THE CONSTRAINTS OF DUAL MORPHOLOGICAL SYSTEMS ON VISUAL WORD PROCESSING IN MALTESE

by

Jonathan A. Geary

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF LINGUISTICS

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA
2022

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Jonathan A. Geary, titled: *The constraints of dual morphological systems on visual word processing in Maltese*, and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Heidi Harley — Date: 11-29-2022
Heidi Harley

K Forster — Date: 12-1-2022
Kenneth Forster

Diane K. Ohala — Date: 11-29-2022
Diane Ohala

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Heidi Harley — Date: 11-29-2022
Heidi Harley
Dissertation Committee Chair

Acknowledgments

I started graduate school at the University of Arizona knowing that I wanted to study the psycholinguistics of Semitic root and pattern morphology as well as contribute to Native language revitalization efforts. The writing of this dissertation represents the culmination of the first interest: Through no fault of the Maltese language, which remains one of the most interesting languages that I have ever studied, my priorities have shifted entirely towards Native language revitalization, and this will probably be the last Maltese research that I conduct. As when I began to work on this dissertation, however, I hope that the MaltLex database, the construction of which I describe here, will continue to be used to contribute to the linguistic scholarship on Maltese for years to come.

I have many people to thank for getting me to this point. First and foremost, I want to thank Skye Anderson for her support at every step of the process, for being my editor and sanity-checking me countless times, and for giving me a voice to look forward to hearing at the ends of the long days of data collection in Malta, among other things. Átawišamaš. Likewise, I want to thank my parents, my grandparents, and the rest of my family for their encouragement and enduring patience. Grandma Spagnola says that there should be a law against children moving too far away from their family, and while I don't agree with her, I have started to understand her reasons. Átawišamataš.

I also want to thank my dissertation committee for sharing their expertise with me here and throughout grad school, and for being patient with me as I wrote this dissertation. Heidi became my advisor during a very stressful time in my life, helped me through several significant transitions during grad school, and managed to push me just enough to get me to write the damn dissertation. Diane was my first friend among the UA faculty, and effectively the first faculty member whom I TA-ed for: I learned much from Diane about psycholinguistics, as well as about how to teach to diverse audiences, present linguistic concepts effectively, and encourage learners. Ken pioneered the experimental paradigms and software that I use throughout my work and so, in an important sense, is probably the one person most responsible for my being able to complete my dissertation research. More importantly, I consider Ken to be the model academic because of his commitment to understanding how our experimental and statistical methods work and why we should use the methods that we use, rather than just accept common practice. More academics should be like Ken.

I have been fortunate to have several other mentors who deserve special recognition. Luis Barragan first connected me with the Piipaash language, supported me as I transitioned away from academia into a more fulfilling career in language revitalization, and continues to encourage me to pursue Yuman language research. Sue Kalt introduced me to Quechua as an undergraduate, often gives me a productive outlet for my coding interests, and likewise helped me transition into a rewarding non-academic career. Táwtalikš (Fred Hill Sr.) shared the Umatilla language with me, warmly welcoming me to Nixyáawi and treating me like a long lost son. K̓ʷałanáwašamaš, natútas! Katrina Miller and Gretchen Kern offered me my first job outside of academia, and have been supportive mentors and great friends ever since. Amadeo Rea, whose ethnobiological scholarship I admire, taught me a lot about developing meaningful, lifelong relationships with the people with whom we interact as researchers: I was fortunate to spend three weeks with Amadeo during the COVID-19 pandemic, digitizing his audio tape collection while living in his shed. Átawišamataš.

I want to thank the Institute of Linguistics and Language Technology and the Department of Cognitive Science at the University of Malta, especially Albert Gatt and Ray Fabri, for making this research possible; Jessica Formosa for facilitating this research and making me feel welcome while in Malta; Abigail Galea for helping to prep the MaltLex stimuli; and Sean Fenech, Norbert Grech, Nicole Mifsud, Nicole Tabone, and Matthew Bonanno for their warm friendships in Malta.

Many other folks deserve thanks for their support and friendship during grad school, among whom I wish to thank individually: John Coopie IV; Thomas Naylor; Carl Ceraolo, MD; John Powell; Andrew Zupon; Maria Alexeeva; Jian Gang Ngui; Ryan Walter Smith; Lucy Hall Hartley; Max Mulé; Ivette González; Amy Fountain; Marian Wiseley; Kimberley Young; Diana Archangeli; Simin Karimi; the Piipaash Elders; my friends from the Gila River Indian Community and the Salt River Pima-Maricopa Indian Community: Robert Johnson, Kelly Washington, Sierra Ward, and Ron Carlos; my friends from the Confederated Tribes of the Umatilla Indian Reservation: Kakínaš (Thomas Morning Owl), Alikalát (Damien Totus), Cisyawak (Mildred Quaempts), Lisa Minthorn, Claire Huesties, and Aaliyah Dick; Annie Smith; Tyusmákt (Mersaedy Atkins); Tiičám K̓iluɫá (Bobby Fossek); Wiwéeletitpe (Brosnan Spencer); Jeanne Raymond; and my new friends from the Shingle Springs Rancheria: Daniel Fonseca, James Sarmento, Richenda Ervin, and Julian Fonseca.

Finally, I want to thank all birds (Figure 1), who again and again have reminded me that there is much more to life than linguistic research and data analysis. K̓ʷaɫanáwašamataš.

This dissertation was funded by the National Science Foundation (Award #BCS-1918143).

Figure 1. Female Black Redstart (*Phoenicurus ochruros*). Photo taken by the author at the University of Malta on November 23, 2019, during a rare break from data collection.

Dedication

To female Black Redstarts (*Phoenicurus ochruros*).

Table of contents

List of figures

List of tables

Abstract

In this dissertation, I conduct a visual lexical decision megastudy for the Semitic language Maltese and describe the construction of "MaltLex", a database of visual lexical decision responses to 11,000 real words and 11,000 non-words by native Maltese speakers. MaltLex is the first such megastudy for a Semitic language, which characteristically use nonconcatenative morphological processes: Nonconcatenative morphology poses novel challenges for theories of word recognition and morphological decomposition which may be explored through the MaltLex dataset. Maltese is also unique among Semitic languages in several respects, such as in its use of the Latin alphabet and in its split lexicon (roughly half of all Maltese words are of non-Semitic origin and use only concatenative morphological processes), and so Maltese provides unique testing conditions for the role of letters, morphology, etc. in visual word processing. I conduct several "virtual experiments" (i.e. where I analyze subsets of the MaltLex dataset) in order to (1) validate MaltLex by replicating well-established findings from the lexical decision literature and (2) build upon our understanding of lexical processing in Maltese and Semitic languages by demonstrating several novel findings.

In Chapter 1, I explore how word frequency and individual differences in language use and proficiency mediate lexical processing in Maltese. I replicate the canonical word frequency effect (readers judge more frequent words faster and more accurately) and show that a frequency measure that is based on the number of documents in which a word occurs in a corpus (contextual diversity) better predicts lexical processing performance than does a more traditional measure that is based on a word's total number of occurrences (word frequency), which is consistent with megastudy-based research on non-Semitic languages (e.g. Brysbaert and New 2009). All MaltLex participants were bilingual in Maltese and English, and I also show that the more English-dominant participants

exhibited smaller processing differences in judging the lexicality of non-Semitic versus Semitic Maltese words than did the more Maltese-dominant participants, which is broadly consistent with "cognate advantage" effects found in multilingual lexical processing (e.g. Poort and Rodd 2017).

In Chapter 2, I explore how the orthographic form of a word mediates lexical processing in Maltese. I replicate standard orthographic neighborhood density effects: MaltLex participants judge real words that have more orthographic neighbors faster and more accurately than those with fewer neighbors, though this advantage becomes smaller as words increase in frequency; and they judge non-words with more neighbors slower and less accurately than those with fewer neighbors (e.g. Andrews 1989). Beyond the standard neighborhood density effect, I also show that MaltLex participants are even slower and less accurate at judging the lexicality of non-words that differ from an existing Maltese word except by the addition or omission of a diacritic, suggesting that Maltese readers process certain pairs of letters alike during lexical processing (e.g. "g" and "ġ").

In Chapter 3, I explore the role that the consonant letters that comprise individual words play in visual lexical processing in Maltese. In a novel visual masked priming study, I show that similar priming effects obtain for non-Semitic versus Semitic words when primed by strings that consist of their consonant letters (e.g. *pnġ* priming *PINĠA* 'to paint'), which is consistent with subset priming effects in non-Semitic languages (e.g. Anderson and Geary 2018, Duñabeitia and Carreiras 2011) but inconsistent with Geary and Ussishkin (2018), who found such priming effects for Semitic Maltese words but not non-Semitic Maltese words. I then analyze the MaltLex dataset, showing that the frequency of a word's set of consonant letters across the Maltese lexicon mediates unprimed lexical processing and thus may mediate these differences in observed priming effects.

I have hardly scratched the surface of the range of analyses that may be performed using the MaltLex dataset, which I make freely available to other researchers. I conclude by describing the structure of the MaltLex dataset and how other researchers may access it in Chapter 4.

Chapter 1

MaltLex: A database of visual lexical decision responses to 11,000 Maltese words

1.1 Introduction[1]

In this chapter, I report on a visual lexical decision "megastudy" for the Semitic language Maltese and describe the construction of a database of Maltese lexical decision responses called "MaltLex", the first of its kind for a Semitic language. In a megastudy, researchers collect many behavioral responses to a wide range of stimuli, often from a wide range of participants. In lexical decision megastudies, researchers assemble lexical decision responses to stimuli that differ across a range of characteristics to produce a massive database of lexical decision responses. This often includes words that are atypical of the stimuli used in traditional experiments (e.g. both uninflected and inflected words, words of various lengths) but which better reflect the linguistic stimuli that individuals encounter in everyday language use (Ernestus and Cutler 2015, Tucker et al. 2018). Any researcher can then analyze a subset of the total dataset to explore some variable(s) of interest, while simultaneously being able to control for other factors known to influence lexical processing.

Megastudies, which have grown in popularity as an alternative to traditional experiments in recent years, circumvent many of the shortcomings of factorial studies (i.e. where researchers select and test a set of items that differ minimally except in some variable of interest); see Keuleers and Balota (2015) for some of the advantages of the megastudy approach to conducting behavioral research. For example, in a typical megastudy participants encounter a diverse set of linguistic items during the experiment that better approximates real language experience, thus circumventing potential context effects that reflect an artificial similarity among test items. Megastudies avoid other potential experimenter biases in selecting items for different conditions, and the massive size

[1] I presented an earlier version of this chapter at the 33rd Annual CUNY Human Sentence Processing Conference.

of megastudy databases provides greater power than is typical of traditional experiments. Further, researchers may be able to study individual differences among language users, depending on the range of participants surveyed and the types of participant information collected by the researchers.

To date, visual lexical decision megastudies have been conducted on a range of languages that includes: Cantonese (Chinese Lexicon Project (CLP): Tse et al. 2017), Dutch (Dutch Lexicon Project (DLP): Keuleers, Diependaele, and Brysbaert 2010; Dutch Lexicon Project 2 (DLP2): Brysbaert et al. 2016), American English (English Lexicon Project (ELP): Balota et al. 2007) and British English (British Lexicon Project (BrLP): Keuleers et al. 2012), French (French Lexicon Project (FLP): Ferrand et al. 2010; Chronolex: Ferrand et al. 2011; MEGALEX: Ferrand et al. 2018), Malay (Malay Lexicon Project (MLP): Yap et al. 2010), Mandarin (MELD-SCH: Tsang et al. 2018), European Portuguese (Soares et al. 2019), and Spanish (González-Nosti et al. 2014). Table 1.1 summarizes the characteristics of previous megastudies. In terms of the overall number of responses collected, MaltLex is consistent with some of the smaller megastudies cited here.

A handful of auditory lexical decision megastudies have been conducted more recently on Dutch (BALDEY: Ernestus and Cutler 2015), Canadian English (MALD: Tucker et al. 2018), and French (MEGALEX: Ferrand et al. 2018), while other lexical decision megastudies have focused on visual masked form priming (Adelman et al. 2014) and on changes to lexical processing that occur across the lifespan of children (Schröter and Schroeder 2017). Even more recently, SPALEX (Aguasvivas et al. 2018, 2020) and the English Crowdsourcing Project (ECP: Mandera et al. 2020) have tapped into online crowdsourcing to create even larger visual lexical decision databases.

Megastudy datasets are designed to be so general as to be able to answer research questions that the creators themselves may have never envisioned. In recent years, researchers have used megastudy data to evaluate the predictions of competing word recognition models (e.g. Adelman

Table 1.1. Summary of visual lexical decision megastudies. I include only non-primed, in-lab studies conducted with adult, native speakers. Note that ELP (Balota et al. 2007), MLP (Yap et al. 2010), Chronolex (Ferrand et al. 2011), and Soares et al. (2019) also included speeded naming tasks, and Chronolex (Ferrand et al. 2011) also included a progressive demasking task, which I do not report. The number of items differed systematically across sessions in some studies, which I note in the "Items/Session" column in parentheses. For some studies, I had to estimate certain values based on other values provided by the authors, which I note using a tilde "~".

	Language	Participants	Real words	Non-words	Items/Session	Sessions/Participant	Datapoints
ELP (Balota et al. 2007)	American English	816	40,481	40,481	2,000 (1st), 1,372–4 (2nd)	2	2,749,324
FLP (Ferrand et al. 2010)	French	975	38,840	38,840	1,000	2	1,946,988
DLP (Keuleers, Diependaele, and Brysbaert 2010)	Dutch	39	14,089	14,089	500 (1st–56th), 178 (57th)	57	1,098,942
MLP (Yap et al. 2010)	Malay	40	1,510	1,510	1,020	3	~122,400
Chronolex (Ferrand et al. 2011)	French	35	1,826	1,826	912 (1st, 3rd), 914 (2nd, 4th)	4	~127,820
BrLP (Keuleers et al. 2012)	British English	78	28,730	28,730	500 (1st–56th), 230 (57th)	57	2,240,940
González-Nosti et al. (2014)	Spanish	36	2,765	2,765	~922	6	199,080

DLP2 (Brysbaert et al. 2016)	Dutch	81	30,016	29,601	500	62	2,495,448
CLP (Tse et al. 2017)	Cantonese	594	25,286	25,286	936-938	3	~1,670,000
MEGALEX (Ferrand et al. 2018)	French	96	28,466	28,466	356	50	2,596,095
MELD-SCH (Tsang et al. 2018)	Mandarin	504	12,578	12,578	2,096–2,100	1	1,056,552
Soares et al. (2019)	European Portuguese	55	1,920	1,920	980	4	211,200
MaltLex	**Maltese**	**104**	**11,040**	**10,954**	**400**	**1-35**	**237,476**

and Brown 2008); to compare the relative contributions of different lexical characteristics to word processing (e.g. Soares et al. 2019); to assess the relative efficacy of similar predictors of lexical processing efficiency, such as different word frequency norms (e.g. Brysbaert and New 2009, Ernestus and Cutler 2015, Keuleers, Brysbaert, and New 2010, Tucker et al. 2018); to contrast the factors that influence visual versus auditory word processing (e.g. Ferrand et al. 2018); to study the effects of repeated exposure to a given linguistic stimulus (e.g. Keuleers, Diependaele, and Brysbaert 2010); and to model word processing computationally (e.g. Baayen et al. 2011). To date, no megastudy has focused on a language that productively uses nonconcatenative morphology, which poses novel challenges for word recognition (e.g. Frost et al. 1997), which I describe in the next section. I aim to fill this gap by conducting the first visual lexical decision megastudy on a Semitic language, namely Maltese, which employs nonconcatenative word-formation strategies.

1.1.1 Why Maltese?

Traditional analyses (e.g. Berman 1978, McCarthy 1979) posit that native word stems in Semitic languages, such as Arabic, Hebrew, and Maltese, typically consist of two discontinuous morphemes: A consonantal "root", which assigns the word to a broad semantic field, and a vocalic and consonantal "pattern", which contributes grammatical and thematic information. For instance, Maltese words consisting of the root *k-t-b* pertain to 'writing' (e.g. *kiteb* 'to write; he wrote', *kitba* 'writing', and *ktieb* 'book'), while words containing the pattern *1V22ie3*[2] comprise agentive nouns (e.g. *kittieb* 'writer', *tebbieħ* 'chef', and *għalliem* 'student'). These nonconcatenative morphemes interleave to form word stems, which may combine with concatenative morphemes (i.e. prefixes and suffixes) to form more complex words (e.g. *kittieb* 'writer' + *-a* 'PL' > *kittieba* 'writers').

[2] I use numerals to indicate the positions of the consonants of the root morpheme, and *V* to indicate a short vowel.

The non-linear nature of Semitic root and pattern morphology poses unique challenges for decomposition-based models of word recognition (e.g. Taft 1979, 2004; Taft and Forster 1975) which posit that individuals decompose complex words into their constituent morphemes during processing, and that these morphemes are stored lexically and mediate lexical access (Frost et al. 1997, 2000). Nonetheless, psycholinguistic research has supported a role for root morphemes in mediating word processing in reading Arabic (e.g. Boudelaa and Marslen-Wilson 2001, 2004) and Hebrew (e.g. Deutsch et al. 1998; Feldman and Bentin 1994; Frost et al. 1997, 2000, 2005; Velan and Frost 2007, 2009, 2011), as well as Maltese (e.g. Geary and Ussishkin 2018, Twist 2006). For instance, Maltese readers judge word lexicality faster when subliminally primed by another word that shares the target's root morpheme (Twist 2006) or by a nonce string that consists of the letters of the root (Geary and Ussishkin 2018). Further, letter transpositions that involve word-internal root and non-root letters impede sentence comprehension in Hebrew (Velan and Frost 2007, 2011).

The structure of the Maltese lexicon provides unique testing conditions for the role of root morphology in word processing: Maltese descends from Siculo-Arabic (Agius 1996), a Maghrebi Arabic variety spoken in Sicily and Malta from the end of the ninth century until the end of the twelfth century (Brincat 2011, Comrie 1991), and thus, as a Semitic language, native word stems in Maltese typically consist of root and pattern morphemes. However, Maltese has spent the last millennium separated from other Semitic languages, instead developing in close contact with a series of Indo-European languages: first Sicilian and Latin, then Italian, and now English. Today, Maltese represents a mixed language combining features from its Semitic ancestry with those of the non-Semitic languages with which it has co-existed (Borg and Azzopardi-Alexander 1997). For instance, native words comprise approximately half the Maltese lexicon, with the bulk of non-native terms having been borrowed from Sicilian or Italian (Bovingdon and Dalli 2006, Brincat

2011, Comrie and Spagnol 2016). Except for early borrowings, many of which were integrated into the root and pattern system (Hoberman and Aronoff 2003, Mifsud 1995a), borrowed words usually do not consist of root and pattern morphemes and instead are subject only to concatenative morphological processes. In general, nonconcatenative morphology has become less productive (e.g. Mifsud 1995b) while concatenative morphology has flourished: For example, Maltese has borrowed and extended to native forms several inflectional suffixes of Indo-European origin (Borg 1994: 57, Gardani 2008: 75, Gatt and Fabri 2018, Hoberman 2007, Mayer et al. 2013).

Little research has investigated the processing of non-native words in Maltese, nor how Maltese's dual morphological systems interact in constraining lexical access. One exception is Geary and Ussishkin (2018), who compared subliminal priming by nonce letter strings consisting of the target's consonant letters: They found facilitation for Semitic words, for which these strings represented the word's root morpheme (e.g. *frx* priming *FIREX* 'to spread'; root: *frx*), but not for non-Semitic words, for which they were non-morphemic (e.g. *pnġ* priming *PINĠA* 'to paint'). The lack of priming for non-Semitic Maltese words contrasts with priming by comparable consonant-letter strings (e.g. *csn* priming *CASINO*) found in Indo-European languages like English (Anderson and Geary 2019) and Spanish (Duñabeitia and Carreiras 2011), and suggests that language-specific morphological patterns may influence the role of consonant letters in constraining lexical access (Anderson and Geary 2019). To facilitate further study of the non-Semitic stratum of Maltese, I have included words of Semitic and non-Semitic origin, as well as uninflected and inflected words that exhibit concatenative and nonconcatenative morphology, as targets in the MaltLex dataset.

Additionally, Maltese is, together with Cypriot Arabic, one of two Semitic languages that are written using the Latin alphabet. Other Semitic languages, like Arabic and Hebrew, use abjad writing systems that prioritize representing consonants and often leave vowels unwritten, and thus represent the roots of words but leave patterns (partly) unwritten. In contrast, all consonants and vowels are written in Maltese (i.e. neither morphemic element is privileged). Differences in writing systems may constrain the role of Semitic morphology in mediating lexical access: For instance, whereas Velan and Frost (2007, 2011) found that letter transpositions that involve root letters disrupt processing in Hebrew, Perea et al. (2012) found that letter transpositions, including those that involve root letters, generally fail to disrupt sentence/lexical processing in Maltese, which they ascribe to orthographic differences between the two languages (cf. Maltese patterns like Indo-European languages that are written using the Latin alphabet in that readers tolerate letter transpositions). Further, the use of the Latin alphabet in Maltese allows for more straightforward comparisons of the contributions of orthographic characteristics between Maltese and non-Semitic languages that use the same writing system, such as English (e.g. Balota et al. 2007) and French (e.g. Ferrand et al. 2018), for which visual lexical decision megastudy datasets already exist.

The rest of this chapter is organized as follows: In Section 1.2, I describe the construction of MaltLex and the data collection procedures. In Section 1.3–1.5, I demonstrate and validate the use of the MaltLex dataset for studying visual word recognition and the factors that mediate lexical processing by conducting analyses wherein I compare different measures of word frequency as predictors of lexical processing efficiency (Section 1.3), I compare responses to Semitic versus non-Semitic Maltese words (Section 1.4), and I explore how language dominance influences the processing of words of the two lexical strata (Section 1.5). I summarize my results in Section 1.6.

1.2 Methods

1.2.1 Participants

I collected responses from 104 native or near-native Maltese speakers,[3] all of whom were recruited through the University of Malta. Participants were at least 18 years old, with a mean age of 24.0 years (range: 18−77 years). Fifty-three participants identified as female and 51 as male; 87 participants identified as right-handed and 17 as left-handed. All participants were expected to be bilingual in Maltese and English, and so they completed the Bilingual Language Profile (BiLP; Birdsong et al. 2012) after their initial session to provide a composite, continuous measure of their language dominance (Appendix 1): Possible scores range from −218 to 218, with negative scores indicating greater English dominance, positive scores indicating greater Maltese dominance, and "0" indicating perfect balance. Participants exhibited a mean dominance score of 38.0 (range: −52.2−113.2), reflecting that participants tended to be dominant in Maltese rather than English. Participants whose performance met my inclusion criteria were invited to participate in additional sessions of the experiment (Section 1.2.3), and participants received €5 for each session completed.

Data from an additional 29 participants was collected but has been excluded from analysis and from the final dataset. These participants were excluded for identifying as late, L2 learners of Maltese, rather than (near-)native speakers; for having an accuracy rate below 80% (Diependaele et al. 2012); or for failing to follow study procedures (e.g. reporting that they chose not to respond as quickly as possible, which was reflected in slower response times relative to other participants; reading the potential words aloud; interacting with other participants or the experimenter, or using their phone during the session; and mishandling and damaging the experiment hardware).

[3] Two participants identified as having learned Maltese starting at the age of 1 years old and English at the ages of 3 and 5 years old, respectively. However, as they had also identified as having lived in Malta their entire lives and did not disclose having learned any other language since birth, I infer that they must have been exposed to Maltese since birth. A third participant identified as having learned English from birth but Maltese only starting at age 3.

1.2.2 Materials

I planned for participants to judge the lexicality of 11,000 real Maltese words and 11,000 non-words. The real-word targets were randomly selected from Korpus Malti v3.0 (Gatt and Čéplö 2013), a 250-million-token corpus of written Maltese which I trimmed to remove non-Maltese texts and non-words (e.g. URLs) as well as personal names, and supplemented with words from other sources. The selected words were checked against Ġabra (Camilleri 2013), a Maltese lexical database containing 16,593 lemma-based entries, and vetted by a native speaker to ensure that they were not offensive, archaic, or otherwise unlikely to be recognized as words by participants. Real-word targets included both uninflected and inflected forms, monosyllabic and multisyllabic words, words of Semitic and non-Semitic origin, negated forms, and words of various syntactic categories (e.g. nouns, verbs, adjectives, numerals, prepositions), such that the target set better reflects the language encountered in daily use than those used in more traditional lexical decision studies.

For each real-word target, I constructed a non-word counterpart by replacing at least one consonant letter to create a potential nonce form unattested in the trimmed version of Korpus Malti. Non-word construction respected properties of the Semitic and non-Semitic strata of the Maltese lexicon (e.g. the letter *ħ* was never used in creating a non-word from a non-Semitic target, as non-Semitic words rarely contain this letter; while the letter *p*, being exclusive to non-Semitic words, was never used to create a non-word from a Semitic target). Real-word targets and their non-word counterparts were matched exactly in length and approximately in orthographic neighborhood density, weighted according to neighbors' word frequency in Korpus Malti in order to compensate for non-words, misspellings, and other low frequency forms.[4] A native speaker vetted all potential

[4] I counted as an orthographic neighbor any word listed in Korpus Malti v3.0 (Gatt and Čéplö 2013) that differs from the target via (1) the substitution, (2) addition, or (3) deletion of a single letter (e.g. for *kittieb* 'writer', (1) *kittien* 'flax, hemp', (2) *kittieba* 'writers', and (3) *kitteb* 'to write regularly; he wrote regularly' were counted as orthographic neighbors). I weighted neighborhood density by summing each neighbor's word frequency value in Korpus Malti.

non-words to ensure that they were non-words, that they comprised plausible (i.e. phonotactically legal) non-words, and that they were not too similar to an existing word such that participants might easily confuse the two, and all problematic non-words were replaced and re-vetted.[5]

I divided the 22,000 targets into 55 lists, each containing 200 real- and 200 non-word targets. To mitigate repetition priming effects, real-word targets having the same lemma (e.g. *jikteb* 'he writes' and *tikteb* 'she writes', both of the lemma *kiteb* 'to write') were assigned to different lists (lemmas were obtained from Ġabra). To maintain length and neighborhood-density matching within lists, real-word targets and their non-word counterparts were assigned to the same list.

