

Researcher:
OK. Is it visible?

T1:
Yes.

Researcher:
OK, so the first set of the first task that I want you to look at is. It's called the effects non word task. So basically this was a task in which you were asked, Sorry this is called the word attack task. This is a task in which you were asked to read the word that was presented to you out loud one time.

T1:
OK.

Researcher:
What you can do if you want during the task, right, during this session is you can read the word out loud and explain to me exactly how you went about reading the word, whether there were certain syllables that you were stressing on, whether you like, split the word into two, you get what I mean? Just what went through your head, whether it was just like one thing and then you were like, “yeah, OK on to the next.” OK, let's begin.

T1:
Mm hmm yeah.

Researcher:
OK, so you can either read this out loud or just tell me how you went about reading this.

T1:
OK. OK I think I'm gonna pronounce it as Vanhip because the VUN looks like VAN, which is like Wan.

Researcher:
Mmhm

T1:
So I just put the ah sound for U and I was like, Van. And then hip is like an onward. So OK, I was like something like. Vanhip

Researcher:
Mm hmm.
Yeah.
Yeah, OK. So I'm not gonna be judging you on pronunciation. OK? This is not like a test of

your pronunciation. I just want to see how you're reading the word. OK. So you're saying you're splitting the word into vun and hip right? And you're doing this because one of the words sort of reminds you - both of the words remind you of other words, you know? OK, OK, next word.

T1:
OK, I think I'll pronounce it like Adjex - one is ad, which is like advertisements. So that made me remember the word. Oh, ad we say ads everything else. Jex I think the sound of X made me say je odha sethu (with Je), so I was like OK, Jex, adjex.

Researcher:
Yeah.
Mm hmm.
OK, nice again splitting into 2 sub words. OK, next.

T1:
Zidnt, OK.
See this again same splitting method – I was like OK what's first? ZI and then the apostrophe thing. It reminds me of the word didn't or whenever we use the apostrophe. So it's something with that sound. So I was like, yes, Zirdnt with the UNT sound.

Researcher:
Alright.
Mm hmm. OK.
Were any of these words so far particularly challenging when you were reading them the first time?

T1:
Hmm, not the first two. Maybe this was a little challenging 'cause I had to comprehend what it was.

Researcher:
OK. So for the next few words, also tell me that whether it was challenging or not to read, OK, I forgot to mention in the first half, OK.
Next.

T1:
Straced. I think it's straced, OK, first what I think the word that came to me was starraised because I saw it as sta and then my mind came up with star.

Researcher:
OK. Again, not judging your pronunciation.

T1:
Like star and then I saw raised. So I was like star-raised and then I realized it's not star. So I was like OK ‘firstu straced apdinu irundhaen’ (first I was like straced), But then I was like, no, if it's a ‘star-aced apdi nu irundhudhu na’ (if it was like star aced, like that)

Researcher:
Hmm.

T1:
So maybe there's a ah sound. So I saw it as two different words, but then because of the STR I thought, it was star and then raised. But then I was like, oh, sounds like straced.

Researcher:
OK. OK. You still somehow split this word, right?

T1:
Yeah.

Researcher:
OK, next one.

T1:
Cedge? Sedgee?

Researcher:
Yeah. OK. OK.

T1:
Yeah ‘poga poga’ (as it goes) I think it's more hard to try to comprehend the word – from the third. I've been noticing that like, how do I say this word again? But I think with -

Researcher:
OK. But if yeah, if you were to just sort of explain what you were doing, would you still say you're splitting the word into 2?

T1:
I'm splitting the word into 2 but I'm also using elements of whatever which was familiar, like

with star, I did that. Now with this CE immediate association was college. I was like OK, there was A C and E, the C and E, ‘immediate-a therinja word’ (the word that I knew immediately) was college. But then I was like OK, it looks like college, but no it's not college because the CEGE, so it's like OK Sed-Ge. But then GE makes a difference on OK, that's different. So I think I'm my brain with this. How it works is I use.

Researcher:
OK. Mm hmm. That's all right, yeah.