For each list, 20 real-word and 20-non-word targets were used as practice items. Practice items comprised targets that had been assigned to other lists. To mitigate possible long-term repetition priming effects, participants were assigned to lists across the study such that they never encountered an item which they had previously judged as a practice item as an experimental target.

The final set of targets includes 11,040 unique real-word targets and 10,954 unique non-word targets. [6] The former comprise 6,514 Semitic words, 4,429 non-Semitic (i.e. Indo-European)

[5] I systematically explore how neighborhood density impacts visual lexical processing performance in Maltese in Chapter 2, where I replicate several established orthographic neighborhood density effects. In particular, I find that Maltese readers judge the lexicality of real words with more orthographic neighbors faster and more accurately than words with few neighbors; that this advantage diminishes as words increase in frequency; and that readers judge the lexicality of non-words with more neighbors slower and less accurately than those with few neighbors.

[6] Following data collection I discovered that, due to coding errors and despite the vetting procedures in place for potential targets, a handful of items needed to be recoded as real or nonce prior to data analysis. Participants did not receive feedback during the experiment (Section 1.2.3) and so would not have known if they had given the "incorrect" response at any point, hence these inconsistencies will not have affected their lexical decision performance. Rather, these changes slightly alter the total number of real-word and non-word targets used across the experiment and within lists, as well as alter the overall characteristics of real-word versus non-word targets, which I report below. (1) Six words were used twice across the experiment, either twice coded as a real word, or once as a real word and once as a non-word. For each of these six pairs, the two instances of the same target had occurred in different lists. I recoded as real-word targets all such targets that had originally been coded as non-words, and corrected participants' response accuracies to reflect the target's corrected lexicality. Further, to mitigate any long-term repetition priming effects, I have excluded from analysis and from the final dataset a total of 16 datapoints where the participant judged the relevant target a second time. (2) One item, intended to be *nsabi* 'nets', was misspelled as *nasbi* (error rate = 73%), which I have recoded as a non-word target and corrected participants' response accuracies to reflect its corrected lexicality. (3) Lastly, 42 items were used as non-word targets that I have subsequently recoded as real words, and corrected participants' response accuracies to reflect the target's corrected lexicality: These items are attested as words in Ġabra (38 items were generated as potential non-words due to their non-occurrence in Korpus Malti v3.0) and were first identified after data collection due to their exhibiting error rates of 70% or higher in the lexical decision megastudy.

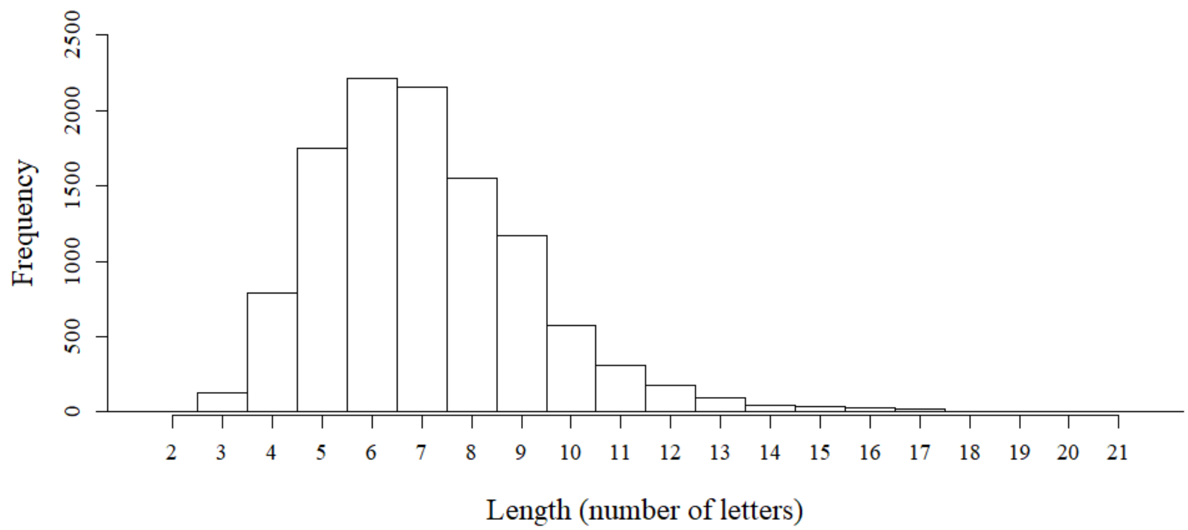

Figure 1.1. Distribution of word lengths for real-word targets (MaltLex).

words, and 97 words of uncertain origin (e.g. because their form is consistent with both a Semitic word and a non-Semitic Maltese word, or because their etymology is unknown; etymologies were taken from Aquilina 1987–1990). Each list contained 199−203 real-word targets (*M* = 200.8 real-word targets) and 197−201 non-word targets (*M* = 199.2 non-word targets). Real-word targets ranged in length from 2−21 letters (*M* = 7.1 letters), in word frequency from 0−20,446.1 occurrences per million words in Korpus Malti v3.0 (*M* = 36.7 occurrences per million words), and in word frequency-weighted orthographic neighborhood density from 0−49,584.6 occurrences per million words in Korpus Malti v3.0 (*M* = 157.8 occurrences per million words). I present the distributions of length, log-transformed word frequency, and log-transformed word frequency-weighted orthographic neighborhood density for real-word targets in Figures 1.1−1.3.

This is not to say that all non-word targets that had a high error rate were recoded: Legitimate non-word targets could exhibit high error rates resulting from, for instance, their bearing a strong resemblance to an existing word (e.g. the nonce target **xeffaq* resembles *xefaq* ‘horizon’), and I have maintained such items as non-words. Likewise some legitimate words exhibited high error rates, yet I consider them real words due their having passed stringent selection criteria: Each word (1) occurs in Korpus Malti v3.0 (Gatt and Čéplö 2013) or another Maltese source, (2) is attested in Ġabra (Camilleri 2013), and (3) was approved by a native speaker prior to data collection. High error rates for some targets are to be expected, and such data may prove informative regarding the words that Maltese readers actually know, the cues that they use in recognizing words (e.g. what spelling variation they subconsciously tolerate), etc.

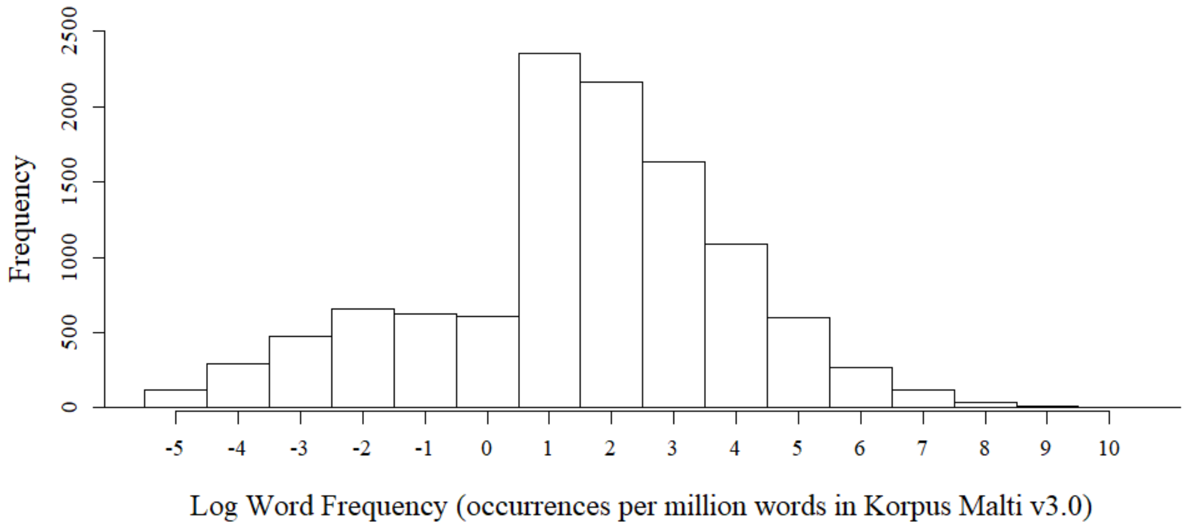

Figure 1.2. Distribution of log-transformed word frequency for real-word targets (MaltLex). Prior to applying a log-transformation, I added 1 to each raw word frequency value to account for items having a raw word frequency value of 0.

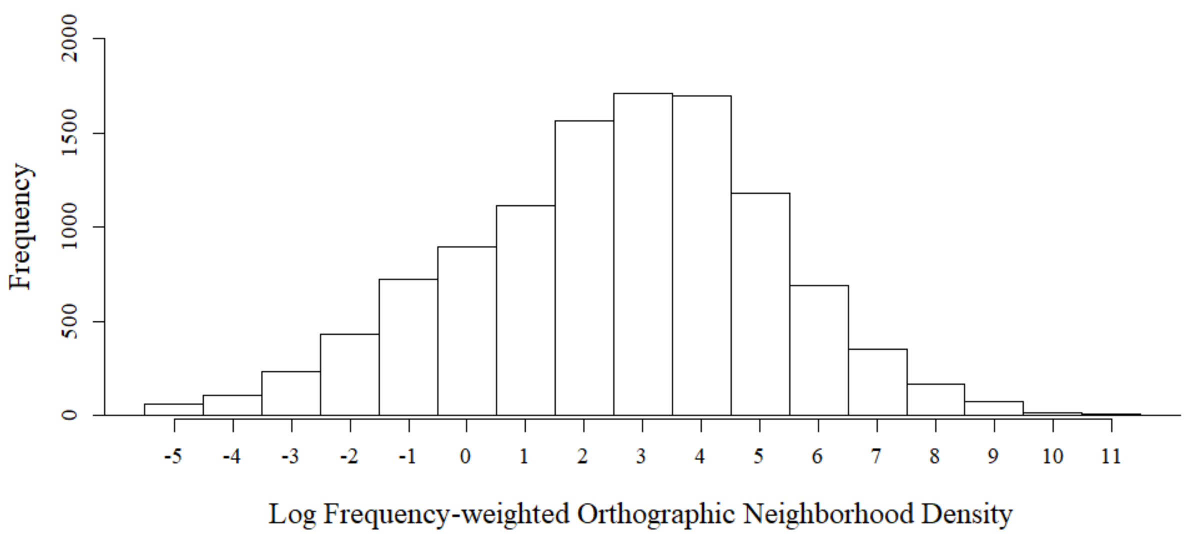

Figure 1.3. Distribution of log-transformed word frequency-weighted orthographic neighborhood density for real-word targets (MaltLex). Prior to log-transformation, I added 1 to each raw word frequency-weighted orthographic neighborhood density value to account for items having a neighborhood density of 0.

The results of a series of Welch's unequal variances *t*-tests comparing word length, word frequency, and word frequency-weighted orthographic neighborhood density for Semitic versus non-Semitic targets revealed (1) that non-Semitic targets had a reliably greater average length than Semitic targets ($M_{\text{Semitic}} = 6.5$, $M_{\text{Non-Semitic}} = 7.9$ letters; $t(6{,}959.3) = 32.12$, $p < 0.001$); (2) that Semitic targets had a reliably greater average frequency than non-Semitic targets ($M_{\text{Semitic}} = 40.4$, $M_{\text{Non-Semitic}} = 29.9$ occurrences per million words; $t(8{,}358.9) = -2.27$, $p < 0.05$); and (3) that Semitic targets had a reliably greater average neighborhood density than non-Semitic targets ($M_{\text{Semitic}} = 213.7$, $M_{\text{Non-Semitic}} = 70.1$ occurrences per million words; $t(7{,}367.7) = -7.27$, $p < 0.001$). The greater average frequency for Semitic targets likely reflects multiple factors, such as the fact that Maltese function words, which are among the most frequent words, tend to be of Semitic origin (see Comrie and Spagnol 2016 for discussion of the distribution of Semitic/non-Semitic words across semantic fields and parts of speech in Maltese). The shorter average length of Semitic words contributes to their greater average neighborhood density (shorter words tend to have more neighbors), and these differences in length and neighborhood density likely also reflect differences in source language phonotactics for Semitic versus non-Semitic Maltese words. Researchers should account for these differences in comparing lexical decision performance for targets of the two etymological strata.

Non-word targets ranged in length from 2−21 letters ($M = 7.1$ letters) and in word frequency-weighted orthographic neighborhood density from 0−54,575.6 occurrences per million words in Korpus Malti ($M = 124.3$ occurrences per million words). The results of a series of Welch's unequal variances *t*-tests comparing length for real-word versus non-word targets in each list were all non-significant, as were a similar series of *t*-tests comparing word frequency-weighted orthographic neighborhood density for real-word versus non-word targets per list in all but one list, suggesting that participants could not rely on these sub-lexical cues to judge target lexicality.

1.2.3 Procedures

The experiment took place in a room at the Institute of Linguistics and Language Technology at the University of Malta. Participants sat in front of one of three laptop computers during the experiment, facing away from other participants and seated approximately seven feet away from the nearest participant (Figure 1.4: A−C). Prior to the experiment participants received

Figure 1.4. Layout of the experiment room. The experiment took place in a room at the Institute of Linguistics and Language Technology at the University of Malta. Participants sat in front of one of the computers at stations (A−C). Depending on the number of participants present in a session, the experimenter sat either in a chair in the corner of the room (D) or at the near end of station (C), where they quietly monitored the session.

verbal instructions in English and written instructions in Maltese, and they were then given an opportunity to request clarification. Participants were instructed that they would see a series of possible words appear onscreen, one at a time, and that for each potential word they would need to decide whether it was a real Maltese word or not “as quickly and as accurately as possible”. Participants were instructed to respond during the study by pressing one of two bumper buttons on a Logitech F310 Gamepad to register their responses. I used a gamepad to record responses in order to ensure consistent and accurate response time (RT) measurements (Witzel et al. 2013).

As many as three participants completed a session simultaneously, providing a potential distraction for one another. Additionally, the Institute of Linguistics and Language Technology is located in a noisy parking lot and was adjacent to a loud construction project at the time of data collection, providing further potential distractions for participants. Thus, each participant wore a pair of active noise-cancelling headphones during the experiment to mitigate such distractions.

After receiving the instructions and consenting to participate, participants completed the 40 practice trials followed by the 400 experiment trials. Items were presented pseudo-randomly such that participants never received more than four real-word targets or four non-word targets in a row (this was not mentioned to participants). Participants were able to take an untimed break following the final practice trial and following the 150th and 300th experiment trials (during which participants remained seated, being able to stretch, take a drink, etc.). After completing the experiment trials, participants were asked to quietly remain seated while other participants finished the experiment. Following their first session of the experiment, participants were asked to complete a short background questionnaire, which included the BiLP (Birdsong et al. 2012) and additional questions asking about handedness and whether they had been diagnosed with a vision problem.

Stimuli were presented using DMDX (Forster and Forster 2003). Stimuli were presented in lowercase, 12-point Courier New font. On each trial, a fixation cross "+" first appeared in the middle of the screen for 500 ms, then was replaced by a blank screen for 100 ms, and then by the target for 3,000 ms. DMDX recorded responses starting from target onset: If a response was not made within 3,000 ms, DMDX recorded a non-response and proceeded to the next trial. During the experiment, participants received feedback only on non-response trials ("Ebda risposta", i.e. 'No response'), which DMDX displayed for approximately 1,500 ms before proceeding to the next trial; otherwise, DMDX displayed a blank screen for approximately 1,500 ms before proceeding.

Each session lasted approximately 40 minutes to 1 hour. After completing their first session, participants were debriefed about the nature of the study, and then those participants (1) who identified as native (or near-native) speakers of Maltese in the background questionnaire, (2) who had achieved an overall accuracy rate exceeding 80% (Diependaele et al. 2012), and (3) who had not exhibited other performance issues (e.g. reporting that they did not respond as quickly and as accurately as possible on each trial) were invited to participate in additional sessions of the experiment, during which they were assigned to a new list of items (the order of list assignment was randomized). Performance was monitored across sessions, and further participation ceased if a participant's accuracy rate dropped below 80% or they exhibited other performance issues.[7]

Participants were able to sign up for as many as three sessions per day (they took at least a ten-minute break between sessions), and participants completed between 1 and 35 sessions apiece (*M* = 5.8 sessions). I originally planned to have a smaller number of participants complete a larger number of sessions apiece, namely for 24 participants to complete 30 total sessions apiece.

[7] Although no participant who completed multiple sessions had an overall accuracy rate that dropped below 80%, a number of participants exhibited sub-80% accuracy rates during their first session and so were not invited to participate further; their data is omitted from the final dataset. Additionally, several participants began engaging in disruptive behavior as sessions accrued (e.g. interacting with other participants, using their cellphones), necessitating that their participation cease and that data for their affected session(s) likewise be omitted from the final dataset.

However, participants varied considerably in the amount of time they could commit to the study and in their focus across sessions of the study, and so I adopted the current procedures wherein participants could participate in any number of sessions up to a max of 35. To account for potential effects of practice and fatigue, I recorded for each participant not only trial number within a given session (expecting performance to diminish across sessions due to fatigue), but also session number in a given day and overall session number for each participant (expecting performance to improve with practice across sessions), and I have included these values in the final dataset.

I collected approximately 237,500 analyzable lexical decision responses in total, which includes 10−13 responses per target (*M* = 10.8 responses).[8,9] Thus, MaltLex is consistent in terms of the number of responses collected with some of the smaller visual lexical decision megastudies (cf. Table 1.1). I believe that the size of MaltLex will prove sufficient for researchers who wish to study the properties of visual word recognition through the Maltese language (e.g. to investigate the nuances of lexical processing in a language that uses both concatenative and nonconcatenative morphology). In the next sections I conduct several "virtual experiments" using MaltLex data, i.e. where I analyze subsets of the total dataset in lieu of conducting a novel experiment (cf. Keuleers and Balota 2015, Kuperman 2015), in which I aim to replicate findings from prior lexical decision studies on Maltese and other languages in order to demonstrate the validity of the MaltLex dataset.

---

[8] The actual number of responses is 237,476, reflecting four factors: (1) I had planned for participants to judge 500 non-practice items per session, and in the first three sessions List 1 contained 500 non-practice targets. This proved to be too many items for participants to judge and complete the background questionnaire in an hour, and for subsequent sessions I trimmed List 1 to 400 targets and assigned the remaining 100 targets to new lists. This produced a total of 237,500 lexical decision responses (= 500 responses/session * 3 sessions + 400 responses/session * 590 sessions). However, (2) I replaced the word *nittien* 'pervert' and its non-word counterpart with new targets after having collected 3 responses apiece, after one participant conveyed to me that *nittien* could be perceived as offensive. Because I was able to collect few responses to these items, I have omitted them from the final dataset. Additionally, (3) DMDX recorded display issues on two trials, and (4) I eliminated 16 datapoints where participants had judged a given target a second time (Section 1.2.2), bringing the total number of responses to 237,476 (= 237,500 – 6 – 2 – 16).

[9] I collected responses from 12 participants for the list which originally contained *nittien*. Because I replaced *nittien* after having already collected 3 responses, I collected only 9 responses for the replacement word, *referendum* 'referendum', and for its non-word counterpart. For the six targets which were reused across lists, I collected 16−20 responses apiece (*M* = 18.3 responses). Otherwise, I collected 10−13 responses per target (*M* = 10.8 responses).

Following common practice in lexical decision megastudies (e.g. Tse et al. 2017, Soares et al. 2019), I trimmed outliers prior to data analysis by removing datapoints for which the RT was faster than 200 ms (13 datapoints) or slower than 2,500 ms (2,382 datapoints). Then, I calculated the mean and standard deviation (SD) of the RTs for each participant and removed datapoints for which the RT was ±2.5 SDs from each participant's mean RT (6,723 datapoints). This reduced the dataset to 228,358 total datapoints (96.2% of the total dataset), comprising 115,612 lexical decision responses to real-word targets and 112,746 responses to non-word targets. All analyses reported below are based on the trimmed dataset. Within the trimmed dataset, participants were faster to judge real-word targets (*M* = 850 ms, *SEM* = 1.02 ms) than non-word targets (*M* = 990 ms, *SEM* = 1.14 ms) on trials on which they gave the correct response, consistent with previous Maltese visual lexical decision studies (e.g. Twist 2006) and typical of lexical decision studies generally: Possible non-words require an exhaustive "search" of the lexicon to confirm their non-lexicality, and so normally take longer to reject in a lexical decision task than real words require to accept (e.g. Forster 1976). Participants judged both real-word and non-word targets with greater than 90% accuracy, but they were more accurate at judging non-words (94.3%) than real words (90.5%).

1.3 Virtual experiment 1: Analysis of word frequency measures

Perhaps the most robust and well-studied variable that is known to mediate lexical processing is word frequency: All else being equal, readers and listeners more efficiently process words that they have encountered more often than words that they have encountered less often (e.g. Preston 1935, Forster and Chambers 1973, Taft and Hambly 1986, Connine et al. 1990; see Brysbaert et al. 2011, 2017 for review of recent advances in the study of word frequency effects in lexical processing). In lexical decision, participants are faster and more accurate at judging the lexicality of high-frequency words than low-frequency words, and word frequency may explain

30−40% of the variation observed in RTs and response accuracy in lexical decision experiments (Brysbaert et al. 2016, 2017). Of course, researchers do not know how frequent a particular word has been in a given individual's linguistic experience: Instead, we estimate word frequency based on corpus counts, expecting that our estimates will hold across the population of language users.

Word frequency has traditionally been estimated by counting the number of times that a word appears in a corpus (henceforth: WF). However, recent research has found that contextual diversity (CD), a measure that is based on the number of unique documents in a corpus in which a word appears, actually serves as a better predictor of lexical processing efficiency (e.g. Adelman et al. 2006, Adelman and Brown 2008, Brysbaert and New 2009, Keuleers, Brysbaert, and New 2010, Tse et al. 2017). For instance, Brysbaert and New (2009) compared WF and CD measures taken from SUBTLEX-US (a corpus of American English film and television show subtitles, part of the SUBTLEX family of subtitle-based corpora) in analyzing English lexical decision RTs using the ELP dataset (Balota et al. 2007): They found that the SUBTLEX-US CD measures outperformed its WF measures (and the WF measures of other corpora) by accounting for a greater amount of variance in the RT data. Keuleers, Brysbaert, and New (2010) and Tse et al. (2017) similarly observed an advantage for SUBTLEX CD measures over its WF measures in explaining variance in lexical decision RTs using Dutch and Cantonese megastudy data, respectively.

There are two general explanations for the superiority of CD measures over WF measures in explaining performance in lexical decision tasks. For one, it may be that CD better approximates the average language user's linguistic experience than does WF, which may over- or underestimate word exposure to a greater degree than does CD: For instance, a word that occurs frequently but in a limited range of contexts (i.e. high WF, low CD; e.g. as could be the case for certain technical vocabulary) may be experienced frequently by some individuals but remain effectively unknown to others who have not encountered those particular contexts, and CD measures may better account for such differences than do WF ones. Alternatively, it may be that language users actually track the range of contexts in which they experience different words, either instead of tracking the total number of times they have experienced each word (i.e. the word frequency effect actually reflects an effect of contextual diversity) or in addition to tracking WF. By prioritizing words according to the range of contexts in which they have occurred, and treating as more "frequent" those words that have been encountered in a wider range of contexts, language users may be better able to predict which words they will need in future situations, for instance, by ignoring lexical candidates that occur frequently but in a limited range of contexts and thus that one is less likely to need in a future context (see Adelman et al. 2006, Adelman and Brown 2008 for similar explanations).

I leave aside the question of why CD-based measures better explain lexical decision performance than WF measures, and instead aim simply to investigate the predictive ability of WF and CD in MaltLex, using WF and CD norms calculated across Korpus Malti v3.0 (Gatt and Čéplö 2013). If I replicate the previously reported advantage for CD, this will serve as a first validation of the MaltLex dataset for its use in studying visual lexical processing. I also leave aside related questions concerning the appropriate size and composition of the corpus that is used to estimate word frequency: For example, previous studies have found that corpora that are based on different sources may explain word processing performance better for different participant populations, such as corpora based on film and television subtitles for university students (Brysbaert and New 2009) versus more traditional, book-based corpora for older participants (Cuetos et al. 2012). Frequency counts based on corpora larger than 20-million words generally outperform those based on smaller corpora (Brysbaert and New 2009, Brysbaert et al. 2017). Korpus Malti v3.0 comprises approximately 250-million words taken from written sources of a range of genres (e.g. academic, legal, and religious texts; Maltese literature). Likewise, I do not ask whether Korpus Malti WF and

CD measures could be improved by incorporating other measures (e.g. Jones et al. 2012 found that a CD measure that incorporates the semantic redundancy between contexts better accounts for visual word processing performance than do raw CD counts) and instead focus on comparing the WF and CD measures alone as predictors of Maltese visual lexical decision performance.

1.3.1 RT analysis and results

I analyzed RTs to real-word targets on trials on which participants provided the correct response (104,644 responses to 10,951 unique targets) by comparing a series of three maximum likelihood-fitted linear mixed-effects regression (LMER) models using the lme4 package (Bates, Maechler et al. 2015) in R (R Core Team 2021). Models were fitted using the bobyqa optimizer. Each model included log-transformed RT as the dependent variable, participant and target as random effects, and the following control variables as fixed effects: log-transformed word frequency-weighted orthographic neighborhood density (I added 1 to each value prior to applying the log transformation to account for targets having a value of 0); log-transformed target length (in number of letters); participant's age; participant's trial number; participant's overall session number; and participant's session number within a given day. Critically, the three models differed in terms of which Korpus Malti v3.0 frequency measure they included as an additional fixed effect:

- no frequency measure;
- log-transformed WF (WF-only model);
- log-transformed CD (CD-only model);

I added random slopes for log-transformed WF by-participants in the WF-only model, and log-transformed CD in the CD-only model: The results of a series of likelihood ratio tests comparing each model to its random intercepts counterpart supported their inclusion for the WF-only ($\chi^2(2) = 536.49$, $p < 0.001$) and CD-only models ($\chi^2(2) = 503.49$, $p < 0.001$) (Bates, Kliegl et al. 2015).

To compare models, I analyzed each model's Akaike Information Criterion (AIC): AIC represents a composite measure of a model's goodness-of-fit (i.e. the amount of explained variance in the fitted dataset) and complexity. Among a set of models fitted to the same dataset, the model that has the lowest AIC has the best combination of fit and complexity, as it minimizes the loss in explained variance relative to the "true" model of the dataset. Models with higher AIC values have a worse combination of fit and complexity: While adding parameters necessarily improves the fit of a model, if the increase in explained variance does not offset the increase in model complexity, AIC increases (Akaike 1973, Burnham and Anderson 2004). I compare the likelihood of each model using the formula: $\exp(-\Delta_i/2)$, where Δ_i represents the difference in AIC between model *i* and the model with the lowest AIC (i.e. $\Delta_i = AIC_i - AIC_{min}$). This indicates the probability that model *i* minimizes information loss relative to the model with the lowest AIC value, indicating the likelihood of model *i* relative to the best-fitting model (Burnham and Anderson 2004: 270-272). To then assess the significance of fixed effects within the best-fitting model, I used the lmerTest package (Kuznetsova et al. 2017) to simulate Satterthwaite approximations for degrees of freedom.

The results of model comparisons are presented in Table 1.2. Unsurprisingly, both the WF-only model ($\Delta = 3{,}213.64$, $p < 0.001$) and the CD-only model outperformed the model that did not include any measure of target frequency as a fixed effect ($\Delta = 3{,}271.12$, $p < 0.001$). However, the CD-only model outperformed the WF-only model, exhibiting a significantly better combination of goodness-of-fit and model complexity ($\Delta = 57.93$, $p < 0.001$). Thus, CD alone also outperforms WF alone in explaining the variance in visual lexical decision RTs within the MaltLex dataset.

All fixed effects that were included in the CD-only model were significant: Unsurprisingly, participants responded faster as CD increased: More frequent targets elicited faster RTs than did less frequent targets ($\hat{\beta} = -0.03964$; $t(139.8) = -28.82$, $p < 0.001$). Participants were faster to respond as neighborhood density increased ($\hat{\beta}$ = –0.00165; $t(10{,}400) = -2.64$, $p < 0.01$) but slower as target length increased ($\hat{\beta}$ = 0.16430; $t(10{,}080) = 34.47$, $p < 0.001$). RTs increased along with participant's age ($\hat{\beta}$ = 0.00443; $t(267.8) = 3.41$, $p < 0.001$). Participants were slower to respond across trials, likely reflecting an effect of fatigue ($\hat{\beta}$ = 0.00008; $t(100{,}500) = 11.35$, $p < 0.001$), but faster to respond across sessions, reflecting an effect of practice and increasing familiarity with the lexical decision task ($\hat{\beta}$ = –0.00174; $t(88{,}860) = -11.82$, $p < 0.001$). Likewise, participants were faster to respond across sessions within the same day ($\hat{\beta}$ = –0.01459; $t(101{,}000) = -7.97$, $p < 0.001$).

1.3.2 Accuracy analysis and results

I analyzed response accuracy to real-word targets (115,612 responses to 11,040 unique targets) by comparing a series of three generalized linear mixed-effects regression (GLMER) models fitted using the binomial logit link function and the bobyqa optimizer in R. Each model included response accuracy as the dependent variable (0 = incorrect, 1 = correct); log-transformed target length and the participant's within-day session number as fixed effects; and participant and target as random effects. As in the RT analysis, the models differed in whether they also included

Table 1.2. Results of model comparisons for the analysis of word frequency measures. The models are organized in increasing order of AIC, and the Δ- and *p*-values are based on comparison of the model's AIC value with that of the best-fitting model.

	RT			Accuracy		
	AIC	Δ	*p*	AIC	Δ	*p*
CD-only model	18,167.48	-----	-----	50,539.43	-----	-----
WF-only model	18,225.41	57.93	< 0.001	50,743.47	204.03	< 0.001
No frequency measure	21,438.64	3,271.12	< 0.001	53,725.63	3,186.20	< 0.001

no measure of target frequency, log-transformed WF (WF-only model), or log-transformed CD (CD-only model) as a fixed effect. Models that included additional fixed effects (e.g. for neighborhood density or overall session number) either failed to converge or failed to improve model fit; the fixed effects structure used here results in the best-fitting set of models that converged.[10] I included random slopes for log-transformed WF by-participants in the WF-only model ($\chi^2(2) = 97.63$, $p < 0.001$) and for log-transformed CD by-participants in the CD-only model, which the results of likelihood ratio tests comparing each model with its random-intercepts counterpart model support for this dataset ($\chi^2(2) = 96.61$, $p < 0.001$) (Bates, Kliegl et al. 2015).

The results of model comparisons are presented in Table 1.2. Once again, both the WF-only model ($\Delta = 2{,}982.16$, $p < 0.001$) and the CD-only model outperformed the model that did not include a measure of target frequency as a fixed effect ($\Delta = 3{,}186.20$, $p < 0.001$), and the CD-only model outperformed the WF-only model ($\Delta = 204.03$, $p < 0.001$). All fixed effects included in the CD-only model were significant: Participants responded more accurately as CD increased ($\hat{\beta} = 0.60701$; $z = 37.55$, $p < 0.001$) and as target length increased ($\hat{\beta} = 2.14285$; $z = 24.25$, $p < 0.001$). Accuracy decreased across sessions within the same day ($\hat{\beta} = -0.11315$; $z = -4.08$, $p < 0.001$).

1.3.3 Discussion

Consistent with earlier analyses of visual lexical decision megastudy data (e.g. Brysbaert and New 2009, Keuleers, Brysbaert, and New 2010, Tse et al. 2017), the results of the RT and response accuracy analyses both support the hypothesis that a measure of frequency that is based on contextual diversity, rather than token word frequency counts, better predicts Maltese lexical processing performance as measured by visual lexical decision. Given the unavailability of other

[10] Convergence issues are not unusual for a generalized linear mixed-effects regression analysis. Further, the overall high accuracy rate (90.5%) leaves little variation for the models to explain.

large Maltese corpora (e.g. one based on film and television subtitles), I only compared WF and CD measures taken from the 250-million-word Korpus Malti v3.0 (Gatt and Čéplö 2013), but I believe that the overall size of the corpus supports the validity of these measures (Brysbaert and New 2009, Brysbaert et al. 2017). In subsequent analyses I include CD, rather than WF, as a predictor of lexical decision performance, and I recommend that other researchers do the same.

1.4 Virtual experiment 2: Analysis of lexical stratum

One of the most interesting properties of Maltese is its split lexicon: Roughly half of all Maltese words are of Semitic origin, reflecting Maltese's beginnings as a Maghrebi Arabic dialect, while the other half comprise borrowings from non-Semitic languages (namely Sicilian, Italian, and English), reflecting a millennium of contact with these languages as well as widespread multilingualism among Maltese speakers (Bovingdon and Dalli 2006, Brincat 2011, Comrie and Spagnol 2016). To be able to address processing differences between native and borrowed Maltese words, I included both Semitic and non-Semitic words as targets in the visual lexical decision megastudy, yet it is worth acknowledging that words of the two strata are, in many ways, not equal. For example, function words and other core vocabulary are overwhelmingly of Semitic origin; Maltese speakers have primarily borrowed content words (Comrie and Spagnol 2016). Among the MaltLex target set, words of the two strata differ in that Semitic words are reliably shorter, more frequent, and inhabit denser orthographic neighborhoods than non-Semitic words (Section 1.2.2).

A priori there is no reason to expect processing differences between Semitic and non-Semitic Maltese words that are due to their etymologies and not to more basic differences between words of the two strata, for instance in their semantic, morphological, or orthotactic characteristics. Yet studies that explore the processing of non-Semitic Maltese words are scarce, and in a recent visual masked priming study Geary and Ussishkin (2018) found that readers were faster to judge the lexicality of Semitic Maltese words (*M* = 652 ms) than non-Semitic ones (*M* = 682 ms), as well as more accurate at judging Semitic words (94.5% overall accuracy rate) than non-Semitic ones (88.5% overall accuracy rate). The 96 Semitic and non-Semitic targets used in their study were matched according to length and word frequency, and a post-hoc analysis failed to indicate that the two sets of targets differed reliably in orthographic neighborhood density, a third factor which could have explained this difference.[11] Here, I attempt to replicate this finding by comparing lexical decision performance for Semitic versus non-Semitic targets using data from MaltLex, assessing whether Semitic and non-Semitic Maltese words exhibit differences in RTs or response accuracy when other lexical characteristics, such as target length and frequency, are accounted for.

1.4.1 RT analysis and results

I analyzed RTs to real-word targets of Semitic (6,445 unique targets) and non-Semitic origin (4,413 unique targets) on trials on which participants provided the correct response (103,848 total responses) using an LMER analysis: I fitted a restricted maximum likelihood-fitted LMER model in R (R Core Team 2021) using the lme4 package (Bates, Maechler et al. 2015). I fitted the model using the bobyqa optimizer, and I assessed the significance of fixed effects using the lmerTest package to simulate Satterthwaite approximations for degrees of freedom (Kuznetsova et al. 2017). The model included log-transformed RT as the dependent variable, participant and target as random effects, and the following variables as fixed effects: target lexical stratum (levels: Semitic versus non-Semitic; reference: Semitic), log-transformed target CD, log-transformed word

[11] Geary and Ussishkin (2018) obtained frequency values from Korpus Malti v2.0 (Borg et al. 2012), a 130-million word version of Korpus Malti. I obtained WF and CD counts for the 96 targets from Korpus Malti v3.0 (Gatt and Čéplö 2013), as well as recalculated neighborhood density, and compared these values across lexical strata in a series of Welch's unequal variances *t*-tests: Likewise, I did not find that the targets differed significantly in WF ($t(91.2) = 0.46$, *n.s.*), CD ($t(93.3) = 0.04$, *n.s.*), or word frequency-weighted orthographic neighborhood density ($t(82.1) = -0.07$, *n.s.*), consistent with these variables not underlying performance differences for Semitic versus non-Semitic targets.

frequency-weighted target neighborhood density, log-transformed target length, participant's age, participant's trial number, participant's overall session number, and participant's session number within a given day. I also included random slopes for lexical stratum by-participants: The results of a likelihood ratio test comparing this model with the random intercepts model were significant ($\chi^2(2) = 218.78$, $p < 0.001$), justifying their inclusion for this dataset (Bates, Kliegl et al. 2015).[12]

The results of this analysis are consistent with those reported in Section 1.3.1: Participants were faster to respond both as CD increased ($\hat{\beta} = -0.03755$; $t(10{,}980) = -54.46$, $p < 0.001$) and as neighborhood density increased ($\hat{\beta} = -0.00171$; $t(10{,}310) = -2.72$, $p < 0.01$), but slower to respond as length increased ($\hat{\beta} = 0.16790$; $t(10{,}060) = 33.36$, $p < 0.001$). RTs increased with participant's age ($\hat{\beta} = 0.00524$; $t(279.4) = 3.87$, $p < 0.001$). Participants were slower to respond across the trials of a session ($\hat{\beta} = 0.00008$; $t(99{,}830) = 11.13$, $p < 0.001$), but faster across sessions both overall ($\hat{\beta} = -0.00173$; $t(88{,}110) = -11.71$, $p < 0.001$) and within the same day ($\hat{\beta} = -0.01447$; $t(100{,}200) = -7.86$, $p < 0.001$). However, the effect of lexical stratum was not significant ($\hat{\beta} = 0.00018$; $t(157.3) = 0.04$, *n.s.*): Despite a 5 ms advantage in mean RT for Semitic targets ($M = 848$ ms) compared to non-Semitic targets ($M = 853$ ms), I did not find that this difference was statistically reliable. That is, I did not replicate Geary and Ussishkin's (2018) RT advantage for non-Semitic words.

1.4.2 Accuracy analysis and results

I analyzed response accuracy to real-word targets of Semitic (6,514 unique targets) and non-Semitic origin (4,429 unique targets; 114,609 total responses) using a GLMER analysis. I fitted a GLMER model using the binomial logit link function and the bobyqa optimizer, using the lme4 package in R. The model included response accuracy as the dependent variable (0 = incorrect,

[12] The results of the random intercepts model were qualitatively identical in that the effect of lexical stratum failed to reach significance ($\hat{\beta} = -0.00410$; $t(10{,}040) = -1.50$, *n.s.*). All other fixed effects remained significant.

1 = correct); lexical stratum, log-transformed CD, log-transformed target length, and participant's within-day session number as fixed effects; and participant and target as random effects. I added random slopes for lexical stratum by-participants, as supported by the results of a likelihood ratio test comparing this model with the random intercepts model ($\chi^2(2) = 183.9$, $p < 0.001$).

The results of this analysis are consistent with those reported in Section 1.3.2: Participants responded more accurately as CD increased ($\hat{\beta} = 0.67756$; $z = 54.02$, $p < 0.001$) and as target length increased ($\hat{\beta} = 2.44638$; $z = 26.90$, $p < 0.001$), while responding less accurately across sessions within the same day ($\hat{\beta} = –0.11569$; $z = –4.15$, $p < 0.001$). The effect of lexical stratum was also significant, but while participants responded more accurately to non-Semitic words (91.7%) than to Semitic words in terms of raw accuracy rate (89.9%), the model revealed that, when these other variables were accounted for, participants actually responded less accurately to non-Semitic than Semitic words ($\hat{\beta} = –0.76557$; $z = –8.72$, $p < 0.001$), consistent with Geary and Ussishkin (2018).

1.4.3 Virtual experiment 2b: Replication of Geary and Ussishkin (2018)

Sixty of Geary and Ussishkin's (2018) 96 real-word targets were used as targets in the MaltLex megastudy (38 Semitic and 22 non-Semitic targets). To further assess possible processing differences between Semitic and non-Semitic Maltese words, I conducted a virtual replication of their study by analyzing RTs and response accuracy to those targets using the MaltLex dataset.

I analyzed RTs to real-word targets of Semitic and non-Semitic origin on trials on which participants provided the correct response (617 total responses) by fitting a restricted maximum likelihood-fitted LMER model using the bobyqa optimizer. The model included log-transformed RT as the dependent variable; lexical stratum (levels: Semitic versus non-Semitic; reference: Semitic), trial number, overall session number, and session number within a given day as fixed effects (adding further fixed effects for neighborhood density, target length, or participant's age failed to improve model fit, reflecting the lower power afforded by the size of this dataset); and participant and target as random effects. A model with a more complex random effects structure including random slopes for lexical stratum by-participant failed to converge; I report the results of the random intercepts model. Participants were faster to respond as CD increased ($\hat{\beta} = -0.05018$; $t(51.8) = -5.70$, $p < 0.001$) and across sessions overall ($\hat{\beta} = -0.00632$; $t(592.8) = -3.14$, $p < 0.005$), but slower to respond across the trials of a session ($\hat{\beta} = 0.00032$; $t(567.4) = 3.64$, $p < 0.001$) and across sessions within the same day ($\hat{\beta} = 0.06113$; $t(548.2) = 2.11$, $p < 0.05$). Although participants were faster overall to judge non-Semitic targets ($M = 771$ ms) than Semitic targets ($M = 809$ ms), the effect of lexical stratum failed to reach significance ($\hat{\beta} = -0.00602$; $t(52.4) = -0.22$, *n.s.*).

I analyzed response accuracy to real-word targets of Semitic and non-Semitic origin (653 datapoints) by fitting a GLMER model using the binomial logit link function and the bobyqa optimizer in R, using the lme4 package. The model included response accuracy as the dependent variable (0 = incorrect, 1 = correct); lexical stratum and log-transformed CD as fixed effects (including other variables as fixed effects either failed to improve model fit or caused the model to fail to converge); and participant and target as random effects. I also included random slopes for lexical stratum by-participant, which the results of a likelihood ratio test suggest were justified for this dataset ($\chi^2(2) = 7.36$, $p < 0.05$). Participants responded more accurately as CD increased ($\hat{\beta} = 0.9315$; $z = 2.39$, $p < 0.05$), and less accurately to non-Semitic words (88.9% overall accuracy rate) compared to Semitic words (97.8% overall accuracy rate; $\hat{\beta} = -6.1119$; $z = -3.01$, $p < 0.005$).

1.4.4 Discussion

My findings with respect to lexical stratum are mixed: I failed to replicate a processing speed advantage for Semitic Maltese words compared to non-Semitic ones in the overall analysis, and in analyzing the 60 words that recurred as targets in Geary and Ussishkin (2018; further, the overall difference in RTs favored non-Semitic words). The lack of an effect in the latter analysis may simply reflect a lack of power afforded by the smaller size of the dataset (617 datapoints). However, the power afforded by the size of the overall dataset (103,851 datapoints) seemingly should have been sufficient to detect such an effect if it exists: I tentatively conclude that there does not exist a general processing speed advantage for Semitic Maltese words that is due to their lexical stratum, and not to some confounding variable (e.g. target frequency, length), and instead encourage further research which considers more nuanced differences between words of the two strata, such as how consonant letters (which may correspond to a root morpheme) constrain lexical processing differently for Semitic versus non-Semitic words (Anderson and Geary 2019).

I did find an advantage for Semitic words over non-Semitic words in terms of response accuracy in both analyses, which is difficult to explain. One possibility was raised by participants, who, in post-session conversations with the researcher and other participants, would occasionally discuss whether certain words "should" be considered Maltese or not due to their status as English (and, less frequently, Italian) loanwords. These arguments tapped into participants' knowledge of English, and reflected their concerns over the influence of English across the Maltese archipelago (e.g. the possibility that English would supplant Maltese) and on the Maltese language. It could be that a more conscious decision process led participants to reject some non-Semitic Maltese words (as determined by linguists and lexicographers; e.g. Aquilina 1987–1990, Camilleri 2013) at a higher rate in the lexical decision task, though this taps into sociolinguistic factors which go beyond the scope of the present study (crucially, not all loanwords were deemed "inappropriate" to consider as Maltese words). However, Geary and Ussishkin (2018) observed RT and response accuracy advantages for Semitic Maltese words despite having avoided using English loanwords as non-Semitic targets, and so their results cannot be explained by a conscious, sociolinguisticssensitive decision process disfavoring English-origin Maltese words in a lexical decision task.[13] Alternatively, some other factor may mediate the processing of Semitic and non-Semitic words, and in the next section I consider how multilingualism influences word recognition in Maltese, focusing on differences in participants' relative dominance between Maltese and English.

1.5 Virtual experiment 3: Analysis of cognate status and language dominance

All MaltLex participants were bilingual in Maltese and English, reflecting widespread bilingualism stemming from two centuries of English being spoken in the Maltese archipelago: This started when France's brief occupation ended and Malta became a protectorate of the British Empire in 1800, and has lasted into the present day where both Maltese and English have official status in the Republic of Malta. Many speakers, including some participants, are also multilingual in Italian, reflecting continued exposure, for instance, via Italian-language media (Brincat 2011). Yet speakers of the same languages vary in the extent to which they use each language, with some being more dominant in one language than the other(s), and I had participants complete the BiLP (Birdsong et al. 2012) to provide a composite, continuous measure of language dominance. Indeed, some participants proved to be English-dominant, though most were more Maltese-dominant.

In general, all of a multilinguals' languages remain simultaneously active and compete for cognitive resources during language processing, even when only one language is appropriate in the current speech context (e.g. in completing a monolingual lexical decision task). For instance, lexical processing is easier for words in one language that have a cognate in the other language compared to words that lack a cognate (a "cognate" is defined here as a word that shares aspects of its form and meaning with the target, reflecting that the two languages inherited and/or borrowed

[13] I can recall one target that prompted some of these discussions: *gowl* 'goal (in football)', which is transparently the English word *goal* borrowed and adapted to the Maltese orthography. However, *gowl* and its plural form *gowls* exhibited overall accuracy rates of 90% and 80% across the study, respectively, inconsistent with this hypothesis.

these words from a common source), reflecting the simultaneous activation of the target word and its cognate, the latter of which facilitates access to the target word relative to other words in the same lexicon. This "cognate advantage" results in faster RTs to words having a cognate in lexical decision and naming tasks, even when the other language is irrelevant to the task at hand (Costa et al. 2000, Dijkstra et al. 1999). Crucially, language dominance has been found to mediate the cognate advantage: These effects are stronger when participants perform in their non-dominant language than in their dominant language (Jared and Kroll 2001, Poort and Rodd 2017). The non-Semitic stratum of the Maltese lexicon is rife with words having a cognate in English, whether because Maltese and English borrowed from comparable sources (e.g. English *piano* and Maltese *pjanu* were borrowed from Italian *piano*) or because Maltese borrowed directly from English (e.g. Maltese *xelter* 'shelter' was borrowed from English *shelter*; Aquilina 1987–1990). Based on earlier research, I expect more English-dominant MaltLex participants to outperform more Maltese-dominant MaltLex participants on non-Semitic targets that have a cognate in English.

In this analysis, I aim to replicate the cognate advantage on lexical decision performance in Maltese using data from MaltLex. I use lexical stratum as a proxy for cognate status because non-Semitic Maltese words often have a cognate in English while relatively few Semitic Maltese words do, and I compare lexical decision performance for Semitic versus non-Semitic words for English-dominant participants versus Maltese-dominant participants. To minimize differences in the composition of the Semitic and non-Semitic strata of the Maltese lexicon (e.g. most function words are of Semitic origin) and to increase the likelihood that a given non-Semitic target has a transparent cognate in English (and because there are currently no electronic databases that would allow me to automate tagging MaltLex targets for cognate status), I analyze only singular nouns. Such words bear minimal morphology except for the feminine singular suffix -*a* (cf. plural nouns take a wide range of plural suffixes and morphological patterns; likewise, uninflected and inflected verbs may differ considerably in form in Maltese). Although this rough approximation of cognate status is limited (e.g. some non-Semitic words that do not have a transparent English cognate are included in the dataset analyzed here, and it is difficult to account for individual variation in the English cognates that participants actually know), future research may refine this analysis.

In the following analyses, I test for an interaction between lexical stratum/cognate status and language dominance: Finding that more English-dominant participants exhibit a greater RT advantage for non-Semitic words compared to Semitic words than do more Maltese-dominant participants would be consistent with the cognate advantage (which, to my knowledge, has yet to be demonstrated with Maltese-English bilinguals) and so further elucidate processing differences between words of the two etymological strata (albeit one which is not due to stratum itself).

1.5.1 RT analysis and results

I analyzed RTs to real-word singular noun targets of Semitic and non-Semitic origin on trials on which participants provided the correct response (24,106 total responses, including 9,068 responses to 962 Semitic-origin Maltese targets and 15,038 responses to 1,546 non-Semitic-origin targets) using an LMER analysis: I fitted a restricted maximum likelihood-fitted LMER model in R (R Core Team 2021) using the lme4 package (Bates, Maechler et al. 2015). I fitted the model using the bobyqa optimizer. I assessed the significance of fixed effects by simulating Satterthwaite approximations for degrees of freedom using lmerTest (Kuznetsova et al. 2017).

The model included log-transformed RT as the dependent variable, and participant and target as random effects. As fixed effects, the model included lexical stratum/cognate status (levels: Semitic/non-cognate versus non-Semitic/cognate; reference: Semitic/non-cognate), participant's language dominance score, and the lexical stratum by language dominance interaction. Including participants' language dominance scores allows me to assess whether there is a general effect of relative language dominance on lexical processing speed, while including the interaction of lexical stratum by language dominance allows me to assess whether differences in language dominance mediate processing speed differences for Semitic versus non-Semitic Maltese words. As additional fixed effects, the model included the following control variables: log-transformed target CD, log-transformed word frequency-weighted neighborhood density, log-transformed target length, participant's age, trial number, overall session number, and session number within a given day.

I also included random slopes for lexical stratum by-participants: The results of a likelihood ratio test comparing this model and the random intercepts model suggest that they are justified ($\chi^2(2) = 57.79$, $p < 0.001$). Including random slopes for dominance by-targets caused the model to fail to converge, and so I do not analyze models with more complex random effects structures.

As in the overall analysis of lexical stratum (Section 1.4.1), participants responded faster as CD increased ($\hat{\beta} = -0.03893$; $t(2{,}510) = -23.71$, $p < 0.001$), but slower as target length increased ($\hat{\beta} = 0.15090$; $t(2{,}269) = 14.50$, $p < 0.001$). Participants responded slower across the trials of a given session ($\hat{\beta} = 0.00009$; $t(22{,}920) = 6.12$, $p < 0.001$), but faster both across sessions overall ($\hat{\beta} = -0.00146$; $t(20{,}350) = -4.86$, $p < 0.001$) and across sessions within the same day ($\hat{\beta} = -0.01662$; $t(23{,}120) = -4.34$, $p < 0.001$). However, in this analysis of a smaller and more restricted subset of the MaltLex dataset, the effects of orthographic neighborhood density ($\hat{\beta} = 0.00024$; $t(2{,}329) = 0.19$, *n.s.*) and participant's age failed to reach significance ($\hat{\beta} = 0.00247$; $t(146.1) = 1.60$, *n.s.*)

The effects of both lexical stratum/cognate-status ($\hat{\beta} = -0.00770$; $t(105.6) = -0.86$, *n.s.*) and language dominance failed to reach significance ($\hat{\beta} = -0.00035$; $t(103.0) = -0.89$, *n.s.*). However, the interaction of lexical stratum by language dominance was significant ($\hat{\beta} = 0.00032$; $t(62.9) = 2.14$, $p < 0.05$): As participants became more Maltese-dominant, they responded slower to non-Semitic targets that may have a cognate in English compared to Semitic targets. In other words, Semitic and non-Semitic Maltese targets induce less of a difference in RTs for the more English-dominant participants, which is consistent with previous findings on the cognate advantage.