T1:
It's sort of like crystallised intelligence, where in the familiar thing I bring it in because I see similar elements to it and then I'm like, oh, OK, maybe it's this. OK. I'm getting familiar with this word with that type of sense. And then I try to split it being like, OK Sed-Je.

Researcher:
Mm hmm. OK. OK, that's fine. What about this one?

T1:
Yeah, same thing happened. First I thought it was. Womb because it had the um sound with umm and it had starting W and end with B. So I thought you know, it was womb, but then it was different. It sounded like womb, but then it doesn't. The womb that I know. And then the humb. When I saw ‘thania’ (alone) I was like, oh, that reminds me of humble.

Researcher:
OK. Yes. Next.

T1:
I think when I'm pronouncing this, I use the splitting method. I was like OK, no ink, no ink. But then with the K and then this is again reminds me of Kodak Studios or same element.

Researcher:
OK, next one.

T1:
Baf-mot-bem – Ok this I split into 3. And I think this I pronounced it with Baf-mot-beam or Bem, yeah.

Researcher:
OK. But you split it into three parts, yeah.

T1:
Yeah.

Researcher:
Mm. Ok. What about this?

T1:
Mong Lus-lus tamer, but I think because first I thought, OK, mong, lust amor, but then it didn't make sense for me I wanted - For me to understand the word I needed to have some sense to it, so I was like, OK, mong lus and tamer, ok that sounds like a word. So, I split it into three in that sense.

Researcher:
OK. Right. And this one?

T1:
Ces, Mina, Dolt. Ces,Mina,Dolt.. Same. Ces, split. Mina, I split and then Dolt was separate.

Researcher:
Ok, so that was the first task.
To summarize, you seem to be splitting the words whenever you can, and you also seem to be kind of going back to other things that you're familiar with. So let's say, like other words, that may look similar or, you know, sort of have the same context and you're using that as a reference point. Do you think, the more complex the word got, the slower you were reading because you couldn't comprehend the word?

T1:
Yeah.
Yeah, definitely. I think with small words, you'll have the ability to read quicker, but then with big words first it was the unfamiliar aspect and then I was like, OK, no, this definitely is a big word. So when I broke it down, only I slowly started getting familiar with the previous aspects. So as the word got big or even when it was, I think the third word when it was quite different when it had like.

Researcher:
OK.

T1:
A symbol I was like. Oh, OK. Like this is difficult.

Researcher:
The apostrophe, yeah.
OK.
Awesome. that was very actually quite interesting. OK, so the second talk that I want you to look at is called the affix non Word choice task. In this task, just to refresh your memory, you were given a sentence with a missing word and to fill that gap, you were given four nonsense words to choose from.
And all I'm gonna be asking you is:
What made you choose the word that you chose? If you can't remember, just try to try to remember. If not to make your decision. It's not a right or wrong - I'm not going to be giving you a score. I'm not going to tell you if it's right or wrong either. All I want to know is what made you choose that particular word. OK, whether it like feels right or says something about the grammatical structure of the word. OK, just give me that. I'm not going to tell you if it is right or wrong. So you're under no pressure. Yeah, not judging you either. That goes without saying.

T1:
Mm hmm.

Researcher:
OK, I'll present my screen again. Let me know if you can see it. Can you see my screen?

T1:
Yes.

Researcher:
OK. That's good. Your first sentence. Just say ABCD, don't bother reading the words.

T1:
I think I'm gonna go with D

Researcher:
OK.

T1:
Yeah.

Researcher:
What made you choose D?

T1:
When I was reading the sentence, I was trying to see which made the most sense sound wise and grammar wise. I was like, OK, all those models are strictly -ify, no vitist? No, like ambilemptify, amblilemptity no, Ok. You know, it sounded most sense when it ended with the -tive. It was like, ah, OK, this sounds good.

Researcher:
So you were specifically focusing on that one like last bit?

T1:
Yeah. Last. And when we say sentences. Oh, she ran quickly. She ran in a quick manner. You end with certain alphabets. And that tells the tense of the sentence, 'illaya?' (right?). So idhulavandhu (in this) I was focusing on the last particular thing. And I was like, OK, which makes more sense, like, the fitting manner, strictly -ative and outdated as well. Made sense for me.

Researcher:
OK, that's great. Moving on to the next one.
'Desert animals are not normally _____.' Again, you can just say ABCD.

T1:
I'm gonna go with B because it said animals, so my mind instantly went to something with plural and then I saw OUS and that made most sense.

Researcher:
Mm hmm, Again, you were drawn to the last bit is it?

T1:
Yes, last bit

Researcher:
OK, OK. Next one you can't even begin -

T1:
And also that said specifically animals. So when it was saying animals, it went OK, it's not something singular like this document or 'indha' (this) bag or 'indha' (this) fruit it said all animals. So it included an aspect which is majorly plural. So 'Adhula paaka modhu' (when you look at that) OUS also you use sometimes for plural tense. So I was like oh OK something made - this is is what is making sense for me.

Researcher:
Mm hmm. OK.

T1:
Yeah.

Researcher:
OK.
‘You can't even begin to _____ without modern equipment.’

T1:
OK, can't even begin to.
Yeah, I'm gonna go with A cause again, It's the last thing that makes sense as a grammatical tone for me. I was like, you can't even begin to -nize without modern equipment. I don't think for me, -ble or -ty or -nive made sense than -nize.

Researcher:
OK. That makes sense.

Researcher:
Next one,
‘Please try to be as totally _____ as possible.’

T1:
Yeah, I'm gonna go with C.

Researcher:
Mm hmm.

T1:
I saw the sentence in a way where someone was telling a particular crowd, OK, try to be in this particular sense. So ‘Adulah pakamodu’ (when we look at that) when I imagine telling them this sentence, I saw the OUS sound again. That made the most sense.

Researcher:
Hmm. So in this time, you place the sentence in sort of another context to understand it, OK.

T1:
Yeah. I saw it like a statement someone was selling to a particular crowd and I then

imagined it. OK, if someone was instructing me something and they told please try to be as totally _____ as possible, then I'll be like, OK, they're saying this only.

Researcher:
Hmmm Ok.
The next one, 'the most _____ samples, were discarded.'

T1:
I'm gonna go with A cause. This time one thing was different. I started noticing from the last round.

Researcher:
Mm hmm.

T1:
'Adhukku munnadi varikkum' (Before all that) all the other – which, how much? Which card is this? 6th or 5th?

Researcher:
I think it's the fifth one, but again these are not all the items on the task either.

T1:
So imagine if this was the first one and 'idhukku munnadi' (before this) was the 4th one. I think until third one, I was very sure that OK, this one thing only made sense, but from the last round I've been having thoughts like oh, but this also makes sense like this also can be 'apdi nu' (like that). For previous round, I had to think over. I was like, OK, if I had to say one option, I would go with this. But with this it's like OK, since it it feels like a very general statement where you can put any sort of tense and it'll make sense.

Researcher:
Hmm.

T1:
I'm like, OK, I have options now, but 'indha option layum' (in this option as well), I'm going with what like my gut says - OK, I think this is the tense that I am choosing and this is the tense that I'm putting. 'Andha vidhatthula' (In that way), I'm choosing this answer. So 'adhila' (in that) I can say I'll choose A.

Researcher:
OK. So in your decision to choose like one option out of the multiple options that you feel are correct, are you looking at like any particular part of the word and making your decision or

are you thinking of like previous examples and situating this word with your previous knowledge and then sort of coming to choose it?

T1:
With this particular sentence, I'm just looking at the same last – how does it end with? And with what sound? OK -ment, -lize, -defy, -dal. 'Adhu dhaan paakaren' (I am looking at that only).

Researcher:
Oh, OK.

T1:
I think with this series, only that particular element.

Researcher:
OK, perfect. Next one.
'The meeting was highly _____ and invigorating.'

T1:
Yeah, I'm gonna go with B because when I said the meeting was highly -lism -lize -fy, -rial made the most sense.

Researcher:
Mm hmm. OK. Right. Was this like an easy choice? Like an oT1ious choice for you, or you had to think about it?

T1:
Ah, this was easy. This was not too challenging like the last one.

Researcher:
OK. So off the like 6 items you were shown so far, how many would you say were challenging?

T1:
Hmm. 2.