1.5.2 Accuracy analysis and results

I analyzed response accuracy to singular noun targets of Semitic and non-Semitic origin (26,605 total responses, including 10,249 responses to 973 Semitic targets and 16,356 responses to 1,548 non-Semitic targets) by fitting a GLMER model using the binomial logit link function and the bobyqa optimizer. The model included response accuracy as the dependent variable (0 = incorrect, 1 = correct); participant and target as random effects; and lexical stratum/cognate status (levels: Semitic/non-cognate versus non-Semitic/cognate; reference: Semitic/non-cognate), participant's language dominance category (levels: English-dominant versus Maltese-dominant; reference: English-dominant), and the interaction of lexical stratum by participants' language dominance category as fixed effects. To circumvent convergence issues introduced by the original, continuous dominance score variable, I converted dominance score to a categorical variable by binning the 52 participants with a dominance score below the median of 36.735 into an "English-dominant" group, and the 52 participants with a dominance score above the median into a "Maltese-dominant" group. As additional fixed effects, the model included the following control predictors: log-transformed CD, log-transformed target length, and the within-day session number.

I added random slopes for lexical stratum by-participants, which were supported by the results of a likelihood ratio test comparing this model with the random intercepts model ($\chi^2(2)$ = 30.22, p < 0.001). Models with more complex random effects structures (e.g. including random slopes for dominance category by-targets) failed to converge, so I do not analyze them here.

Consistent with prior analyses, participants responded more accurately as CD increased ($\hat{\beta}$ = 0.83294; z = 24.94, p < 0.001) and as target length increased ($\hat{\beta}$ = 1.09542; z = 5.58, p < 0.001), but less accuracy across sessions within the same day ($\hat{\beta}$ = –0.14497; z = –2.48, $p < 0.05$). The main effect of dominance category was significant ($\hat{\beta}$ = 0.57638; z = 2.89, $p < 0.005$): The Maltese-dominant participants responded more accurately to Semitic targets (90.7% overall accuracy rate) compared to the English-dominant participants (87.1% overall accuracy rate). However, neither the main effect of lexical stratum/cognate status ($\hat{\beta}$ = –0.21312; z = –1.27, *n.s.*) nor the interaction of lexical stratum by dominance category reached significance ($\hat{\beta}$ = –0.26986; z = –1.41, *n.s.*).

1.5.3 Discussion

Analyzing a subset of the MaltLex dataset, namely singular nouns, I found that MaltLex participants who were more English-dominant exhibited a smaller RT advantage in judging the lexicality of Semitic Maltese words compared to non-Semitic Maltese words than did Maltese-dominant participants. In other words, the more English-dominant participants, performing lexical decision in their non-dominant language, experienced less difficulty in judging the lexicality of non-Semitic words compared to Semitic words than did their Maltese-dominant counterparts. Taking lexical strata as a rough approximation of cognate status (since the non-Semitic stratum of the Maltese lexicon contains many words that have cognates in English, while the Semitic stratum contains few such words), these results are consistent with prior research showing an advantage for cognates in lexical processing when participants perform in their non-dominant language (e.g. Costa et al. 2000, Dijkstra et al. 1999, Jared and Kroll 2001, Poort and Rodd 2017). Still, future research on Maltese should take a more nuanced approach to cognatehood, as a cognate advantage may be found within the non-Semitic stratum when the processing of non-Semitic words having an English cognate is compared with that of non-Semitic words lacking such a cognate. Ultimately, my results are not consistent with prior research finding a RT advantage for Semitic Maltese words

(Geary and Ussishkin 2018), but they do suggest that language dominance is an important factor mediating the processing of words of the two strata which has been overlooked by prior studies.

I found that more Maltese-dominant participants were more accurate at judging Semitic-origin Maltese words than were English-dominant participants. However, I did not find that the effect of lexical stratum on response accuracy differs for English-dominant and Maltese-dominant participants, which I believe is also consistent with research on the cognate advantage (which, to my knowledge, is typically found for lexical processing speed, not accuracy): The activation of an English cognate may help activate the Maltese target and so speeds its recognition, but as long as participants know the Maltese target they should ultimately be able to recognize it and judge it accordingly, and so the presence of a cognate form should not affect response accuracy.

1.6 General discussion

The analyses performed here have illuminated the roles that three variables play in the processing of Maltese words: The frequency at which language users have encountered individual words, the etymology of different words, and Maltese-English bilingual speakers' relative levels of language dominance. None of these results are particularly surprising in light of prior research: Rather, the fact that I was able to replicate established findings from the lexical decision literature supports the validity of the MaltLex dataset for studying visual word recognition in Maltese.

I found that a measure of frequency that is based on the number of unique documents in Korpus Malti v3.0 (Gatt and Čéplö 2013) in which a word appears (CD) serves as a better predictor of lexical processing efficiency, as measured by both RTs and response accuracy in a visual lexical decision task, than does a more traditional measure based on the total number of times that a word appears in the corpus (WF). This result is consistent with prior analyses of visual lexical decision megastudy data for English (Brysbaert and New 2009), Cantonese (Tse et al. 2017), and Dutch

(Keuleers, Brysbaert, and New 2010), and I suggest that Maltese researchers who study lexical processing should attend to words' Korpus Malti CD values rather than to their WF estimates.

Contrary to Geary and Ussishkin (2018), I did not observe a general RT advantage for Semitic Maltese words compared to non-Semitic words, though I did observe an advantage for Semitic words in terms of response accuracy. Although all MaltLex participants self-identified as Maltese-English bilinguals, I found that participants differed in their relative levels of language dominance, with some being more English-dominant and others being strongly Maltese-dominant: I found that Maltese-dominant participants judged Semitic-origin targets more accurately than did English-dominant participants, reflecting their having greater use of the Maltese language. Furthermore, I found that more English-dominant participants exhibited a smaller difference in RTs for Semitic versus non-Semitic targets than did more Maltese-dominant participants (both groups responded slower numerically to non-Semitic targets compared to Semitic targets, although the main effect of lexical stratum did not reach significance in this analysis). This may reflect the fact that many non-Semitic Maltese words, and few Semitic words, have cognates in English, and is consistent with prior research on the cognate advantage in lexical processing: When performing in their non-dominant language, multilinguals are usually faster to process words having a cognate in their dominant language than words lacking such a cognate, owing to interlingual coactivation that ultimately facilitates recognition of the non-dominant target form (e.g. Jared and Kroll 2001). These results underscore the need for Maltese researchers to account for individual differences in language use/proficiency and suggest that language processing megastudies that use multilingual participant populations should collect information about participants' language dominance.

In the chapters that follow, I continue to validate the MaltLex dataset by replicating other established findings from the visual lexical decision literature. I explore how the orthographic form of a word mediates lexical processing in Chapter 2, and the role of consonant letters in Chapter 3.

Chapter 2

Orthographic similarity effects in a Maltese visual lexical decision megastudy

2.1 Introduction[14]

In Chapter 1, I described the development and composition of the MaltLex database of Maltese visual lexical decision responses, and I used data from MaltLex to investigate how two factors – (1) how frequently participants have encountered a given word as estimated using corpus measures, and (2) whether an individual word is of Semitic or non-Semitic origin – impact lexical processing performance in Maltese. In particular, I found (1) that a measure of word frequency that is based on the number of documents in which a word appears in a corpus (called "contextual diversity") outperforms a more traditional measure that is based on the total number of times that a word appears in a corpus ("word frequency") in predicting visual lexical decision response times and response accuracy, consistent with previous studies that focused on English (e.g. Brysbaert and New 2009), Dutch (Keuleers, Brysbaert, and New 2010), and Cantonese (Tse et al. 2017). I also found (2) that English-dominant Maltese readers (all of whom were bilingual speakers of Maltese and English) are faster to judge the lexicality of non-Semitic-origin Maltese words than Semitic words, but this advantage diminishes as participants become more Maltese-dominant.

In Chapter 1, I also found that the length of a given word, in number of letters, correlates positively with response times (RTs) and response accuracy in the MaltLex dataset: Participants respond slower to real-word targets as they increase in length, most likely due to the fact that more information must be processed in order to confirm the target's identity as a real word; and they respond more accurately to real-word targets as they increase in length, perhaps reflecting a speed-accuracy tradeoff. In this chapter, I continue to explore how the orthographic form of a target

[14] I presented an earlier version of this chapter at the 2021 Words in the World International Conference (WOW2021).

impacts lexical processing in Maltese, focusing on how targets may resemble other Maltese words. The analyses presented here inform our understanding of lexical processing in Maltese, a language that has a unique combination of orthographic and morphological properties that may mediate the effects of orthographic form on lexical processing, while also validating the MaltLex dataset by replicating effects on visual word recognition that are commonly found in lexical decision.

Prior research has repeatedly shown that the number of words that are orthographically similar to a given target, called its orthographic "neighbors", strongly influences lexical processing performance for that target as measured in visual lexical decision (e.g. Andrews 1989, 1992, 1997; Balota et al. 2004; Keuleers, Diependaele, and Brysbaert 2010). An orthographic neighbor may include other words that differ from the target by the substitution, insertion, or deletion of a letter, or by the transposition of two adjacent letters (e.g. *train*, *trails*, *rail*, and *trial* represent neighbors of *trail*). All else being equal, participants tend to judge the lexicality of words that occupy "high-density" neighborhoods (i.e. that have many neighbors) faster and more accurately than words that occupy "low-density" neighborhoods (i.e. that have few neighbors). Indeed, I found such effects of orthographic neighborhood density in the analyses reported in Chapter 1: MaltLex participants responded faster and more accurately to real-word targets as their neighborhood density increased.

However, prior research has also found that other lexical variables, namely target lexicality and word frequency, mediate the effects of neighborhood density on lexical decision performance. Whereas orthographic neighborhood density has a generally facilitatory effect for real-word targets in visual lexical decision, it has an inhibitory effect for non-word targets: All else being equal, participants tend to judge the lexicality of high-density non-words slower and less accurately than low-density non-words, possibly due to increased competition with potential real-word candidates (e.g. Andrews 1989, Balota et al. 2004, Carreiras et al. 1997, Coltheart et al. 1977, Forster and Shen 1996, Hendrix and Sun 2021, Keuleers, Diependaele, and Brysbaert 2010, Yap et al. 2015).

On the other hand, the facilitatory orthographic neighborhood density effect for real-word targets has been found to diminish as targets increase in frequency, with high-frequency targets that are of low-density versus high-density exhibiting no processing difference (e.g. Andrews 1989, 1992; Lim 2016; Sears et al. 1995). I attempt to replicate these findings using the MaltLex dataset in the analyses presented below, by assessing how orthographic neighborhood density interacts with the lexicality of the target (Section 2.3) and with target frequency for real-word targets (Section 2.4).

In Section 2.5, I consider what it means for two words to be orthographically similar, focusing on similarities between pairs of letters of the Maltese alphabet. Namely, Maltese has three pairs of letters that are identical except in the omission or inclusion of a diacritic: "g" and "ġ", "h" and "ħ", and "z" and "ż".[15] Due to the prevalence of English keyboards that lack diacritic-bearing letters, Maltese writers often omit these diacritics (e.g. using "g" instead of "ġ", but not vice versa; cf. Mifsud and Mitterer 2020). During the post-session debriefings, participants mentioned having noticed that some non-word targets resembled real words except in whether a diacritic had been omitted from or added to a letter, and they reported that such similarities had made rejecting these non-words in the lexical decision task especially difficult. I test this observation here, assessing whether diacritic-based word-similarity has an inhibitory effect on lexical decision performance for non-word targets that goes beyond the standard orthographic neighborhood density effect.

Maltese is unique among Semitic languages in that it is written using the Latin alphabet, allowing for a more-straightforward comparison with orthographic similarity effects in languages in which orthographic neighborhood density has been previously explored. Maltese is also unique among the languages of prior study in that it uses nonconcatenative morphology that is typical of Semitic languages: Because changes to a word's internal letters are likely to alter its morphological

[15] An additional letter, "ċ", includes a diacritic. While its diacritic-less counterpart, "c", is not used in standard Maltese, Maltese readers are familiar with it as a letter of English. Further, like "ġ", "ħ", and "ż", "c" is often substituted for "ċ" in more casual written Maltese (Mifsud and Mitterer 2020). However, I did not use "c" in constructing nonce Maltese stimuli for use in the MaltLex visual lexical decision study, and so I do not focus on "c" here.

composition, I may find different effects of orthographic similarity in a Semitic language such as Maltese. To my knowledge, only Frost et al. (2005) have ever explored orthographic neighborhood density effects in a Semitic language, namely Hebrew: As in non-Semitic languages, they found a facilitatory effect of neighborhood density on responses to real-word targets in a masked priming visual lexical decision study. Likewise, little research has explored the processing of non-words in Maltese (cf. Twist 2006, who found effects of verb pattern for both real and non-words in lexical decision). Research on the role of diacritics in lexical processing is in its earliest stages (e.g. Marcet et al. 2020, Mifsud and Mitterer 2020). Thus, the analyses reported below explore neighborhood density effects in a language that has a novel combination of orthographic and morphological properties, while also illuminating the processing of Maltese-like non-words by Maltese readers.

2.2 Methods

In all analyses reported below, I analyze lexical decision data from the MaltLex database of Maltese visual lexical decision responses, the development and structure of which was described in greater detail in Chapter 1. The MaltLex dataset consists of approximately 237,500 total lexical decision responses to 21,994 targets that were collected from 104 native or near-native Maltese speakers who were recruited from the University of Malta. Participants ranged in age from 18−77 years (*M* = 24.0 years). Participants judged the lexicality of 200 real- and 200 non-word targets during a single lexical decision session. Individual participants could participate in up to three sessions per day, and in up to 35 total sessions across days (*M* = 5.8 sessions per participant).

Real-word targets comprised Maltese words that ranged in length from 2−21 letters (*M* = 7.1 letters), in word frequency from 0−20,446.1 occurrences per million words in the 250-million-word/253-thousand-document Korpus Malti v3.0 (*M* = 36.7 occurrences per million words), and in contextual diversity from 0−851.2 occurrences per thousand contexts in Korpus Malti v3.0 (*M*

= 11.9 occurrences per thousand contexts). For each real-word target, I calculated orthographic neighborhood density by (1) counting the number of words in Korpus Malti v3.0 that differ from the target by the substitution, insertion, or deletion of a letter (e.g. *għana* 'wealth', *għalaq* 'it shut', and *ħala* 'wasting' represent neighbors of *għala* 'why') to calculate a total number of neighbors, and then (2) weighing the target's number of neighbors by summing the word frequency/contextual diversity values of the target's neighbors, in order to compensate for non-words and various typos that may be included in Korpus Malti v3.0 and in order to account for effects of neighbor frequency on lexical processing performance (cf. Andrews 1997 for a review of neighbor frequency effects). Real-word targets ranged in word frequency-weighted neighborhood density from 0−49,584.6 occurrences per million words in Korpus Malti v3.0 (M = 157.8 occurrences per million words), and in contextual diversity-weighted neighborhood density from 0−3,071.4 occurrences per thousand contexts in Korpus Malti v3.0 (M = 41.9 occurrences per thousand contexts).

For each real-word target, I created a corresponding non-word target by replacing at least one consonant letter to generate a nonce form that was matched with the original target in word frequency-weighted neighborhood density, which was then vetted by a native Maltese speaker. Non-word targets ranged in length from 2−21 letters (M = 7.1 letters), in word frequency-weighted neighborhood density from 0−54,575.6 occurrences per million words in Korpus Malti v3.0 (M = 124.3 occurrences per million words), and in contextual diversity-weighted neighborhood density from 0−3,895.0 occurrences per thousand contexts (M = 28.8 occurrences per thousand contexts)

As in the analyses reported in Chapter 1, I calculated the mean and standard deviation (SD) of the RTs for each participant and removed 6,723 datapoints for which the RT was ±2.5 SDs from each participant's mean RT prior to data analysis. This reduced the dataset to 228,358 lexical decision responses to 21,994 unique targets (i.e. 96.2% of the dataset), comprising 115,612 lexical decision responses to real-word targets and 112,746 responses to non-word targets.

2.3 Virtual experiment 4: Analysis of orthographic neighborhood density x target lexicality

Prior research has consistently shown that lexical decision performance for real versus non-word targets is differently affected by changes in the target's orthographic neighborhood density: Generally, as neighborhood density increases, RTs decrease and response accuracy increases for real-word targets. For non-words, the opposite pattern holds: As neighborhood density increases, RTs increase and response accuracy decreases, owing to increased competition with potential real-word candidates which impedes readers' ability to confirm the target's status as a non-word (e.g. Andrews 1989, Balota et al. 2004, Carreiras et al. 1997, Coltheart et al. 1977, Forster and Shen 1996, Hendrix and Sun 2021, Keuleers, Diependaele, and Brysbaert 2010, Yap et al. 2015). In Chapter 1, I showed that the expected neighborhood density effect holds for real-word targets in the MaltLex dataset: Participants responded faster and more accurately to real-word targets as they increased in neighborhood density. I did not explore neighborhood density effects for non-word targets, nor has anyone explored neighborhood density effects for non-word targets in a Semitic language (cf. Frost et al. 2005 found a facilitatory neighborhood density effect for real-word targets in a Hebrew visual masked priming lexical decision task, but they did not explore neighborhood density effects for non-word targets; Twist 2006 analyzed performance for non-word targets in a Maltese visual lexical decision experiment, but did not analyze neighborhood density effects).

I address this gap in the following analyses, assessing whether neighborhood density effects differ for real versus non-word targets in Maltese. In particular, I compare lexical decision performance to the MaltLex real-word versus non-word targets, testing the interaction between the target's orthographic neighborhood density value (weighted according to the contextual diversity values of its neighbors) and lexicality. I expect to find that participants respond faster and more accurately as real-word targets increase in neighborhood density (cf. the effects that I reported in Chapter 1), but slower and less accurately as non-word targets increase in neighborhood density.

2.3.1 RT analysis and results

I analyzed RTs on trials on which participants provided the correct response (210,960 lexical decision responses to 21,900 unique targets, including 104,644 responses to 10,951 real-word targets and 106,316 responses to 10,949 non-word targets) by fitting a maximum likelihood-fitted LMER model using the bobyqa optimizer in R (R Core Team 2021), using the lme4 package (Bates, Maechler et al. 2015). I assessed the significance of fixed effects by using the lmerTest package (Kuznetsova et al. 2017) to simulate Satterthwaite approximations for degrees of freedom.

The model included log-transformed RT as its dependent variable; target lexicality (levels: real versus nonce, reference: real), log-transformed contextual diversity-weighted orthographic neighborhood density (to account for targets having a value of 0, I added 1 to each value prior to the log transformation), and the interaction of target lexicality by log-transformed contextual diversity-weighted orthographic neighborhood density as fixed effects; and participant and target as random effects. The model also included the following control variables as fixed effects:

- log-transformed target length (in number of letters);
- participant's age;
- participant's trial number;
- participant's overall session number;
- participant's session number within a given day.

I included random slopes for both target lexicality by-participants and log-transformed contextual diversity-weighted orthographic neighborhood density by-participants: The results of a series of likelihood ratio tests indicated that including random slopes for target lexicality by-participants improved model fit compared to the random intercepts model ($\chi^2(2) = 2{,}926.2$, $p < 0.001$), as did including random slopes for neighborhood density by-participants compared to the random intercepts model ($\chi^2(2) = 1{,}210.1$, $p < 0.001$). Further, including random slopes for both variables improved model fit compared to the model that only included random slopes for target lexicality by-participants ($\chi^2(3) = 646.0$, $p < 0.001$) and compared to the model that only included random slopes for neighborhood density by-participants ($\chi^2(3) = 2{,}362.0$, $p < 0.001$), justifying the inclusion of both random slopes for this dataset (Bates, Kliegl et al. 2015). A model that included random slopes for target lexicality, neighborhood density, and the interaction by-participants failed to converge, and so I do not consider models with more complex random effects structures.

All fixed effects were significant and patterned in expected directions: Participants were slower to respond as target length increased ($\hat{\beta} = 0.22790$; $t(20{,}350) = 62.51$, $p < 0.001$). RTs increased both with participant's age ($\hat{\beta} = 0.00823$; $t(416.9) = 7.40$, $p < 0.001$) and across trials, likely reflecting an effect of fatigue ($\hat{\beta} = 0.00002$; $t(201{,}200) = 5.14$, $p < 0.001$). On the other hand, RTs decreased both across sessions generally ($\hat{\beta} = -0.00239$; $t(185{,}100) = -22.86$, $p < 0.001$) and across sessions within the same day ($\hat{\beta} = -0.01839$; $t(202{,}000) = -14.42$, $p < 0.001$).

Consistent with prior research, the effect of target lexicality was significant ($\hat{\beta} = 0.04762$; $t(181.7) = 4.87$, $p < 0.001$), with participants responding slower to non-word targets ($M = 990$ ms) than real-word targets ($M = 850$ ms). The effect of neighborhood density was also significant ($\hat{\beta} = -0.01800$; $t(261.1) = -19.97$, $p < 0.001$), with participants responding faster to real-word targets as neighborhood density increased. Moreover, the interaction of target lexicality and neighborhood density was significant ($\hat{\beta} = 0.01983$; $t(20{,}760) = 28.29$, $p < 0.001$), with the facilitatory effect of neighborhood density diminishing for non-word targets compared to real-word targets (Figure 2.1).

To investigate this interaction further, I split the dataset by target lexicality and re-fitted separate LMER models to the real-word and non-word data. Each model included log-transformed RT as its dependent variable. As fixed effects, each model included log-transformed contextual diversity-weighted orthographic neighborhood density, log-transformed target length, participant's

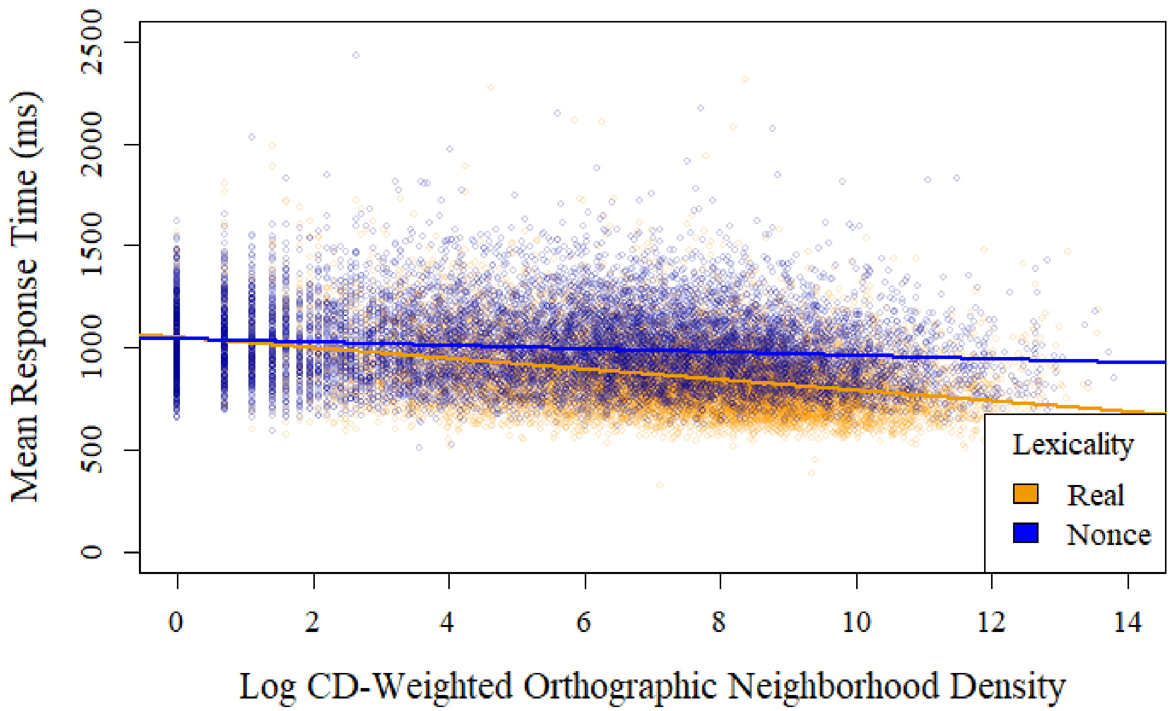

Figure 2.1. Mean RTs by log CD-weighted orthographic neighborhood density for real versus non-word targets. Note that log RT was analyzed, but raw RT is presented here.

age, participant's trial number, participant's overall session number, and participant's session number within a given day. Each model included participant and target as random effects, as well as random slopes for neighborhood density by-participants: The results of likelihood ratio tests comparing these models with the random intercepts models suggest that they are justified for both the real-word ($\chi^2(2) = 175.8$, $p < 0.001$) and non-word datasets ($\chi^2(2) = 561.0$, $p < 0.001$).

All fixed effects were again significant for both the real-word and non-word data: For real-word targets, the effects of target length ($\hat{\beta} = 0.13820$; $t(10{,}110) = 25.76$, $p < 0.001$), age ($\hat{\beta} = 0.00474$; $t(275.6) = 3.61$, $p < 0.001$), trial number ($\hat{\beta} = 0.00008$; $t(99{,}200) = 11.01$, $p < 0.001$), session number ($\hat{\beta} = -0.00136$; $t(94{,}350) = -8.99$, $p < 0.001$), and same-day session number were significant ($\hat{\beta} = -0.01447$; $t(99{,}740) = -7.85$, $p < 0.001$). For non-word targets, the effects of target length ($\hat{\beta} = 0.32630$; $t(10{,}360) = 68.66$, $p < 0.001$), age ($\hat{\beta} = 0.01035$; $t(272) = 7.46$, $p < 0.001$), trial number ($\hat{\beta} = -0.00003$; $t(102{,}500) = -3.81$, $p < 0.001$), session number ($\hat{\beta} = -0.00346$;

$t(88{,}780) = -24.38$, $p < 0.001$), and same-day session number were significant ($\hat{\beta} = -0.02205$; $t(102{,}700) = -12.56$, $p < 0.001$). Consistent with previous research, moreover, the neighborhood density effect was facilitatory for real-word targets ($\hat{\beta} = -0.02196$; $t(196.1) = -21.83$, $p < 0.001$) but inhibitory for non-word targets ($\hat{\beta} = 0.00678$; $t(147.9) = 7.63$, $p < 0.001$): As neighborhood density increased, participants were faster to respond for real words but slower for non-words.