Researcher:
OK, the next one. 'Too much _____ is bad for the economy.'

T1:
I'm gonna go with A. Because when the word economy popped up, it reminded me with the

previous experience I was like, OK, something with -ty, -ty like ITY makes sense, because it kinda means some sort of money or saving came up and then I was like, OK. Too much, maybe. It's indicating of something which is a source, like a type of resource like water or something. And then I was like, OK, purity – 'water nu sonnonae' (when I said water, I remembered Purity which came with Ty and then I was like, OK, I remember the sound of purity. And then I was like, OK, Malburuity, Malburuify, -cious, -able. And then I was like no, I think -TY makes the most sense.

Researcher:
OK. Yeah, that does make a lot of sense.

Researcher:
The next one. 'She met her first _____ when she moved out West.'

T1:
I'm gonna go with A because, with this, with the sentence I framed, I framed it as a person meeting someone 'innuru edathu la' (in another place), and when it started off saying she met her first something maybe like an individual or like a 'position irukura individual' (an individual with a position), The ending of tist, made more sense like with a profession like a physiotherapist.
Like someone with a position, I think I've labelled it with ending with the IST. So when I saw she met her first dash, I was confused between A and D, but then when I saw the familiarity of already known words with the person of a position ending with TIST, I went with A.

Researcher:
OK. That makes sense.
The next one.
'You must _____ them quickly or you'll ruin the colours.'

T1:
This again was confusing. I was confused between B&D. When I saw the sentence 'You must dash them quickly or you will ruin the colours' I thought of it as OK. Something involving with colours. So it's a process. Something like OK sugar pannamodhu (making) You know how they caramelize. So again, I saw the last - this time I saw the entire word and I just tried. OK. Premanic or premanicize and then since it's a process which involves maybe 'time bound a irukalam' (could be time bound) or 'task bound a irukalam' I thought of cooking. And I remember the word caramelize. So with the ze element. And then I was like, OK, but D also makes sense. Maybe in painting if there's something in this element. I tried to kept on saying OK Premaniic apdinu. It was like -ciize makes more sense because it feels more familiar. It feels like a word that you'll use for a process like this because because it says you

must do this or the colours will get ruined. So that indicates that an action – a particular action needs to be taken or else the process will get ruined. So I've gone with B.

Researcher:
Hmm. You've gone with D, OK.

T1:
B.

Researcher:
B for ball?

T1:
B for ball, yeah.

Researcher:
OK. This is your last one.
'The new equipment will _____ everything automatically.'

T1:
OK. Yeah, I'm going to go with A cause. When you said - when you were reading. This new equipment will dash everything automatically. My mind immediate without looking the options, The word transfer came into the mind. And then when I was looking at the options, since everything was with a touch of trans, A and D made sense, but then when I added the D word with the sentence and I started to pronounce it, it didn't make much sense. So I was like, OK, I think I'll go with A.

Researcher:
Hmm Ok. So that was it for the second task. So just to summarise for my benefit.
For most of the words you found that you were looking only at the suffix - like the last bit of the word, but there was some words that you found challenging enough to look at the whole word, and then you manually, you know, put them into the sentence, and then you're like, OK, this feels right. And for those words, you also kind of drew on your past knowledge of English, like this might work - even though this is like a nonsense word.

T1:
Yeah.

Researcher:
OK, that's good. OK. Alright. Thank you so much. I'll just stop the recording here.

Appendix V – Breakdown of L1 proficiency self-rating

Participant L1 self-ratings						
	Reading		Writing		Speaking	
Rating	**Tamil**	**Hindi**	**Tamil**	**Hindi**	**Tamil**	**Hindi**
1	1	0	5	4	0	0
2	2	0	10	15	0	0
3	1	4	13	9	1	1
4	23	21	12	4	5	7
5	26	21	13	14	47	38
Total	53	46	53	46	53	46

Raw data of participant self-rating scores

	Tamil		Hindi	
Skill	**Mean**	**Std. Dev**	**Mean**	**Std. Dev**
Reading	4.34	.854	4.37	.645
Writing	3.38	1.228	3.15	1.475
Speaking	4.83	.470	4.85	.132