2.3.2 Accuracy analysis and results

I analyzed response accuracy on all trials (228,358 responses to 21,994 unique targets, including 115,612 responses to 11,040 real-word targets and 112,746 responses to 10,954 non-word targets) by fitting a GLMER model using the binomial logit link function and the bobyqa optimizer in R (R Core Team 2021), using the lme4 package (Bates, Maechler et al. 2015). The model included response accuracy as the dependent variable (0 = incorrect, 1 = correct). As fixed effects, the model included target lexicality, log-transformed contextual diversity-weighted target orthographic neighborhood density, and the target lexicality by orthographic neighborhood density interaction, along with log-transformed target length and the participant's within-day session number as control predictors. The model included participant and target as random effects. Models that included random slopes for target lexicality by-participants or for neighborhood density by-participants failed to converge, and so I analyze the results of the random intercepts model.

All fixed effects were significant: Participants responded more accurately as target length increased ($\hat{\beta} = 0.98502$; $z = 12.70$, $p < 0.001$) but less accurately across sessions within the same day, likely reflecting an effect of fatigue ($\hat{\beta} = -0.11335$; $z = -5.64$, $p < 0.001$). The effect of target lexicality was significant ($\hat{\beta} = 2.61071$; $z = 25.21$, $p < 0.001$), with participants responding more accurately to non-words (94.3%) than to real words (90.5%). Perhaps this stems from the fact that many obscure or archaic, and thus low-frequency words were used as targets to expand the real-word target set. The effect of neighborhood density was also significant ($\hat{\beta}$ = 0.24535; z = 20.60, p < 0.001), with participants responding more accurately to real-word targets as neighborhood density increased. Moreover, the interaction of target lexicality and neighborhood density was also significant ($\hat{\beta}$ = –0.29887; z = –20.48, p < 0.001), with the facilitatory effect of increases in neighborhood density diminishing for non-word targets compared to real-word targets (Figure 2.2).

To investigate this interaction further, I split the dataset by target lexicality and re-fitted separate GLMER models to the real-word and non-word data. Each model included response accuracy as its dependent variable. Each model included log-transformed contextual diversity-weighted orthographic neighborhood density, log-transformed target length, and the participant's within-day session number as fixed effects. Each model also included subject and target as random effects. For the real-word data, including random slopes for neighborhood density by-participants

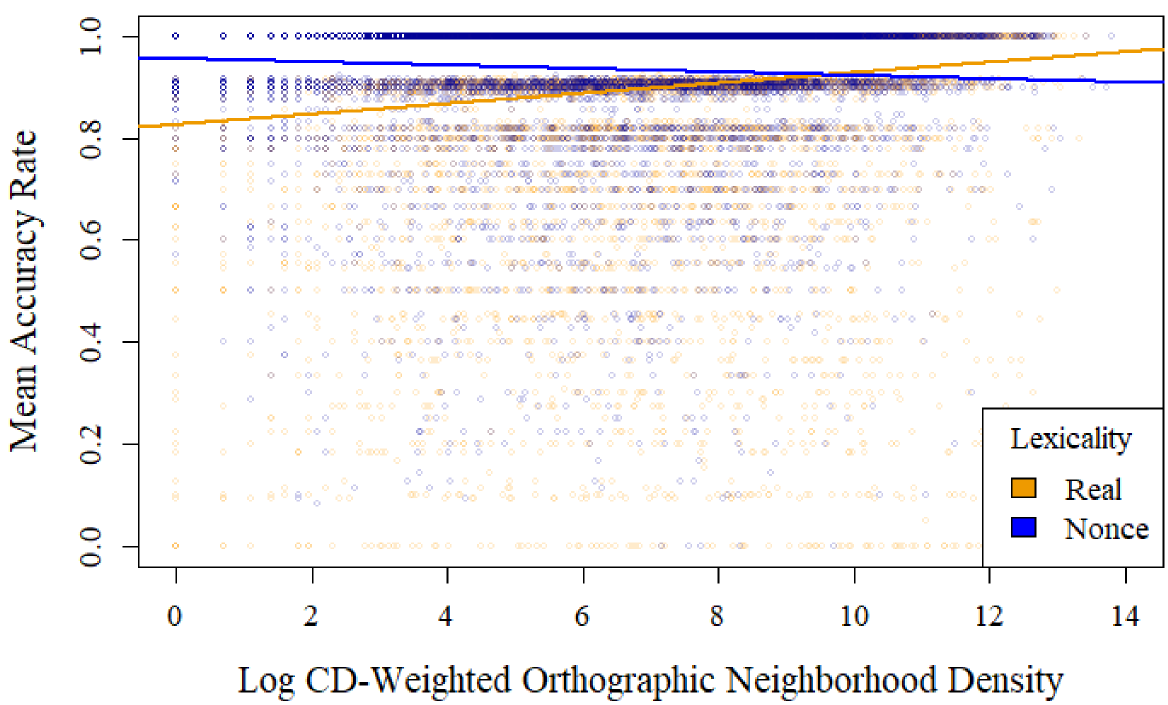

Figure 2.2. Mean accuracy by log CD-weighted orthographic neighborhood density for real versus non-word targets.

improved model fit relative to the random intercepts model ($\chi^2(2) = 69.8$, $p < 0.001$), whereas for the non-word data the model failed to converge when random slopes for neighborhood density by-participants were included. For ease of comparison of the two analyses, I did not include random slopes in either model. (However, all fixed effects remained significant and patterned in the same directions in the analysis of the real-word data when such random slopes were included).

For real-word targets, the effects of target length ($\hat{\beta} = 3.05355$; $z = 26.33$, $p < 0.001$) and same-day session number were significant ($\hat{\beta} = –0.13792$; $z = –4.99$, $p < 0.001$), with participants responding more accurately as target length increased but less accurately across sessions within the same day. For non-word targets, the effect of target length was significant ($\hat{\beta} = –1.32662$; $z = –13.11$, $p < 0.001$): Participants responded less accurately as target length increased. The effect of same-day session number was not significant for non-word targets ($\hat{\beta} = –0.05743$; $z = –1.87$, *n.s.*). Consistent with previous research, moreover, the effect of orthographic neighborhood density was significant and facilitatory for real-word targets ($\hat{\beta} = 0.38607$; $z = 27.43$, $p < 0.001$) but inhibitory for non-word targets ($\hat{\beta} = –0.17535$; $z = –17.04$, $p < 0.001$): As neighborhood density increased, participants responded more accurately to real-word targets but less accurately to non-words.

2.3.3 Discussion

The results of both analyses are consistent with my predictions and with the findings of prior studies that have compared orthographic neighborhood density effects for real- versus non-word targets (e.g. Andrews 1989, Coltheart et al. 1977, Hendrix and Sun 2021, Yap et al. 2015): MaltLex participants responded faster and more accurately to real-word targets as they increased in orthographic neighborhood density, but they responded slower and less accurately to non-word targets as they increased in orthographic neighborhood density and became more similar to real

Maltese words. To my knowledge, this is the first demonstration of an orthographic neighborhood density effect on visual lexical decision performance for non-words in any Semitic language.

In analyzing responses to real- and non-word targets, I excluded any measure of word frequency, which for all non-words should effectively be null, as a predictor of lexical decision performance in order to facilitate the comparison of performance for real and non-word targets. While the effect of orthographic neighborhood density is generally facilitatory for real-word targets, prior research has shown that the advantage that neighborhood density affords for visual lexical decision diminishes as real-word targets become more frequent (e.g. Andrews 1989, 1992; Lim 2016; Sears et al. 1995). In the next section, I attempt to replicate these findings by assessing how contextual diversity (i.e. a measure of frequency) interacts with orthographic neighborhood density in predicting lexical decision performance for real-word targets in the MaltLex dataset.

2.4 Virtual experiment 5: Analysis of neighborhood density x contextual diversity (real words)

Consistent with prior research, I observed significant facilitatory effects of orthographic neighborhood density on lexical decision performance to Maltese words in the MaltLex dataset: Participants responded faster and more accurately to real-word targets as neighborhood density increased and the targets became more orthographically similar to other Maltese words. However, prior research has also found that the neighborhood density effect for real words exhibits a complex relationship with word frequency, in that the advantage for high-density words compared to low-density words diminishes and may even disappear as targets increase in frequency (Andrews 1989, 1992; Lim 2016; Sears et al. 1995). That is, lexical decision participants respond faster and more accurately to low-frequency, high-density real-word targets than to low-frequency, low-density targets, whereas no such difference may be observed for responses to high-frequency targets.

I address these findings in the following analyses, assessing whether the asymmetrical effect of orthographic neighborhood density for low-frequency versus high-frequency real words replicates in Maltese, a language with a unique combination of orthographic and morphological properties compared to the experimental language in previous studies: English. To my knowledge, these will be the first analyses to ever explore the relationship between neighborhood density and word frequency in any Semitic language. In particular, I analyze lexical decision responses to the MaltLex real-word targets, testing the interaction between the target's orthographic neighborhood density value (again, weighted according to the contextual diversity values of its neighbors) and its contextual diversity value. I expect participants to respond faster and more accurately as the targets increase in contextual diversity (cf. the effects reported in Chapter 1), and as the targets increase in orthographic neighborhood density (cf. the effects reported in Section 3.1), but for the advantage for high-density targets to also diminish as targets increase in contextual diversity.

2.4.1 RT analysis and results

To assess how word frequency interacts with the neighborhood density effect on RTs, I analyzed RTs to real-word targets on trials on which participants provided the correct response (104,644 lexical decision responses to 10,951 unique targets) by fitting a maximum likelihood-fitted LMER model using the bobyqa optimizer in R (R Core Team 2021), using the lme4 package (Bates, Maechler et al. 2015). I assessed the significance of fixed effects by using the lmerTest package (Kuznetsova et al. 2017) to simulate Satterthwaite approximations for degrees of freedom.

The model included log-transformed RT as its dependent variable. As fixed effects, the model included the target's log-transformed contextual diversity, its log-transformed contextual diversity-weighted orthographic neighborhood density, and the interaction of contextual diversity by neighborhood density. The model included the following control predictors as additional fixed effects: log-transformed target length, participant's age, participant's trial number, participant's overall session number, and participant's session number within a given day. The model included participant and target as random effects, and random slopes for contextual diversity by-participants and neighborhood density by-participants: The results of a series of likelihood ratio tests revealed that adding random slopes for contextual diversity by-participants ($\chi^2(2) = 504.4$, $p < 0.001$) or for neighborhood density by-participants alone improved model fit relative to the random intercepts model ($\chi^2(2) = 192.9$, $p < 0.001$). Further, including random slopes for both contextual diversity by-participants and neighborhood density by-participants improved model fit relative to the model that included only random slopes for contextual diversity by-participants ($\chi^2(3) = 53.8$, $p < 0.001$) and the model that included only random slopes for neighborhood density by-participants ($\chi^2(3) = 365.3$, $p < 0.001$), justifying their inclusion for this dataset (Bates, Kliegl et al. 2015). An additional model having random slopes for contextual diversity by-participants, neighborhood density by-participants, and the interaction of contextual diversity and neighborhood density by-participants failed to converge, and so I do not analyze models with more complex random effects structures.

All fixed effects were significant and patterned in the expected directions: Participants were slower to respond as target length increased ($\hat{\beta} = 0.16670$; $t(10{,}090) = 35.14$, $p < 0.001$). RTs increased both as participants' age increased ($\hat{\beta} = 0.00438$; $t(268.9) = 3.38$, $p < 0.001$) and across the trials of a given session ($\hat{\beta} = 0.00008$; $t(100{,}500) = 11.32$, $p < 0.001$). On the other hand, RTs decreased both across sessions generally ($\hat{\beta} = -0.00174$; $t(88{,}580) = -11.78$, $p < 0.001$) and across sessions that were completed within the same day ($\hat{\beta} = -0.01464$; $t(101{,}000) = -8.00$, $p < 0.001$).

Consistent with prior research and with analyses performed here, the effect of contextual diversity was significant ($\hat{\beta} = -0.05247$; $t(509.0) = -26.51$, $p < 0.001$), with participants responding faster as CD increased. Further, the effect of neighborhood density was significant ($\hat{\beta} = -0.01444$;

t(1,083) = –9.41, *p* < 0.001), with participants responding faster as neighborhood density increased. However, the interaction of contextual diversity and neighborhood density was also significant ($\hat{\beta}$ = 0.00197; *t*(11,240) = 9.41, *p* < 0.001), with the facilitatory effect of neighborhood density diminishing as targets increased in their contextual diversity: That is, high-frequency, high-density targets yielded a smaller RT advantage compared to high-frequency, low-density targets than did low-frequency, high-density targets compared to low-frequency, low-density targets (Figure 2.3).

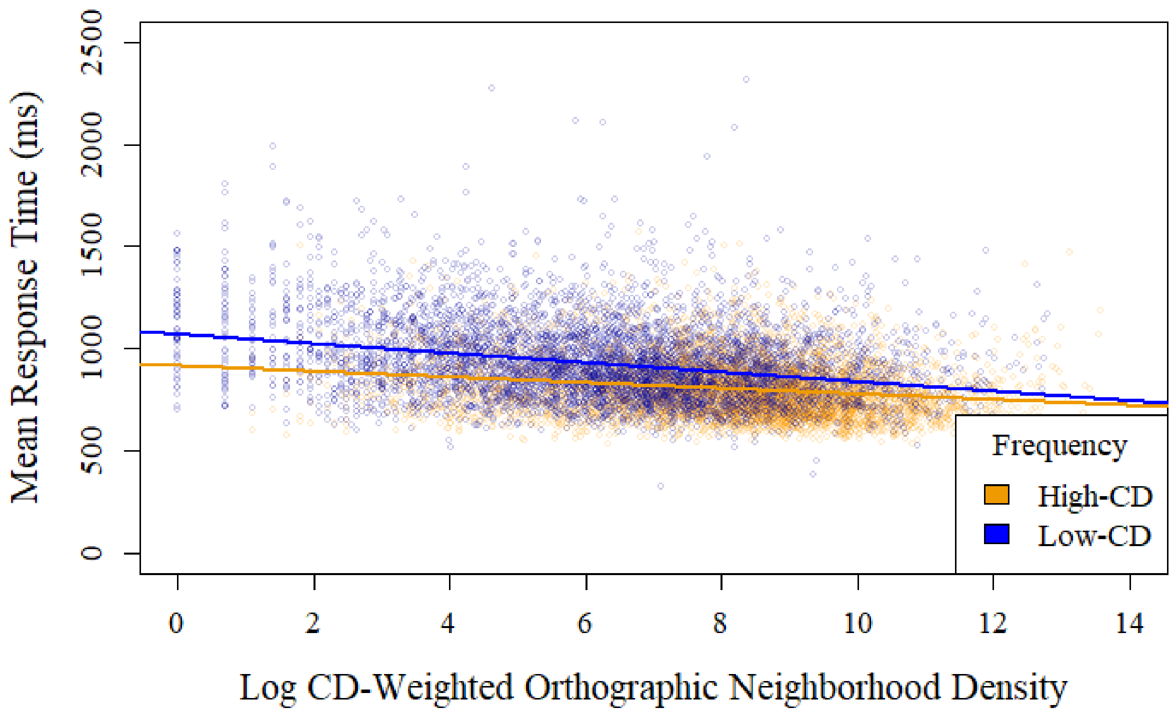

Figure 2.3. Mean RTs by log CD-weighted orthographic neighborhood density for high-CD and low-CD real-word targets. To visualize the interaction of contextual diversity (CD) by neighborhood density, I binned targets according to the median CD value of 595 (high-CD > 595, low-CD ≤ 595). Note that log RT was analyzed, but raw RT is presented here.

2.4.2 Accuracy analysis and results

To assess how word frequency interacts with the neighborhood density effect on response accuracy, I analyzed response accuracy to real-word targets on all trials (115,612 lexical decision responses to 11,040 unique real-word targets) by fitting a GLMER model using the binomial logit link function and the bobyqa optimizer in R (R Core Team 2021), using the lme4 package (Bates, Maechler et al. 2015). The model included response accuracy as the dependent variable (0 = incorrect, 1 = correct). Unfortunately, but not unexpectedly for a GLMER analysis, a model with as "simple" of a fixed effects structure as one that included only the target's log-transformed contextual diversity, its log-transformed contextual diversity-weighted orthographic neighborhood density, and the interaction of contextual diversity by neighborhood density failed to converge.

To compensate for this, I binned the targets by their contextual diversity value, treating as "low-frequency" targets having a contextual diversity value less than or equal to the median value of 595 (i.e. 2.4 occurrences per thousand contexts; this included 56,884 responses to 5,524 real-word targets), and as "high- frequency" targets with a contextual diversity value greater than 595 (58,728 responses to 5,516 targets). As fixed effects, the model included the target's contextual diversity category (levels: low-frequency versus high-frequency; reference: low-frequency), its log-transformed contextual diversity-weighted orthographic neighborhood density, and the contextual diversity category by orthographic neighborhood density interaction, along with log-transformed target length and the participant's within-day session number as control predictors.

The model also included participant and target as random effects, as well as random slopes for contextual diversity category by-participants and for neighborhood density by-participants: The results of a series of likelihood ratio tests revealed that adding random slopes for contextual diversity category by-participants ($\chi^2(2) = 66.9$, $p < 0.001$) or neighborhood density by-participants alone improved model fit relative to the random intercepts model ($\chi^2(2) = 56.2$, $p < 0.001$). Further, including random slopes for both contextual diversity category by-participants and neighborhood density by-participants improved model fit relative to the model that included only random slopes for contextual diversity category by-participants ($\chi^2(3) = 35.1$, $p < 0.001$) and the model that included only random slopes for neighborhood density by-participants ($\chi^2(3) = 45.7$, $p < 0.001$), justifying their inclusion here (Bates, Kliegl et al. 2015). Adding random slopes for the interaction of contextual diversity category and neighborhood density by-participants caused the model to fail to converge, and so I do not analyze models with more complex random effects structures.

Target length and same-day session number had significant effects on response accuracy: Participants responded more accurately as target length increased ($\hat{\beta} = 2.53197$; $z = 22.95$, $p < 0.001$) but less accurately across sessions held on the same day ($\hat{\beta} = –0.11765$; $z = –4.25$, $p < 0.001$). Moreover, the effect of contextual diversity category was significant ($\hat{\beta} = 2.26459$; $z = 10.65$, $p < 0.001$), with participants responding more accurately to high-frequency targets (96.6%) than low-frequency targets (84.3%). The effect of neighborhood density was also significant ($\hat{\beta} = 0.17289$; $z = 10.04$, $p < 0.001$), with participants responding more accurately to low-frequency targets as orthographic neighborhood density increased. In contrast, the interaction of contextual diversity category and neighborhood density was not significant ($\hat{\beta} = –0.03665$; $z = –1.42$, *n.s.*), thus providing insufficient evidence to suggest that the effect of neighborhood density on response accuracy reliably differed for high-frequency targets compared to low-frequency ones (Figure 2.4).

2.4.3 Discussion

The results of the RT analysis were consistent with my predictions and with the findings of earlier research on orthographic neighborhood density effects (Andrews 1989, 1992; Lim 2016; Sears et al. 1995): MaltLex participants responded faster to real-word targets as they increased in

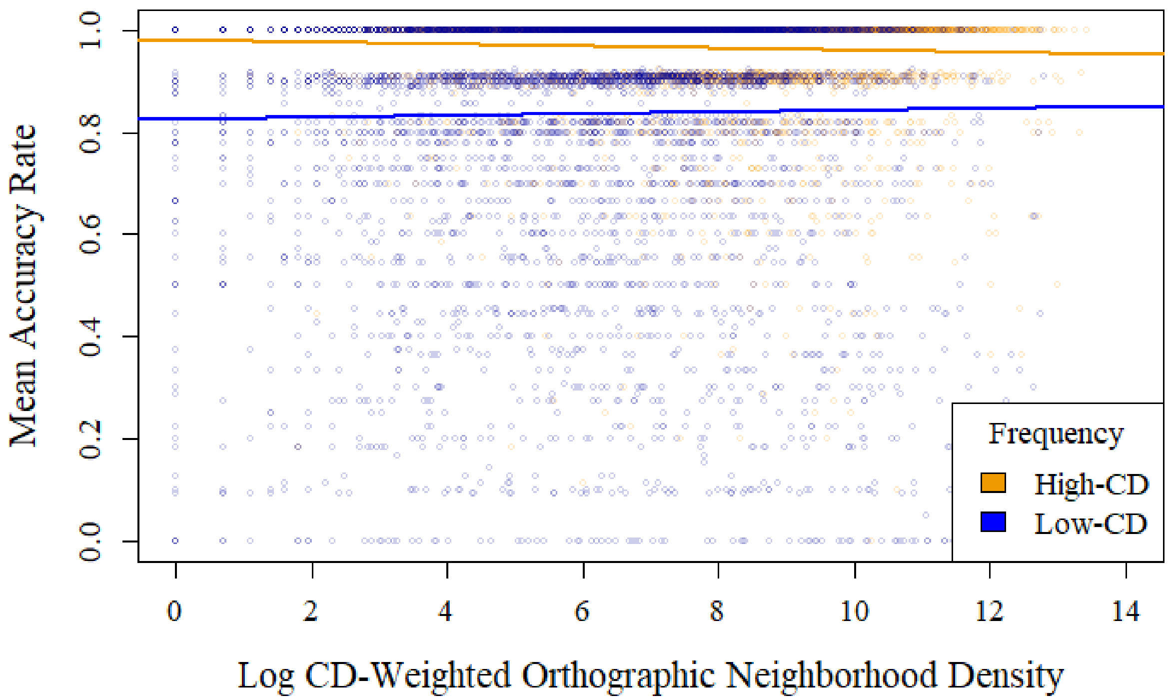

Figure 2.4. Mean accuracy by log CD-weighted orthographic neighborhood density for high-CD and low-CD real-word targets. To visualize the interaction of CD by neighborhood density, I binned targets according to the median CD of 595 (high-CD > 595, low-CD ≤ 595).

orthographic neighborhood density, but this advantage for high-density targets compared to low-density targets was also greater for low-frequency targets than for high-frequency targets.

On the other hand, the results of the response accuracy analysis were only partly consistent with my predictions: MaltLex participants responded more accurately to real-word targets as they increased in orthographic neighborhood density, but the orthographic neighborhood density by contextual diversity interaction was not significant, and so there is insufficient evidence to assess whether the neighborhood density effect on response accuracy reliably differed for low-frequency versus high-frequency targets. This may seem surprising, in that the size of the MaltLex dataset should afford the analysis sufficient power to detect such an effect (if it exists), given that earlier studies detected such effects in smaller datasets obtained through traditional factorial experiments.

However, the need to bin the targets by contextual diversity to mitigate convergence issues, which are an unfortunate but common issue in conducting GLMER analyses, may have limited the power of the present analysis to detect subtle effects that involve contextual diversity such as this.

2.5 Virtual experiment 6: Analysis of diacritic-based word-similarity (non-words)

Three pairs of letters of the Maltese alphabet differ only in whether they include diacritics, including: “g” and “ġ”, “h” and “ħ”, and “z” and “ż”. In practice, however, Maltese writers often omit diacritics from “ġ”, “ħ”, and “ż” (as well as “ċ”, which lacks a diacritic-less counterpart “c” in standard Maltese), in part due to the prevalence of English keyboards which lack the diacritic-bearing letters in Malta and to the absence of Maltese spell checkers (Mifsud and Mitterer 2020). Thus, Maltese readers often encounter Maltese words that are written without diacritics. During the post-session debriefings, multiple participants mentioned having noticed that some of the non-word targets resembled real words except in that a diacritic either had been omitted from some letters or had been added superfluously to other letters, and they believed that such similarities had made rejecting these non-words in the lexical decision task more difficult. Indeed, a small number of the non-word targets were (unintentionally) constructed such that they differed from an existing word only in whether some diacritic had been included on 1–2 letters (M = 1.04 letters). For example, the non-word target *gera* resembles the real word *ġera* ‘he ran, roamed, flowed’ except in the absence of a dot above the letter “g”, while the non-word target *vaġun* resembles the real word *vagun* ‘wagon, carriage (of a train)’ except in the inclusion of a dot above the letter “ġ”.

To identify which non-word targets resembled Maltese words in this manner, I searched the Maltese lexical database Ġabra (Camilleri 2013) for real words that could be derived from each of the 10,954 non-word targets by the addition of a diacritic to the letters “g”, “h”, or “z”, or by the omission of a diacritic from the letters “ġ”, “ħ”, or “ż”. My search revealed 51 such potential real words that are listed in Ġabra (i.e. corresponding to 0.46% of the non-word targets), which are presented in Appendix 2. However, it is worth noting that Ġabra lists many infrequent and/or archaic words that may be unknown to participants, as well as many morphologically possible yet semantically implausible verb conjugations: For example, the non-word *raħbuhom* resembles *rahbuhom* 'they joined them a religious order', which is listed in Ġabra as a form of the intransitive verb *rahab* 'to join a religious order' that is inflected for perfect aspect and a third person plural subject (cf. *rahbu* 'they joined a religious order'), plus a third person plural direct object (-*hom*).

It is unlikely that participants would have perceived any of these non-word targets as their infrequent, archaic, or implausible "real"-word counterparts. To account for such discrepancies between Ġabra and participants' language experience, I also obtained each potential real word's word frequency and contextual diversity values from Korpus Malti v3.0 (Gatt and Čéplö 2013), the 250-million-token/253-thousand-document corpus of written Maltese: 23 of the potential real words do not occur in Korpus Malti, while the remaining 28 potential real words (i.e. 0.26% of the total non-word targets) range in word frequency from 0.004–8.9 occurrences per million words (*M* = 0.9 occurrences per million words) and in contextual diversity from 0.004–6.2 occurrences per thousand contexts (*M* = 0.7 occurrences per thousand contexts). That participants noticed these non-words and perceived them to be more difficult, despite their comprising a small percentage of the non-word target set and despite their real-word counterparts' being relatively infrequent overall (cf. actual real-word targets had an average word frequency of 36.7 occurrences per million words, and an average contextual diversity of 11.9 occurrences per thousand contexts), underscores the saliency of the apparent processing difficulty that these non-words presented for participants.

I address two questions regarding the processing of Maltese non-words and of the role that diacritics play in Maltese lexical processing in the following analyses: (1) Do Maltese readers subconsciously tolerate changes to letter diacritics, such that they process non-words that differ from real words except in their diacritics as their real-word counterparts? That is, were participants correct in suggesting that such non-words were more difficult to reject in the lexical decision task? I address this question by analyzing lexical decision responses to the MaltLex non-word targets, assessing whether non-words that resemble real words except in their diacritics are more difficult to reject than other non-word targets: I expect participants to respond slower and less accurately to such non-word targets. In particular, I analyze whether the potential real-word counterpart's contextual diversity value impacts lexical processing performance for non-word targets that differ from a potential real word except in their diacritics, expecting that it will be more difficult for participants to reject such non-words as the potential real-word counterpart increases in frequency.

(2) Is the effect of diacritic-based word-similarity separate from the effect of orthographic neighborhood density on non-word processing? As I have shown in Section 2.3, Maltese readers respond slower and less accurately to non-word targets as neighborhood density increases and non-word targets become more similar orthographically to real words, stemming from the activation of potential real-word candidates (cf. Andrews 1989, Hendrix and Sun 2021). I ask whether the similarity of real and non-words that differ only in their diacritics influences lexical processing beyond the more general effect of orthographic similarity that orthographic neighborhood density already captures, and I assess this by including a measure of neighborhood density as a predictor in the following analyses: Any effect of diacritic-based word-similarity that is significant in these analyses is thus independent from the neighborhood density effect (Wurm and Fisicaro 2014).

Current models of visual word recognition typically assume that letter representations are abstract (Kinoshita et al. 2013), and only recently have researchers begun to explore how the visual components of letters mediate visual lexical processing. For example, in visual masked priming lexical decision, Marcet and Perea (2017, 2018) have observed significant facilitatory priming by non-word primes that differ from the target by the replacement of a letter with a visually similar letter/set of letters versus a visually dissimilar letter/set of letters in Spanish (e.g. *nevtral* versus *neztral* priming *NEUTRAL* ‘neutral’, *documento* versus *docusnento* priming *DOCUMENTO* ‘document’). These results suggest that a letter’s visual components play a critical role in language processing, such that visually similar letters may in fact activate the same letter representations.

The role of diacritics in lexical processing appears to be language- and letter-specific. For example, Perea et al. (2016, 2018) obtained no priming in an Arabic visual masked priming lexical decision task by primes that differed from the target by the replacement of a consonant letter with a letter that differs only in a diacritic versus primes that differed from the target by the replacement of one consonant with a dissimilar letter (e.g. صخفية versus صكفية priming صحفية ‘journalist’). They interpret their results as suggesting that diacritics are processed rapidly, such that consonant letters that differ only in their diacritics in Arabic (e.g. “خ” and “ح”) nonetheless do not activate the same lexical representations. Similarly, priming is not obtained by primes that differ from the target by only a diacritic in Spanish (e.g. *rasgó* ‘he tore’ priming *RASGO* ‘I tear; feature’; Domínguez and Cuetos 2018) and French (e.g. *tâper* priming *TAPER* ‘to type’; Chetail and Boursain 2019)

In contrast, Marcet et al. (2020) obtained priming in Spanish by nonce primes that differ from the target by the omission of a tilde above the letter “ñ” compared to primes that differ by the substitution of a visually dissimilar letter (e.g. *muneca* versus *museca* priming *MUÑECA* ‘doll’). However, they obtained no priming by nonce primes that differ by the addition of a tilde above the letter “n” (e.g. *moñeda* versus *moseda* priming *MONEDA* ‘coin’). This suggests that, in the early stages of visual lexical processing, the absence of a tilde permits “n” to activate the letter representations for “n” and “ñ”, whereas the presence of a tilde on “ñ” is rapidly processed thus preventing it from activating any letter representation for its diacritic-less counterpart “n”.

Mifsud and Mitterer (2020) recently explored the role of diacritics in Maltese visual lexical processing, focusing on how their absence from the letters "ċ", "ġ", "ħ", and "ż" impacts word recognition in a visual masked priming lexical decision study and a sentence reading task.[16] In the lexical decision study, they obtained a facilitatory priming effect both by repetition primes and by primes that were derived from the target by the omission of a diacritic compared to primes that were derived by substituting a dissimilar letter (e.g. *girja* versus *dirja* priming *ĠIRJA* 'run'). In the sentence reading task, they varied whether individual words were written using their standard diacritics or not, and they found no effect of diacritic omission on participants' eye movements. The researchers interpret these results as indicating that Maltese readers process words that are missing diacritics the same as their diacritic-bearing counterparts. Further, they propose that their results differ from those of prior studies (e.g. Perea et al. 2016, 2018) because Maltese readers are uniquely exposed to forms that are missing their diacritics in everyday language use, suggesting that language experience modulates the processing of diacritics in visual word recognition.

In the present study, I analyze non-words that are related to a real Maltese word both by the omission of a diacritic and by the addition of a superfluous diacritic for two reasons: (1) Few such non-words were included as targets in the MaltLex non-word target set, and (2) participants explicitly identified both types of non-words as being difficult to reject. It is worth acknowledging, of course, that Mifsud and Mitterer's (2020) hypothesis predicts that the addition of superfluous diacritics should not make a non-word more "word-like" to Maltese readers, since Maltese readers do not encounter words with superfluous diacritics in real language use. If this were the case, then analyzing the two types of non-words together as such makes this a more conservative analysis in that the latter type of non-words will offset any processing disadvantage imposed by the former.

[16] I wish to thank one of the WOW2021 participants for bringing this relevant research to my attention.

2.5.1 RT analysis and results

To assess how lexical processing performance for non-words is affected by the absence of diacritics that would make a non-word target into a real word (e.g. the non-word *gera* resembles the real word *ġera* 'he ran' except in the absence of a dot above the letter "g"), and the inclusion of superfluous diacritics whose absence would make a non-word target into a real word (e.g. the non-word *vaġun* resembles the real word *vagun* 'wagon' except in the inclusion of a dot above the letter "ġ"), I analyzed RTs to non-word targets on trials on which participants provided the correct response (106,316 responses to 10,949 unique targets) by fitting a maximum likelihood-fitted LMER model using the bobyqa optimizer in R (R Core Team 2021), using the lme4 package (Bates, Maechler et al. 2015). I assessed the significance of fixed effects by using the lmerTest package (Kuznetsova et al. 2017) to simulate Satterthwaite approximations for degrees of freedom.

The model included log-transformed RT as its dependent variable. To assess how diacritic-based similarity to a real word affects non-word lexical decision performance, and to account for the fact that potential words that are listed in Ġabra differ in the frequency with which participants will have encountered them and may even be unknown to participants, I included the potential real-word target's log-transformed contextual diversity value as a fixed effect: This included non-zero values for 28 of the non-word targets (raw range: 0.004–6.2 occurrences per thousand contexts).[17] To assess whether diacritic-based word-similarity had an effect on RTs that was independent from the overall orthographic neighborhood density effect, the model also included the target's log-transformed contextual diversity-weighted orthographic neighborhood density as

[17] I also fitted a model in which the non-word targets were instead coded according to whether Ġabra simply listed a real word that differed from it in the addition/omission of a diacritic, without taking into account the potential real-word target's contextual diversity: In this analysis 51 targets were coded as having a potential real-word counterpart. Comparison of the AIC values for random intercepts models in which frequency was versus was not accounted for showed that the model in which frequency was accounted for provided a better fit to the dataset ($\Delta = 8.57$, $p < 0.05$).

a fixed effect. Finally, as control predictors, the model also included the following variables as additional fixed effects: log-transformed target length, participant's age, participant's trial number, participant's overall session number, and participant's session number within a given day.

The model included participant and target as random effects, as well as random slopes for the target's log-transformed contextual diversity-weighted orthographic neighborhood density by-participants: The results of a likelihood ratio test comparing this model and the random intercepts model revealed that adding random slopes improved model fit compared to the random intercepts model ($\chi^2(2) = 560.8$, $p < 0.001$), justifying their inclusion for this dataset (Bates, Kliegl et al. 2015). Adding random slopes for the potential target's log-transformed contextual diversity by-participants instead of random slopes for neighborhood density by-participants failed to improve model fit compared to the random intercepts model ($\chi^2(2) = 4.7$, *n.s.*); adding such random slopes in addition to random slopes for neighborhood density by-participants failed to improve model fit compared to the model that included only random slopes for neighborhood density by-participants ($\chi^2(3) = 3.7$, *n.s.*), and so I do not analyze models with alternative random effects structures.

All fixed effects were significant and patterned in expected directions: Participants were slower to respond as target length increased ($\hat{\beta} = 0.32720$; $t(10{,}370) = 68.89$, $p < 0.001$), as well as with increases in participant's age ($\hat{\beta} = 0.01036$; $t(271.8) = 7.46$, $p < 0.001$). On the other hand, participants were faster to respond across the trials of a given session ($\hat{\beta} = -0.00003$; $t(102{,}500) = -3.85$, $p < 0.001$), across sessions generally ($\hat{\beta} = -0.00346$; $t(88{,}750) = -24.37$, $p < 0.001$), and across sessions completed within the same day ($\hat{\beta} = -0.02206$; $t(102{,}700) = -12.57$, $p < 0.001$).

Consistent with prior research and with the analyses already performed here, the effect of neighborhood density was significant ($\hat{\beta} = 0.00674$; $t(147.9) = 7.59$, $p < 0.001$), with participants responding slower to non-word targets as neighborhood density increased. Moreover, the effect of diacritic-based word-similarity was also significant ($\hat{\beta} = 0.04122$; $t(16{,}350) = 6.08$, $p < 0.001$):

That is, participants responded slower to non-word targets for which the addition or omission of a diacritic would create a real word, and this disadvantage increased as the potential real-word target increased in contextual diversity (Figure 2.5). Furthermore, because I included both orthographic neighborhood density and the potential target's contextual diversity value as fixed effects, I can be certain that the inhibitory effect of diacritic-based word-similarity is statistically independent from the inhibitory effect of neighborhood density on non-word lexical decision performance.

2.5.2 Accuracy analysis and results

To further assess how lexical processing performance for non-words is affected by the absence of diacritics that would make a non-word target into a real word, and the inclusion of superfluous diacritics whose absence would make a non-word target into a real word, I analyzed

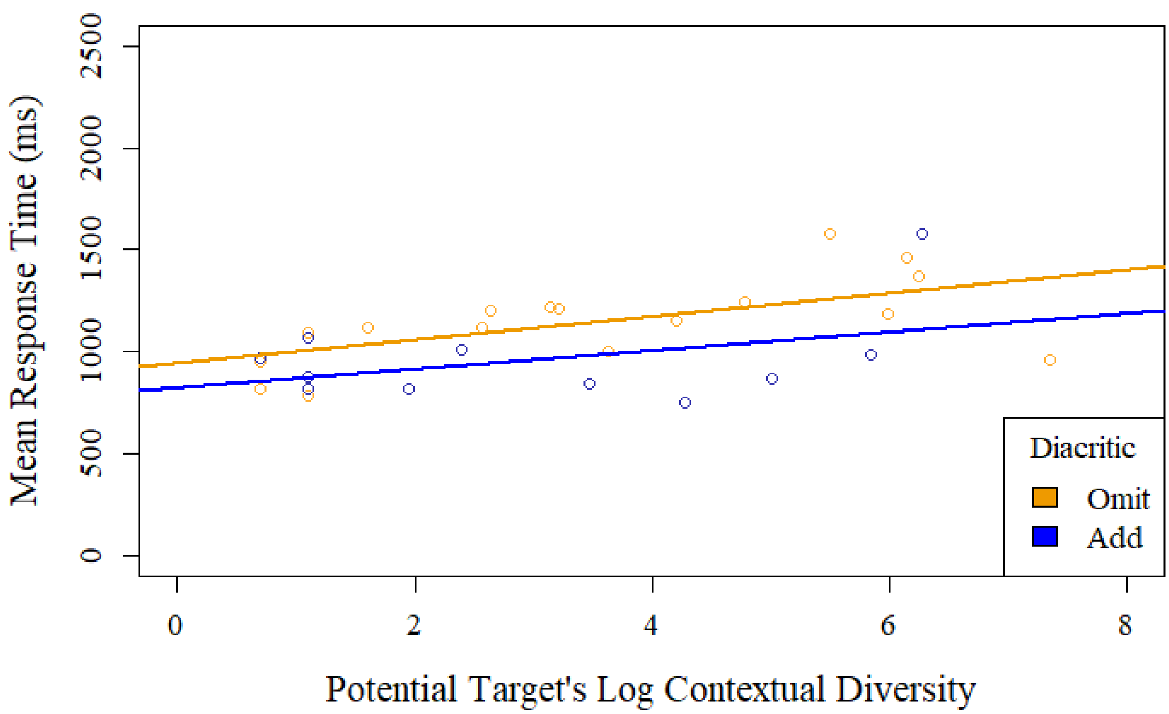

Figure 2.5. Mean RTs by the log contextual diversity of the potential real-word target (via diacritic omission or addition). Note that log RT was analyzed, but raw RT is presented here.

response accuracy to non-word targets on all trials (112,746 responses to 10,954 unique targets) by fitting a GLMER model using the binomial logit link function and the bobyqa optimizer in R (R Core Team 2021), using the lme4 package (Bates, Maechler et al. 2015). The model included response accuracy as the dependent variable (0 = incorrect, 1 = correct). As fixed effects, the model included the potential real-word target's log-transformed contextual diversity,[18] the target's log-transformed contextual diversity-weighted orthographic neighborhood density, and the following control predictors: log-transformed target length and participant's same-day session number.

The model included participant and target as random effects, and random slopes for the potential target's log-transformed contextual diversity by-participants: The results of a likelihood ratio test revealed that adding random slopes improved model fit relative to the random intercepts model ($\chi^2(2) = 10.6$, $p < 0.005$), justifying their inclusion for this dataset (Bates, Kliegl et al. 2015). Adding random slopes for the target's log-transformed contextual diversity-weighted orthographic neighborhood density by-participants instead of random slopes for the potential target's contextual diversity by-participants failed to improve model fit compared to the random intercepts model ($\chi^2(2) = 1.4$, *n.s.*), while including random slopes for both factors introduced convergence issues into the model, and so I do not analyze models with alternative random effects structures.

The effect of target length was again significant ($\hat{\beta} = -1.35810$; $z = -13.47$, $p < 0.001$), with participants responding less accurately as target length increased, whereas the effect of same-day session number was not significant ($\hat{\beta} = -0.05704$; $z = -1.86$, *n.s.*). The effect of neighborhood density was also significant ($\hat{\beta} = -0.17426$; $z = -17.01$, $p < 0.001$), with participants responding less accurately to non-word targets as orthographic neighborhood density increased. Moreover, the

[18] As in the RT analysis, comparison of the AIC values for random intercepts models in which the frequency of the potential real-word target was versus was not accounted for showed that the model in which frequency was accounted for provided a better fit to the dataset ($\Delta = 15.13$, $p < 0.001$), and so this is the variable that I analyze.

effect of diacritic-based word-similarity was significant ($\hat{\beta} = -0.69514$; $z = -6.50$, $p < 0.001$): That is, participants responded less accurately to non-word targets for which the addition or omission of a diacritic would create a real word, and this disadvantage increased as the potential real-word target increased in contextual diversity (Figure 2.6). Again, because I included both orthographic neighborhood density and the potential target's contextual diversity as fixed effects, we know that the inhibitory effect of diacritic-based word-similarity on response accuracy for non-word targets is statistically independent from the inhibitory effect of orthographic neighborhood density.

2.5.3 Discussion

The results of both analyses are consistent with my predictions: Non-words that resemble real words except in the addition or omission of diacritics were harder for MaltLex participants to

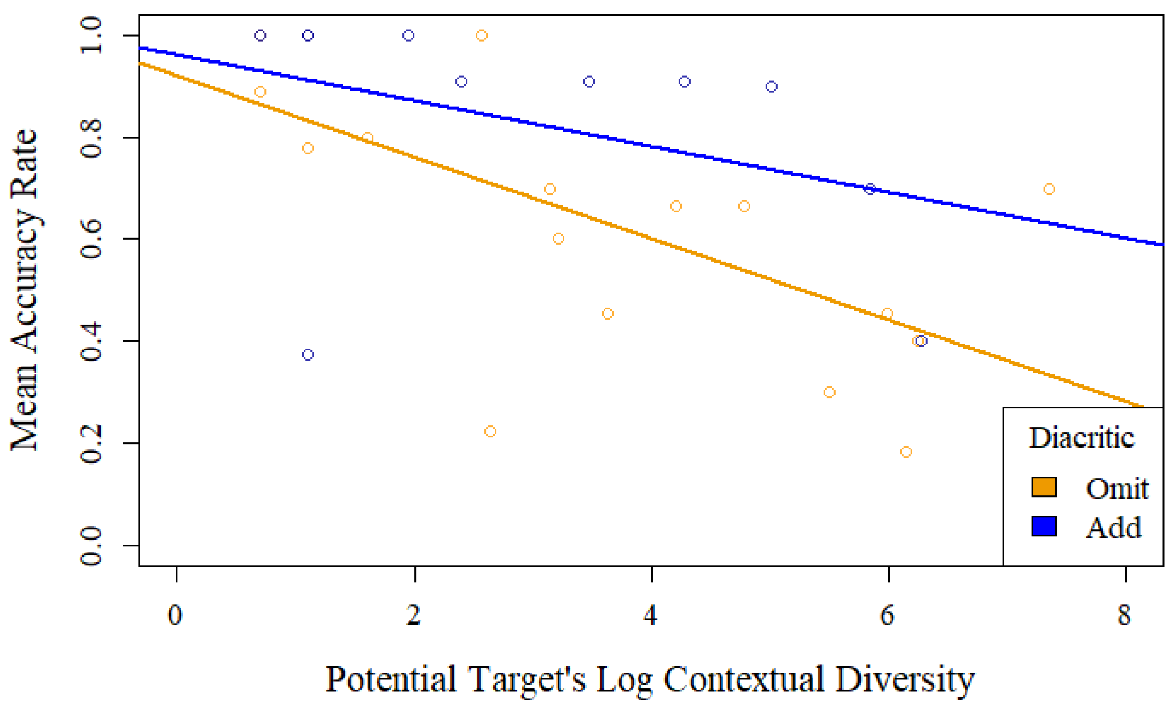

Figure 2.6. Mean accuracy by the log contextual diversity of the potential real-word target (as derived via the omission or addition of a diacritic).

reject, resulting in their responding slower and less accurately compared to other non-words. This suggests that Maltese readers subconsciously tolerate changes to a word's diacritics, such that they process these non-words as their potential real-word counterparts. Moreover, because the analyses included measures of both diacritic-based word-similarity and orthographic neighborhood density as fixed effects, we know that this effect is independent of the general neighborhood density effect.

That non-words became more difficult for participants to reject as the potential real-word counterpart increased in contextual diversity could reflect that high-contextual diversity real words were more easily activated by the non-word targets and competed more for recognition. On the other hand, it may simply be that weighting the potential real words by their contextual diversity value allows me to better approximate participants' language experience and so predict whether a given non-word target will be processed as its real-word counterpart. For example, participants are more likely to know a high-contextual diversity word like *fhimtx* 'I/you did not understand, realize' (which occurs in Korpus Malti 2.1 times per thousand contexts) than a low-contextual diversity word like *mazzu* 'they shuffled cards' (which occurs in Korpus Malti 0.004 times per thousand contexts). Consequently, participants will be more susceptible to processing the non-word *fħimtx* as the word *fhimtx* than they are to processing the non-word *mażżu* as the word *mazzu*. An analysis that includes a measure of the potential real-word counterparts' contextual diversity should be better able to predict which words participants will know and thus whether they could experience this processing ambiguity for a given target, producing the pattern of results observed here.

My results are consistent with those of Mifsud and Mitterer (2020), who hypothesize that Maltese readers process non-words without diacritics as their diacritic-bearing counterparts (e.g. the non-word *girja* as *ġirja*) because of their exposure to such diacritic-less counterparts in writing. I extend these findings both through MaltLex participants' self-reported difficulty in rejecting such non-words in the lexical decision task and through these statistical analyses: Even when they successfully identify such items as non-words, Maltese readers do so slower than on average due to interference from the real-word counterpart. Further, unlike Mifsud and Mitterer, I obtained these results while analyzing non-words that have a superfluous diacritic compared to a real-word counterpart, but I note that this difference must be interpreted with caution: Maltese readers do not normally encounter words that have been written with a superfluous diacritic, and it is possible that non-words that are formed by the omission of a diacritic are harder to reject/more likely to be processed as a real word than are those formed by the addition of a diacritic. Indeed, inspection of Figures 2.5–2.6 suggests that the processing disadvantages on RTs and response accuracy were greater for non-words formed by diacritic omission than for those formed by diacritic addition.

I tested this by re-running the analyses reported in Sections 2.5.1 and 2.5.2, with diacritic relation (contrast coded: –0.5 = diacritic added to form a non-word, 0 = default non-word, 0.5 = diacritic omitted to form a non-word) as well as the interaction of diacritic relation by the potential target's log-transformed contextual diversity as fixed effects. Neither the effect of diacritic relation ($\hat{\beta}$ = –0.07103; $t(11{,}840) = -0.77$, *n.s.*) nor the interaction were significant in the RT analysis ($\hat{\beta}$ = 0.04716; $t(14{,}030) = 1.85$, *n.s.*). Likewise, neither the effect of diacritic relation ($\hat{\beta}$ = –0.92597; $z = -0.59$, *n.s.*) nor the interaction were significant in the response accuracy analysis ($\hat{\beta}$ = –0.12149; $z = -0.31$, *n.s.*). It is likely that the low number of non-word targets that are related to a real word via diacritic-omission ($N = 17$) or diacritic-addition ($N = 11$) affords too lower of statistical power to detect any difference; a traditional factorial experiment comparing lexical decision performance for non-words derived via the omission versus addition of a diacritic may thus be needed here.

My results demonstrate an issue in constructing and selecting non-word targets for use in Maltese word recognition experiments, and so allow me to make the following recommendation to Maltese word recognition researchers: Non-word targets that resemble real words except in their diacritics pose special processing challenges for Maltese readers beyond the standard orthographic neighborhood density effect, and so such non-words need to be used as experiment stimuli with caution. Likewise, MaltLex users may need to avoid analyzing datapoints corresponding to these non-word targets or include diacritic-based word-similarity as a factor in their analyses of the non-word data (although the size of the MaltLex non-word dataset may effectively nullify the effect of diacritic-based word-similarity in any analysis of the non-word data). It seems probable that this finding will extend in some form to other writing systems in which sets of symbols differ only in diacritics, though investigating this beyond Maltese goes beyond the scope of the present study.

2.6 General discussion

The analyses reported here investigated the role that the orthographic form of a given word or non-word plays in lexical processing in Maltese. Consistent with prior research, I found that how orthographically similar a given target was to other Maltese words (here, operationalized as the target's orthographic neighborhood density) influenced visual lexical decision performance as measured in MaltLex, but that the effect of orthographic similarity differed for real-word versus non-word targets (e.g. Andrews 1989, Coltheart et al. 1977, Hendrix and Sun 2021, Yap et al. 2015) and for low-frequency versus high-frequency real-word targets (e.g. Andrews 1989, 1992): That is, increases in orthographic neighborhood density had an overall facilitatory effect for real-word targets but an inhibitory effect for non-word targets. Further, the orthographic neighborhood density advantage became smaller as real-word targets increased in frequency. To my knowledge, these analyses are the first to probe neighborhood density effects in Maltese visual lexical decision, as well as the first to explore neighborhood density effects for non-words in any Semitic language: The results of these analyses suggest that orthographic similarity effects in Maltese visual language processing pattern the same as in non-Semitic languages, despite its unique morphology, and they further support the validity of the MaltLex dataset for studying visual word recognition in Maltese

I also found that, beyond the overall orthographic neighborhood density effect, lexical decision performance for non-word targets was influenced by whether the target resembled a real Maltese word except in its diacritics (e.g. the non-words *gera* and *vaġun* resemble *ġera* 'he ran, roamed, flowed' and *vagun* 'wagon, carriage (of a train)', respectively, except in their diacritics): Consistent with previous lexical processing research on Maltese (Mifsud and Mitterer 2020) and confirming MaltLex participants' subjective impression that such targets were difficult to judge, I found that diacritic-based word-similarity made non-word targets more difficult for participants to reject in lexical decision compared to other non-word targets. This finding suggests that researchers who study visual word recognition in Maltese (and perhaps in other languages whose writing systems differentiate sets of letters only by their diacritics) must pay special attention to the manner in which non-word stimuli resemble real words, particularly in their diacritics, and avoid using non-words that differ from Maltese words in their diacritics as experiment stimuli.

To be sure, I have not nearly exhausted the full range of orthographic similarity effects on visual lexical decision performance that one could investigate using the MaltLex dataset. For example, individual differences in participants' language experiences and language processing abilities may mediate orthographic similarity effects on Maltese lexical processing. Related to this, for instance, Andrews and Hersch (2010) and Andrews and Lo (2012) have found that individual differences in spelling proficiency mediate the effects of neighbor priming in a series of English visual masked priming experiments (e.g. *jury* priming *FURY*): Better spellers exhibit an inhibitory effect of neighbor priming while worse spellers exhibit less inhibition or even a facilitatory effect of neighbor priming on lexical decision performance for high-density real-word targets, possibly because poorer spellers maintain less precise lexical representations that are more susceptible to preactivation by a similar visual input (cf. the lexical quality hypothesis: Andrews 2012; Perfetti 1992, 2007). The MaltLex dataset includes participants' scores on the Bilingual Language Profile (BiLP; Birdsong et al. 2012), which represents a composite, continuous measure of their relative language dominance, as well as participants' responses to individual BiLP questions that elicited self-assessments of their ability to speak, understand, write, and read Maltese versus English. As shown in Chapter 1, these measures could be used to further explore how individual differences in participants' language experiences and abilities mediate lexical processing in Maltese.

Chapter 3

Investigating the role of consonant letters in Maltese lexical processing

3.1 Introduction

In Chapter 1, I described the development and composition of the MaltLex database of Maltese visual lexical decision responses, and I used MaltLex data to compare the effectiveness of two measures of word frequency to predict lexical decision performance in Maltese: I found that a measure of word frequency that takes into account the number of documents in which a word appears ("contextual diversity") outperforms a traditional measure based on the number of times that a word appears in a corpus ("word frequency") at predicting lexical decision performance, consistent with prior megastudy-based research on a range of languages (e.g. English, Cantonese). In Chapter 2, I used MaltLex data to explore how the orthographic form of a word affects lexical processing in Maltese, focusing on how lexical decision targets resemble other Maltese words. In particular, I replicated several established orthographic neighborhood density effects in Maltese: I found that Maltese readers judge the lexicality of real words with more orthographic neighbors faster and more accurately than words with few neighbors; that this advantage diminishes as words increase in frequency; and that readers judge the lexicality of non-words with more neighbors slower and less accurately than those with few neighbors, likely due to increased competition with potential real-word candidates that are activated by the target. In this chapter, I conduct a novel visual masked priming experiment and again use the MaltLex dataset to explore how the consonant letters that comprise individual words influence lexical processing performance in Maltese.

Prior research has consistently shown that a word's consonant letters play a significant role in mediating lexical access and determining lexical identity. For example, visual masked priming studies have obtained priming by non-word strings that contain the same consonant letters as the target but different vowels (e.g. *jalu* priming French *JOLI* ‘pretty’; New et al. 2008, Soares et al. 2014) as well as by nonce letter string that consists of (a subset of) the consonant letters of a real-word target (e.g. *csn* priming *CASINO*; Anderson and Geary 2018, Duñabeitia and Carreiras 2011, Grainger et al. 2006, Peressotti and Grainger 1999), suggesting that exposure to a word’s consonant letters may be sufficient to permit lexical access. These priming effects disappear when the prime contains unrelated letters (e.g. *blcn* versus *bslcrn* priming French *BALCON* ‘balcony’; Peressotti and Grainger 1999), demonstrating the importance of the set of consonants in mediating lexical access. Further, these priming effects disappear when the relative order of the letters that comprise the prime differ from their order in the target (e.g. *arict* versus *acirt* priming *APRICOT*; Grainger et al. 2006), indicating that linear order may play an essential role in lexical processing here.

The aforementioned studies have focused on Indo-European languages. Maltese, a Semitic language, provides unique opportunities for exploring the role of consonant letters in mediating lexical processing and for understanding how orthographic effects on lexical processing are shaped by language-specific morphological patterns. Like other Semitic languages, native word stems in Maltese typically consist of two discontinuous morphemes: a consonantal root, which assigns the stem to a broad semantic field, and a vocalic/consonantal pattern, which contributes grammatical and thematic information to the stem. For example, Maltese words that contain the root *ktb* pertain to ‘writing’ (e.g. *kiteb* ‘to write; he wrote’, *kitba* ‘writing’, and *ktieb* ‘book’) and words containing the pattern *1V22ie3*[19] typically comprise agentive nouns (e.g. *kittieb* ‘writer’, *tebbieħ* ‘chef’, and *għalliem* ‘student’). These nonconcatenative morphemes interleave to form word stems, which may combine with concatenative morphemes (i.e. prefixes and suffixes) to form more complex words (e.g. *kittieb* ‘writer’ + *-a* ‘PL’ > *kittieba* ‘writers’). Experimental research has suggested that root morphemes mediate word processing in Semitic languages, including Hebrew (e.g. Deutsch

[19] I use numerals to indicate the positions of the consonants of the root morpheme, and *V* to indicate a short vowel.

et al. 1998; Feldman and Bentin 1994; Frost et al. 1997, 2000, 2005; Geary and Ussishkin 2019; Velan and Frost 2007, 2009, 2011), Arabic (e.g. Boudelaa and Marslen-Wilson 2001, 2004; Schluter 2013), and Maltese (e.g. Geary and Ussishkin 2018; Twist 2006; Ussishkin et al. 2015, 2016). For example, in Hebrew (Geary and Ussishkin 2019) and Maltese (Ussishkin et al. 2015), listeners are typically faster to judge the lexicality of a word when primed by another word that consists of the same root, though not by another word that consists of the same pattern, suggesting that exposure to one word activates other words that are derived from the same morphological root.

Yet not all Maltese words exhibit nonconcatenative morphology: Although it is a Semitic language, descending from Siculo-Arabic (Brincat 2011, Comrie 1991), Maltese has spent the last millennium developing in close contact with and frequently borrowing lexical items from a series of Indo-European languages, namely Sicilian and Latin, Italian, and English. Today, non-native borrowings comprise as much as half of the Maltese lexicon, with the bulk of non-native terms having been borrowed from Sicilian or Italian (Bovingdon and Dalli 2006, Brincat 2011, Comrie and Spagnol 2016). Except for some early borrowings that were integrated into the root and pattern morphological system (Mifsud 1995), non-native words typically do not consist of root and pattern morphemes and instead are subject only to concatenative morphological processes in Maltese.

Recently, Geary and Ussishkin (2018) used the visual masked priming paradigm to explore how consonant letters mediate lexical processing in Maltese, focusing on their role in the system of nonconcatenative morphology that characterizes the Semitic half of the Maltese lexicon (e.g. Borg and Azzopardi-Alexander 1997). They compared priming by nonce letter strings consisting of the target's letters for native, Semitic-origin Maltese words, for which these strings comprised the target's root morpheme (e.g. *frx* priming *FIREX* 'to spread'), versus non-native, non-Semitic words, for which these strings were non-morphemic (e.g. *pnġ* priming *PINĠA* 'to paint'). They found that participants were faster to judge the lexicality of Semitic Maltese words when primed by such letter strings, but not non-Semitic Maltese words. Unaware of prior research on "subset priming" effects, they interpreted these results as stemming from morphological activation: They hypothesized that, across reading experience, Maltese readers have abstracted out and stored root morphemes for Semitic Maltese words lexically, such that exposure to root morphemes in isolation activates their lexical representations and thus primes their morphological derivates (e.g. exposure to *frx* activates the lexical representation of *frx*, and so activates its morphological derivative *firex*). Their results conflict with those found for Indo-European languages: The lack of subset priming for non-Semitic words suggested that language-specific morphological patterns may determine the role that consonant letters play in constraining lexical access (cf. Anderson and Geary 2018).

In this chapter, I conduct a novel visual masked priming study as a follow-up to Geary and Ussishkin (2018), assessing whether subset priming effects occur for non-Semitic Maltese words. In particular, I test for priming effects by letter strings that comprise the consonants of a non-Semitic Maltese word when said strings also comprise a root versus when they are non-morphemic. In their original study, Geary and Ussishkin failed to control for whether the "root-letter" prime for non-Semitic targets comprised the letters of a Semitic root. If their explanation is correct, i.e. that the priming effect they observed stems from morphological activation, then it may be that they failed to observe priming effects for non-Semitic targets due to inhibition from the activated lexical representation of the root morpheme. If this is the case, then I may observe facilitatory priming by nonce strings corresponding to the consonant letters of a non-Semitic word when those letters do not correspond to a Semitic root morpheme, but not when they do correspond to a morpheme.[20]

[20] Consistent with this hypothesis, a post-hoc analysis of Geary and Ussishkin's results reveals that their participants were numerically faster to judge the lexicality of non-Semitic targets in the root-letter priming condition compared to the control condition when the root-letter primes did not comprise a Semitic root (667ms versus 693ms; effect = −26ms), but slower when they did comprise a Semitic root (735ms versus 696ms; effect = 39ms).

3.2 Factorial experiment 1 – Visual masked priming study

3.2.1 Methods

3.2.1.1 Participants

I collected lexical decision performance data from 44 self-identified native speakers of Maltese, all of whom were recruited from the University of Malta. Participants ranged in age from 18−52 years (*M* = 23.2 years). Twenty-three participants self-identified as female; 21 participants self-identified as male. Thirty-seven participants self-identified as right-handed; 7 participants self-identified as being left-handed. None of the participants whose data was analyzed reported having an uncorrected vision problem that could have affected visual lexical decision performance.

Data from an additional 15 participants was collected but is not analyzed due to their having achieved low accuracy rates during the experiment: They provided the correct response on less than 80% of real-word and/or non-word trials. Data from an additional 4 participants was collected but is not analyzed for various reasons: Two participants for self-identifying as L2 Maltese learners (i.e. not native speakers), and 2 participants for reporting having an uncorrected vision problem.

3.2.1.2 Materials

Participants judged the lexicality of 160 visual targets, including 80 real Maltese words and 80 Maltese-like non-words. All real-word targets comprised Maltese words of non-Semitic origin that contained three distinct consonant graphemes, ignoring gemination (e.g. *figura* 'figure', *tabella* 'street sign', and *kappell* 'hat' served as real-word targets; Appendix 3). Real-word targets ranged in length from 4–7 letters (*M* = 5.8 letters), in word frequency from 0.09−324.2 occurrences per million words in the 250-million-word Korpus Malti v3.0 (*M* = 21.0 occurrences per million words), and in contextual diversity from 0.07−79.9 occurrences per thousand contexts in the 253-thousand-document Korpus Malti v3.0 (*M* = 8.9 occurrences per thousand contexts). Real-word targets ranged in word-frequency weighted orthographic neighborhood density from 0.1−738.7 occurrences per million words in Korpus Malti v3.0 (*M* = 57.7 occurrences per million words), and contextual diversity-weighted orthographic neighborhood density from 0.1−144.1 occurrences per thousand contexts in Korpus Malti v3.0 (*M* = 24.2 occurrences per thousand contexts).

Each real-word target was paired with three primes (Table 3.1):

1. a "repetition" prime, which was identical to the target;
2. a "root-letter" prime; which was the target's three consonant letters; and
3. a "control" prime, which was another set of three consonant letters that shared no letters with the target's root-letter prime, and that did not correspond to a Semitic root.

To assess how root-letter priming interacts with the status of the target's "root" letters as a Semitic root, for half of the real-word targets the root-letter prime corresponded to a Semitic root (e.g. the "root" letters of *klima* 'climate', *klm*, occur as the root of Semitic-origin Maltese words like *kelma* 'word' and *kellem* 'to converse'). For the other half, the root-letter prime did not correspond to any Semitic root (e.g. the "root" letters of *figura* 'figure', *fgr*, are non-morphemic in Maltese). The

Table 3.1. Example prime-target pairs for which the target was a real Maltese word (Experiment 1). Note that all real-word targets comprised Maltese words of non-Semitic origin, and so the "root" letters of each target are non-morphemic for the target itself.

Priming condition	Does the root-letter prime correspond to a Semitic root?			
	Yes		No	
	Prime	Target	Prime	Target
Repetition	*klima*	*KLIMA* 'climate'	*figura*	*FIGURA* 'figure'
Root-letter	*klm*	*KLIMA*	*fgr*	*FIGURA*
Control	*rsf*	*KLIMA*	*snt*	*FIGURA*

results of a series of Welch's unequal variances *t*-tests comparing real-word targets across these categories revealed that they do not differ reliably in length (M_{Yes} = 5.9, M_{No} = 5.8 letters; $t(76.5) = -0.535$, *n.s.*), word frequency (M_{Yes} = 22.0, M_{No} = 19.9 occurrences per million words; $t(70.5) = -0.205$, *n.s.*), contextual diversity (M_{Yes} = 9.0, M_{No} = 8.8 occurrences per thousand contexts; $t(77.8) = -0.075$, *n.s.*), word frequency-weighted orthographic neighborhood density (M_{Yes} = 49.8, M_{No} = 65.5 occurrences per million words; $t(59.8) = 0.662$, *n.s.*), or contextual diversity-weighted orthographic neighborhood density (M_{Yes} = 23.6, M_{No} = 24.8 occurrences per thousand contexts; $t(76.9) = 0.161$, *n.s.*). Thus, it is unlikely that differences in any of these variables would underly differences in any priming effects that are observed across the two categories of real-word targets.

For each real-word target, I constructed a corresponding non-word target by replacing the original target's consonant letters with those of an orthographically legal but unattested root to create a possible non-word target with approximately equal orthographic neighborhood density (Appendix 4). I checked each nonce root against the Maltese lexical database Ġabra (Camilleri 2013), and a native speaker vetted each non-word target to ensure that it was an actual, plausible non-word that was not too similar to any real word; and that each nonce root was a nonce root. Non-word targets ranged in length from 4–7 letters (M = 5.8 letters), in word frequency-weighted orthographic neighborhood density from 0.0−1,811.0 occurrences per million words (M = 65.6 occurrences per million words), and in contextual diversity-weighted neighborhood density from 0.0−546.0 occurrences per thousand contexts (M = 26.6 occurrences per thousand contexts). The results of a series of Welch's unequal variances *t*-tests revealed that real and non-word targets do not differ in length ($t(158) = 0.0$, *n.s.*), word frequency-weighted neighborhood density ($t(114.7) = 0.297$, *n.s.*), or contextual diversity-weighted neighborhood density ($t(112.9) = 0.279$, *n.s.*).

Table 3.2. Example prime-target pairs for which the target was an orthographically legal, Maltese-like non-word (Experiment 1).

Priming condition	Prime	Target
Repetition	*kriga*	*KRIGA*
Root-letter	*krg*	*KRIGA*
Control	*vfd*	*KRIGA*

Each non-word target was similarly paired with three primes: a repetition prime, a root-letter prime, and a control prime (Table 3.2). Unlike for real-word targets, for non-word targets the root-letter and control primes never corresponded to the letters of a Semitic root morpheme.

Another twelve items were used as practice targets. This included six real-word targets of non-Semitic origin and six non-word targets derived from a real Maltese word in the manner described above. Two real-word practice targets and two non-word practice targets were paired with a repetition prime, two apiece with a root-letter prime, and two apiece with a control prime.

Three lists counterbalanced by priming condition were constructed using a Latin square design: Each list included 160 prime-target pairs (80 real-word targets and 80 non-word targets), with approximately one third of the targets (26–27 real-word targets and 26–27 non-word targets) occurring in each priming condition within a given list. Subjects were assigned randomly to one and only one list, and so they judged the lexicality of each target only once during the experiment.

3.2.1.3 Procedures

The study took place in a room at the Institute of Linguistics and Language Technology at the University of Malta. Participants were seated in front of one of three laptop computers during the experiment, facing away from other participants and seated approximately seven feet away from the nearest participant. Participants were instructed that they would see a series of possible words appear onscreen, one at a time, and that for each potential word they would need to decide whether it was a real Maltese word or not "as quickly and as accurately as possible"; no mention was made of the primes. Participants were instructed to respond by pressing one of two bumper buttons on a Logitech F310 Gamepad to register their responses. I used a gamepad to record responses to ensure consistent and accurate response time (RT) measurements (Witzel et al. 2013). As many as three participants completed the study simultaneously: In order to minimize potential distractions (e.g. other participants, loud noises), participants wore a pair of active noise-cancelling headphones during the experiment (see Chapter 1 Section 1.2.3 for more on the room layout).

After receiving the instructions and consenting to participate, participants completed the 12 practice trials followed by the 160 experiment trials. Items were presented pseudo-randomly such that participants never received more than four real-word targets or four non-word targets in a row (this was not mentioned to participants). Participants were able to take an untimed break following the final practice trial and following the 80th experiment trial (during which participants remained seated, being able to stretch, take a drink, etc.). After completing the experiment trials, participants were asked to quietly remain seated while other participants finished the experiment. Participants then completed a short language background questionnaire and were debriefed.

All stimuli were presented in DMDX (Forster and Forster 2003), using the visual masked priming paradigm (Forster and Davis 1984). On each trial, a forward mask consisting of seven hashtags "#######" appeared in the middle of the screen for 500 ms, then the prime replaced the mask and appeared in lowercase for 50 ms, and then the target replaced the prime and appeared in uppercase for 500 ms (Figure 3.1). All stimuli appeared in 12-point Courier New font: I used a monospaced font to make the appearance of the primes less obvious, and a seven-character forward

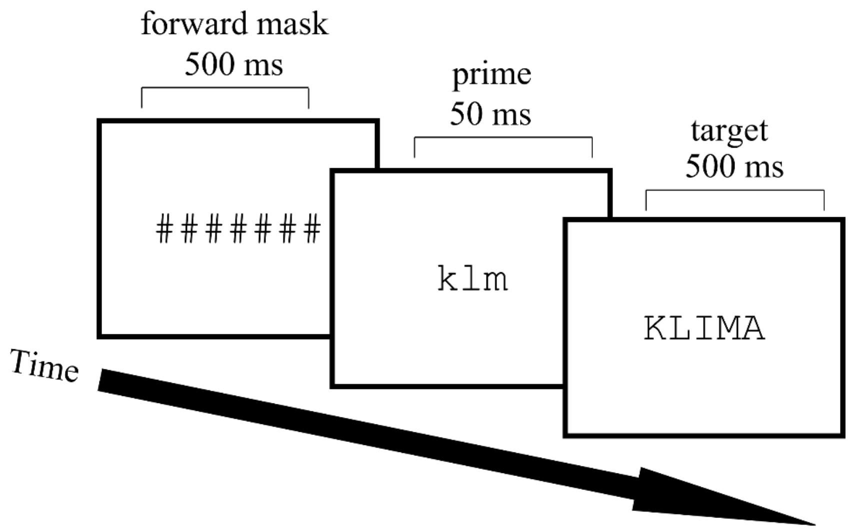

Figure 3.1. Illustration of a visual masked priming trial.

mask so that it would be as wide as the longest possible prime. DMDX recorded responses starting from target onset: If a response was not made within 3,000 ms, DMDX recorded a non-response and proceeded to the next trial. During the experiment, participants received feedback only on non-response trials ("Ebda risposta"), which DMDX displayed at the bottom of the screen for ~1,500 ms before proceeding to the next trial; otherwise, DMDX displayed a blank screen for ~1,500 ms.

3.2.2 Analyses and results

Prior to data analysis, I removed datapoints for one real-word target due to low accuracy, which suggested that it actually represents a non-word for most participants (29.5% of participants judged *milsa* 'spleen' as a non-word; including datapoints for this target does not qualitatively change the results of either analysis of the real-word data that is reported in this section). Thus, I analyzed a total of 3,476 lexical decision responses to 79 unique real-word targets, and 3,520 lexical decision responses to 80 unique non-word targets in the series of analyses reported below.

3.2.2.1 RT analysis and results (real-word targets)

I analyzed RTs on trials on which participants provided the correct response to a real-word target (3,195 lexical decision responses to 79 targets) by fitting a maximum likelihood-fitted LMER model using the bobyqa optimizer in R (R Core Team 2021), using the lme4 package (Bates, Maechler et al. 2015). I assessed the significance of fixed effects by using the lmerTest package (Kuznetsova et al. 2017) to simulate Satterthwaite approximations for degrees of freedom.

The model included negative reciprocal-transformed RT (i.e. –1,000/RT) as the dependent variable: I analyzed reciprocal RT to compensate for the effects of positive skew that are typical of masked priming experiments (Kliegl et al. 2010). The model included priming condition (levels: repetition, root-letter, and unrelated; reference: unrelated), whether the target's consonant letters comprise a Semitic root ("root-status"; levels: no versus yes; reference: no), and the interaction of priming condition by root-status, plus log-transformed target contextual diversity, log-transformed target length, and trial number as fixed effects[21]; and participant and target as random effects.

I included random slopes for priming condition by-targets: The results of a likelihood ratio test indicated that including random slopes for priming condition by-targets improved model fit compared to the random intercepts model ($\chi^2(5) = 14.7$, $p < 0.05$). In contrast, including random slopes for priming condition by-participants or root-status by-participants resulted in a singularity fit, and so I do not consider models with more complex random effects structures. Finally, I applied a model-based trimming procedure which, given a linear model, removes from the fitted dataset datapoints for which the residual standard score is more than 2.5 units from the mean of zero (Kenneth I. Forster p.c.), and then I re-fitted a new LMER model to the trimmed dataset. Applying this trimming procedure removed 52 datapoints from the dataset (1.6% of the data).

[21] Including log-transformed contextual diversity-weighted orthographic neighborhood density as an additional fixed effect either failed to improve model fit and/or resulted in a singularity fit in this and subsequent analyses (likely due to the lack of power afforded by the small size of this dataset), and so I omit this as a fixed effect in all analyses.

I obtained significant effects of contextual diversity, target length, and trial number: Participants responded faster as targets increased in frequency ($\hat{\beta}$ = –0.03631; $t(77.8) = –3.95$, $p < 0.001$) and across trials ($\hat{\beta}$ = –0.00025; $t(3,011.0) = –2.65$, $p < 0.01$), but slower as targets increased in length ($\hat{\beta}$ = 0.26090; $t(77.5) = 2.49$, $p < 0.05$). I also obtained significant priming effects, with participants responding faster to targets whose consonant letters did not correspond to a Semitic root when primed by a repetition prime (*M* = 651 ms; $\hat{\beta}$ = –0.09053; $t(73.0) = –4.86$, $p < 0.001$) or a root-letter prime (*M* = 669 ms; $\hat{\beta}$ = –0.03760; $t(71.8) = –2.16$, $p < 0.05$) compared to an unrelated prime (*M* = 689 ms). Though participants responded faster numerically to targets whose consonant letters do not correspond to a Semitic root (*M* = 669 ms) than to targets whose consonant letters do (*M* = 720 ms), the effect of root-status was not significant ($\hat{\beta}$ = 0.06540; $t(77.0) = 1.96$, *n.s.*). Moreover, the priming condition by root-status interaction failed to reach significance at either the repetition ($\hat{\beta}$ = 0.02856; $t(75.7) = 1.07$, *n.s.*) or root-letter levels ($\hat{\beta}$ = 0.02389; $t(75.0) = 0.95$, *n.s.*), providing insufficient evidence to suggest that priming effects differ across levels of root-status. Still, for targets whose consonant letters correspond to a Semitic root the difference in RTs between the root-letter (*M* = 729 ms) and unrelated priming conditions (*M* = 733 ms) is small (Figure 3.2).

3.2.2.2 Accuracy analysis and results (real-word targets)

I analyzed response accuracy on all trials with a real-word target (3,476 responses to 79 targets) by fitting a GLMER model using the binomial logit link function and the bobyqa optimizer in R (R Core Team 2021), using the lme4 package to fit the model (Bates, Maechler et al. 2015).

The model included response accuracy as its dependent variable (0 = incorrect, 1 = correct). As fixed effects, the model included priming condition (levels: repetition, root-letter, and unrelated; reference: unrelated), root-status (levels: no versus yes; reference: no), the priming condition by

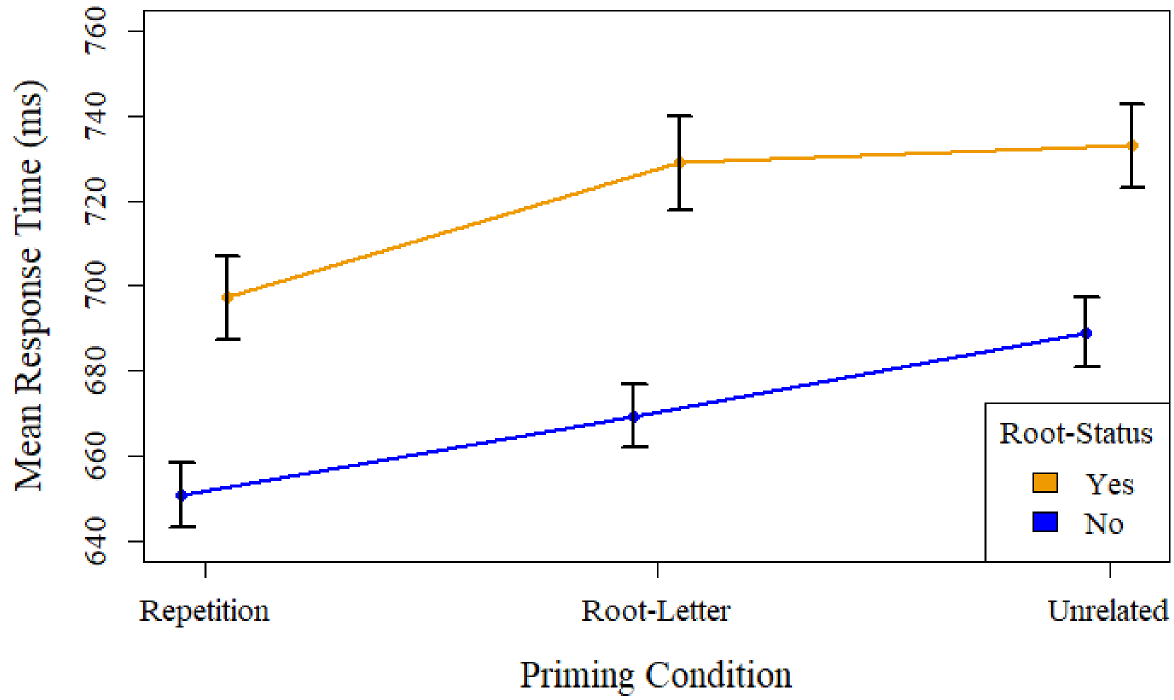

Figure 3.2. Mean RTs by priming condition for real-word targets. Error bars indicate standard error. Note that reciprocal RT was analyzed, but raw RT is presented here.

root-status interaction, log-transformed target contextual diversity, and log-transformed target length.[22] As random effects, the model included participant and target. Including random slopes for priming condition by-participants or by-targets produced a singularity fit, while including random slopes for root-status by-participants failed to improve model fit compared to the random intercepts model ($\chi^2(2) = 0.2$, *n.s.*), and so I report the results of the random intercepts model below.

As in the RT analysis, the effect of contextual diversity was significant, with participants responding more accurately as targets increased in frequency ($\hat{\beta} = 0.33011$; $z = 3.80$, $p < 0.001$). However, neither the effect of target length ($\hat{\beta} = –0.18866$; $z = –0.19$, *n.s.*) nor the effect of root-status were significant ($\hat{\beta} = –0.45148$; $z = –1.37$, *n.s.*): That is, there is insufficient evidence to show that response accuracy differed for targets whose consonant letters corresponded to a Semitic root

[22] Including trial number as a fixed effect here or in the response accuracy analysis for non-word targets caused the model to fail to converge, and so I do not include trial number as a fixed effect in any analysis of response accuracy.

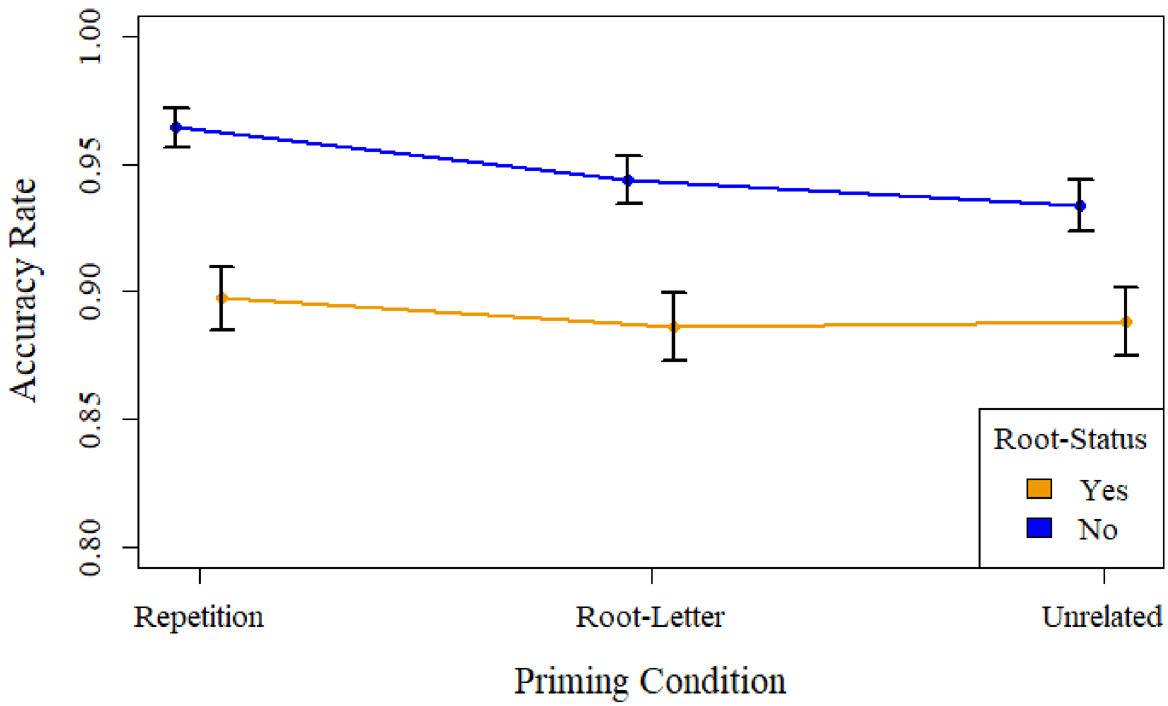

Figure 3.3. Accuracy by priming condition for real-word targets. Error bars indicate standard error.

(89.0%) versus for targets whose consonant letters did not (94.7%). The effect of priming condition was significant at the repetition level, with participants responding more accurately to targets for which the consonant letters did not correspond to a Semitic root that were primed by a repetition prime (96.4%) versus an unrelated prime (93.4%; $\hat{\beta} = 0.67903$; $z = 2.37$, $p < 0.05$). In contrast, the effect of priming condition was not significant at the root-letter level (94.4%; $\hat{\beta} = 0.18611$; $z = 0.73$, *n.s.*), nor were the effects of the priming condition by root-status interaction at the repetition ($\hat{\beta} = -0.56392$; $z = -1.59$, *n.s.*) or root-letter levels ($\hat{\beta} = -0.22555$; $z = -0.69$, *n.s.*; Figure 3.3).

3.2.2.3 RT analysis and results (non-word targets)

I analyzed RTs on trials on which participants provided the correct response to a non-word target (3,328 lexical decision responses to 80 targets) by fitting a maximum likelihood-fitted

LMER model using the bobyqa optimizer in R (R Core Team 2021), using the lme4 package (Bates, Maechler et al. 2015). I assessed the significance of fixed effects by using the lmerTest package (Kuznetsova et al. 2017) to simulate Satterthwaite approximations for degrees of freedom.

The model included negative reciprocal-transformed RT (i.e. –1,000/RT) as the dependent variable. As fixed effects, the model included priming condition (levels: repetition, root-letter, and unrelated; reference: unrelated), root-status (levels: no versus yes; reference: no), and the priming condition by root-status interaction, as well as log-transformed target length and trial number. As random effects, the model included participant and target. Including random slopes for priming condition by-participants failed to improve model fit compared to the random intercepts model ($\chi^2(5) = 6.4$, *n.s.*), while including random slopes for priming condition by-targets resulted in a model with a singularity fit, and so I report the results of the random intercepts model below.

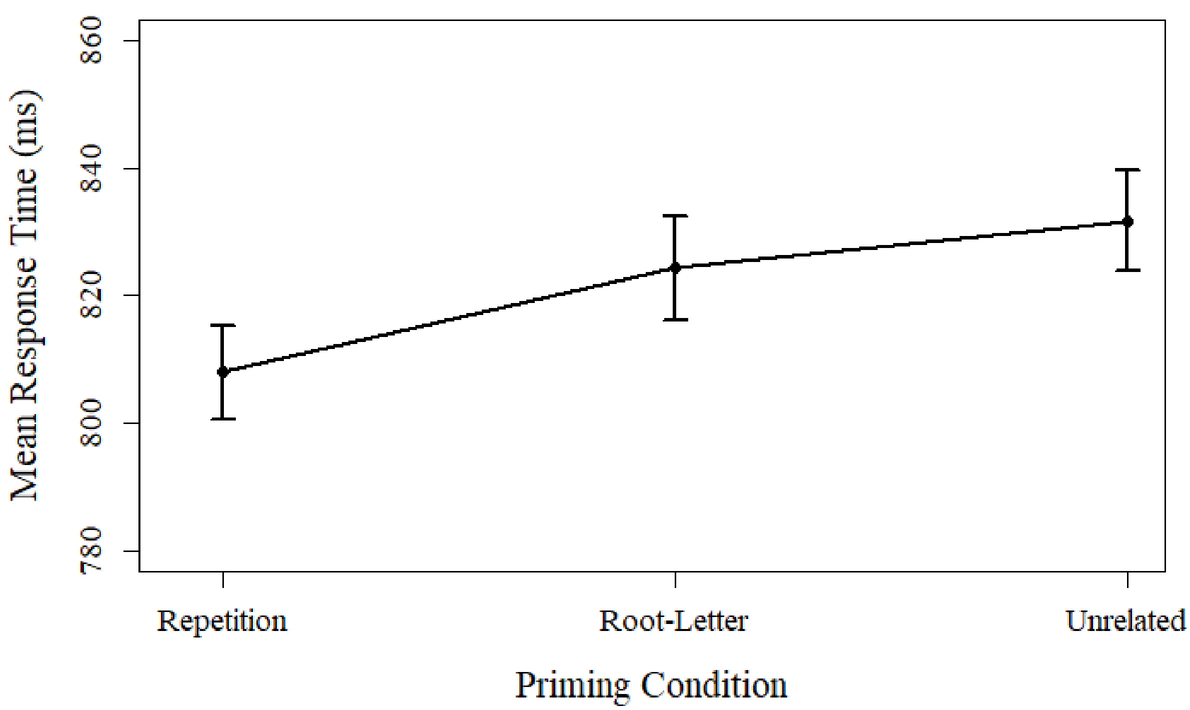

Figure 3.4. Mean RTs by priming condition for non-word targets. Error bars indicate standard error. Note that reciprocal RT was analyzed, but raw RT is presented here.

Finally, I applied the same model-based trimming procedure as in Section 3.2.2.1 to remove 54 datapoints (1.6% of the dataset), and then I re-fitted the model to the trimmed dataset.

As in the analysis of real-word targets, I obtained significant effects of target length and trial number on RTs for non-word targets: Participants responded slower as targets increased in length ($\hat{\beta}$ = 0.21690; $t(78.2) = 3.77$, $p < 0.001$) and across trials ($\hat{\beta}$ = 0.00035; $t(3{,}179.0) = 4.63$, $p < 0.001$). I also obtained significant effects of priming condition at the repetition ($\hat{\beta}$ = –0.03181; $t(3{,}155.0) = -3.74$, $p < 0.001$) and root-letter levels ($\hat{\beta}$ = –0.01861; $t(3{,}155.0) = -2.19$, $p < 0.001$): Participants responded faster to non-word targets when primed by a repetition prime (M = 808 ms) or by a root-letter prime (M = 824 ms) compared to an unrelated prime (M = 832 ms; Figure 3.4).

3.2.2.4 Accuracy analysis and results (non-word targets)

I analyzed response accuracy on all trials with a non-word target (3,520 responses to 80 targets) by fitting a GLMER model using the binomial logit link function and the bobyqa optimizer in R (R Core Team 2021), using the lme4 package to fit the model (Bates, Maechler et al. 2015).

The model included response accuracy as its dependent variable (0 = incorrect, 1 = correct). As fixed effects, the model included priming condition (levels: repetition, root-letter, and unrelated; reference: unrelated) and log-transformed target length. As random effects, the model included participant and target. Including random slopes for priming condition by-participants or by-targets resulted in a singularity fit, and so I report the results of the random intercepts model below.

Unlike in the analysis of response accuracy for real-word targets, the effect of target length failed to reach significance for non-word targets ($\hat{\beta}$ = –0.34373; $z = -0.34$, *n.s.*). Likewise, the effects of priming condition failed to reach significance at either the repetition ($\hat{\beta}$ = 0.06805; $z = 0.36$, *n.s.*) or root-letter levels ($\hat{\beta}$ = 0.20328; $z = 1.06$, *n.s.*), thus providing insufficient evidence to

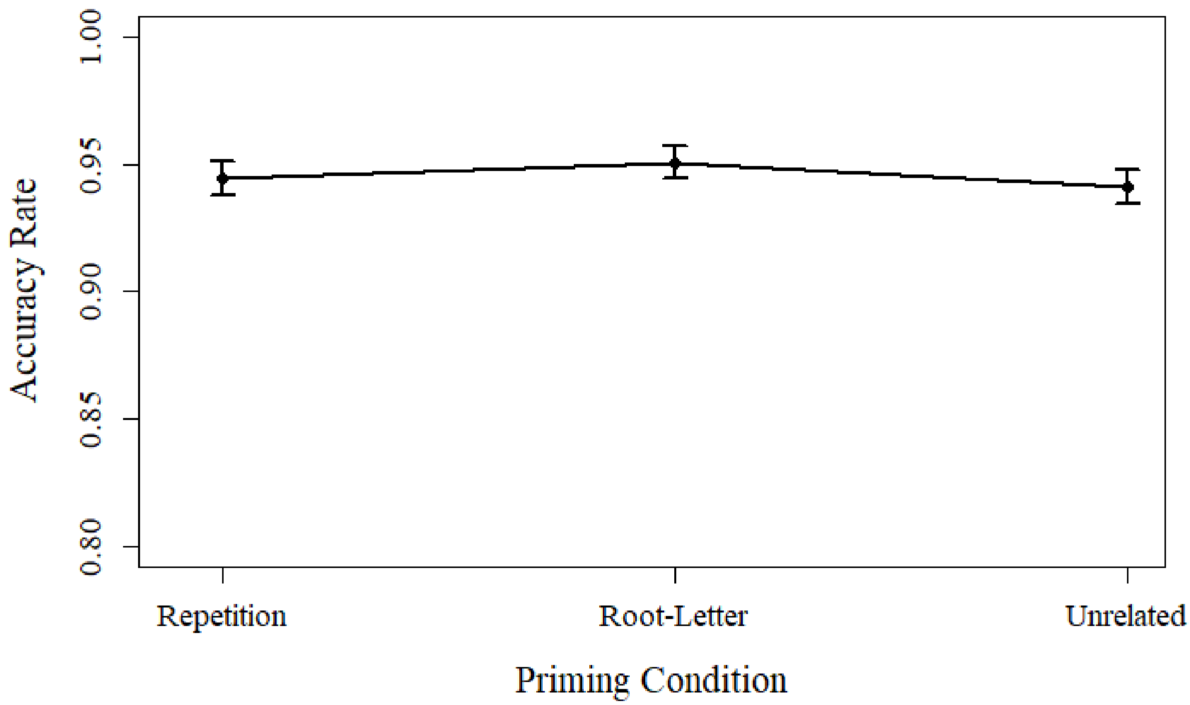

Figure 3.5. Accuracy by priming condition for non-word targets. Error bars indicate standard error.

support that accuracy differed for non-word targets when primed by a repetition prime (94.5%) or a root-letter prime (95.1%) compared to an unrelated prime (94.1%; Figure 3.5).

3.2.3 Discussion

These results are somewhat puzzling. On the one hand, I found a facilitatory priming effect on RTs for non-Semitic Maltese words when primed by strings consisting of their consonant letters, unlike Geary and Ussishkin (2018). These results suggest that, under some conditions, consonant letters play the same role in constraining and facilitating the recognition of non-Semitic Maltese words for which they are non-morphemic as they do for Semitic Maltese words for which they are morphemic: Whether prior exposure to a word's consonant letters facilitates lexical access is not contingent upon the status of those letters as a root morpheme. On the other hand, I failed to find evidence that priming effects differed according to the morphemic status of the target's consonant letters[23], and so the source of the discrepancy between my results and those of Geary and Ussishkin (2018) remain unresolved. I attempt to identify the source of this discrepancy in the next section by exploring the role that a word's consonant letters play in mediating visual word recognition in unprimed lexical decision in Maltese, using lexical decision data from the MaltLex dataset.

I also found a facilitatory priming effect on RTs for non-words when primed either by repetition primes or by strings consisting of their consonant letters. This may seem surprising, in that masked priming is usually assumed to tap into lexical representations, which non-words lack, and so masked priming effects are not usually expected to occur for non-words (e.g. Forster and Davis 1984, Twist 2006: 58). For example, neither Frost et al. (1997) nor Twist (2006) obtained masked priming effects for non-words in a series of Hebrew and Maltese visual masked priming lexical decision experiments, respectively. However, some masked priming effects are prelexical in nature, and masked priming effects do sometimes occur for non-word targets, though they tend to be smaller and therefore less stable/harder to detect than priming effects for real-word targets (e.g. Bodner and Masson 1997, Forster 1992, Masson and Isaak 1999; see Forster 1998 and Forster et al. 2003 for a review of masked priming effects for non-words). For example, masked repetition priming effects have been found for non-word targets that closely resemble a long word that has few neighbors (e.g. *fanulous* priming *FANULOUS*, which differs from *fabulous* by only one letter; Forster 1992, cited in Forster 1998), and it is worth noting that, in this study, non-word targets had been derived from Maltese words by replacing 1–2 consonant letters. Regardless, masked priming effects can occur for non-word targets, and I will explore these priming effects no further here.

[23] It is worth noting, again, that the root-letter priming effect was much smaller numerically for non-Semitic real-word targets whose consonant letters correspond to a Semitic root (M = 4 ms) than for those whose consonant letters do not correspond to a Semitic root (M = 20 ms). When I recode the data so that the reference level of root-status is "yes", and rerun the analysis reported in Section 2.2.1, I find that the effect of priming condition is no longer significant at the root-letter level ($\hat{\beta}$ = –0.01371; t(78.0) = –0.76, *n.s.*). Of course, the priming condition by root-status interaction remains non-significant at the root-letter level of priming condition ($\hat{\beta}$ = –0.02389; t(75.0) = –0.95, *n.s.*), and so this analysis too provides insufficient evidence to support that root-letter priming effects differ across levels of root-status.

3.3 Interim discussion – Analyzing consonant letter competition effects in MaltLex

Prior research has consistently demonstrated that consonant letters play a significant role in mediating lexical access and determining lexical identity. For example, visual masked priming studies have obtained facilitatory priming by non-word strings that contain the same consonant letters as the target but different vowels (e.g. *jalu* priming French *JOLI* 'pretty'), but not by non-word strings that contain the same vowels as the target but different consonants (e.g. *vobi* priming *JOLI*; New et al. 2008, Soares et al. 2014). Prior research has also found that exposure to a nonce letter string that consists of (a subset of) the consonant letters of a real-word target (e.g. *csn* priming *CASINO*) facilitates target recognition (e.g. Anderson and Geary 2018, Duñabeitia and Carreiras 2011, Grainger et al. 2006, Peressotti and Grainger 1999), although such priming effects disappear when unrelated letters are inserted into the prime (e.g. *blcn* versus *bslcrn* priming French *BALCON* 'balcony'; Peressotti and Grainger 1999) or the relative order of the letters that comprise the prime differ from their order in the target (e.g. *arict* versus *acirt* priming *APRICOT*; Grainger et al. 2006).

In contrast, exposure to a nonce letter string that consists of (a subset of) the target's vowel letters (e.g. *aia* priming *ANIMAL*) typically does not influence target recognition (e.g. Duñabeitia and Carreiras 2011). Duñabeitia and Carreiras have shown that the lack of priming by vowel letter substrings is not due to differences in letter frequency, as exposure to consonant letter substrings that consist of low-frequency consonants nonetheless facilitates target recognition; nor due to the repetition of letters in vowel substring primes, as facilitation persists with consonant substrings that contain repeated letters (e.g. *rtr* priming Spanish *FRUTERO* 'fruit'); nor due to phonological processing, as subset priming effects persist at short prime durations when phonological processing does not occur. Instead, Duñabeitia and Carreiras argue that the difference in occurrence of subset priming effects reflects the fact that more words share vowel substrings than consonant substrings (because languages typically have fewer vowel letters than consonant letters, and so allow fewer possible combinations of vowels than consonants). Thus, consonant letters more tightly constrain lexical competitors than do vowel letters, and so a priming effect emerges from consonant letter substrings but not necessarily from vowel substrings. In other words, the distributional properties of consonant letters versus vowel letters shapes their role in constraining lexical access.

In the following set of analyses, I explore the role that consonant letters play generally in mediating lexical access in Maltese by assessing how changes to the frequency of a target's consonant letters impact lexical decision performance within the MaltLex dataset. Prior to these analyses, I calculated the frequency of each MaltLex target's constituent consonant letters across all words in Korpus Malti v3.0 (Gatt and Čéplö 2013) by separately summing the word frequency and contextual diversity values of each word containing that set of consonant letters (in the same order). For example, for the target *kittieb* 'writer', I calculated the frequency of *kttb* by summing these values for all words that consist of this set of consonants, including *kiitieb* 'write' itself as well as *kitteb* 'to write regularly'. Likewise, for the target *privat* 'private (masculine)', I calculated the frequency of *prvt* by summing these values for all other words that contain this consonant set, such as *privat* 'private (mas.)' and its feminine form *privata* 'private (feminine)'. Consonant letter substrings that are attested among the MaltLex targets range in word frequency from 0–39,583.1 occurrences per million words (*M* = 68.6 occurrences per million words) and in contextual diversity from 0–1,579.3 occurrences per thousand contexts (*M* = 22.7 occurrences per thousand contexts).[24]

In the following analyses, I assess whether the frequency with which a given target's set of consonant letters occurs across words influences lexical decision performance for that target. In

[24] I also calculated the frequency of each target's constituent vowels, but I leave assessing whether the frequency of vowel letter substrings mediates lexical processing for future research. Attested vowel substrings range in word frequency from 0.004–77,356.7 occurrences per million words (M = 13,329.2 occurrences per million words) and in contextual diversity from 0.004–15,307.2 occurrences per thousand contexts (M = 4,347.3 occurrences per thousand).

particular, I compare consonant-subset frequency effects on lexical processing for real Maltese words versus non-words (Section 3.4), and for Semitic versus non-Semitic Maltese words (Section 3.5). If consonant letters mediate lexical access in Maltese, then I expect to find that participants respond faster and more accurately to real-word targets that consist of high-frequency consonants but slower and less accurately to non-words, since targets that consist of a high-frequency set of consonants should be more "word-like". On the other hand, if consonant letters play different roles in mediating access to Semitic words than non-Semitic words because of Maltese's morphological characteristics (cf. Geary and Ussishkin 2018), I may obtain greater consonant-subset frequency effects for Semitic Maltese words. Further, by including neighborhood density in these analyses, I assess whether any effect of consonant-subset frequency is independent of the more general orthographic neighborhood density effect on lexical processing. By exploring the role of consonant letters in mediating lexical access, I ultimately aim to reconcile the divergent masked priming results that I obtained in Section 3.2 with those reported by Geary and Ussishkin (2018).

3.4 Virtual experiment 7 – Analysis of consonant-subset frequency x target lexicality

In this analysis, I assess whether the frequency of a given target's set of consonant letters across all Maltese words influences lexical decision performance for that target by analyzing visual lexical decision responses from the MaltLex dataset. In particular, I compare consonant-subset frequency effects for real and non-word targets, while also assessing whether any consonant-subset frequency effect is independent of the general neighborhood density effect on lexical processing.

As in Chapters 1–2, I calculated the mean and standard deviation (SD) of the RTs for each participant and removed 6,723 datapoints for which RT was ±2.5 SDs from each participant's mean RT prior to data analysis. This reduced the dataset to 228,358 lexical decision responses to 21,994 unique targets (115,612 responses to real-word targets, 112,746 responses to non-words).

3.4.1 RT analysis and results

I analyzed RTs on trials on which participants provided the correct response (210,960 lexical decision responses to 21,900 unique targets, including 104,644 responses to 10,951 real-word targets and 106,316 responses to 10,949 non-word targets) by fitting a maximum likelihood-fitted LMER model using the bobyqa optimizer in R (R Core Team 2021), using the lme4 package (Bates, Maechler et al. 2015). I assessed the significance of fixed effects by using the lmerTest package (Kuznetsova et al. 2017) to simulate Satterthwaite approximations for degrees of freedom.

The model included log-transformed RT as its dependent variable. As fixed effects, the model included target lexicality (levels: real versus nonce, reference; real), the log-transformed contextual diversity value of the target's consonant letters (i.e. "consonant contextual diversity"), and the interaction of target lexicality by log-transformed consonant contextual diversity as fixed effects. The model also included the following control variables as additional fixed effects: the target's log-transformed contextual diversity-weighted orthographic neighborhood density, log-transformed target length, participant's age, trial number, overall session number, and session number within a given day. The model included participant and target as random effects.

I included random slopes for target lexicality by-participants, log-transformed consonant contextual diversity by-participants, and the lexicality by log-transformed consonant contextual diversity interaction by-participants. The results of a series of likelihood ratio tests indicated that including this set of random slopes improved model fit compared to the random intercepts model ($\chi^2(9) = 3{,}971.3$, $p < 0.001$), as well as compared to models that included random slopes only for target lexicality by-participants ($\chi^2(7) = 1{,}076.8$, $p < 0.001$) or for log-transformed consonant contextual diversity by-participants ($\chi^2(7) = 1{,}287.4$, $p < 0.001$; an additional model that included random slopes for target lexicality by-participants and for log-transformed consonant contextual diversity by-participants, but not for their interaction by-participants, failed to converge).

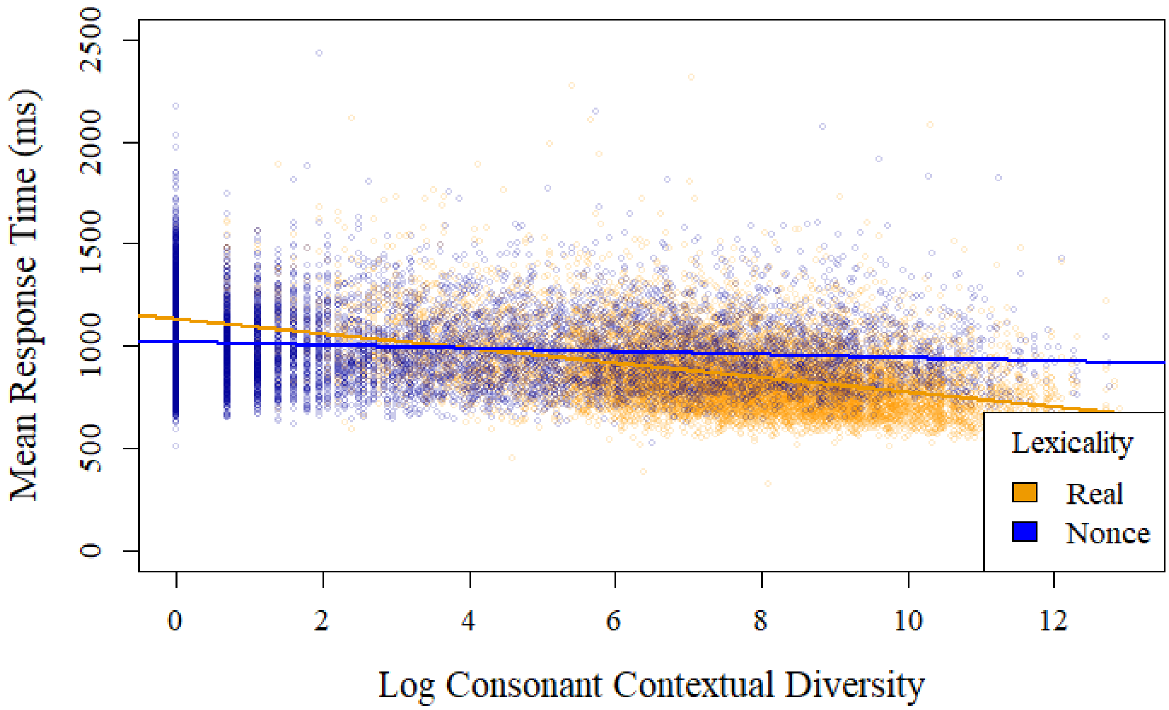

Figure 3.6. Mean RTs by log consonant contextual diversity. Note that log RT was analyzed, but raw RT is presented here.

All fixed effects were significant in the RT analysis: Overall, participants were slower to respond as target neighborhood density increased ($\hat{\beta}$ = 0.00334; $t(20{,}100) = 8.14$, $p < 0.001$) and as target length increased ($\hat{\beta}$ = 0.22120; $t(20{,}410) = 61.21$, $p < 0.001$). RTs increased both with participant's age ($\hat{\beta}$ = 0.00832; $t(413.0) = 7.54$, $p < 0.001$) and across trials ($\hat{\beta}$ = 0.00003; $t(202{,}000) = 5.28$, $p < 0.001$). RTs decreased across sessions generally ($\hat{\beta}$ = –0.00251; $t(181{,}400) = –24.27$, $p < 0.001$) and across sessions within the same day ($\hat{\beta}$ = –0.01852; $t(202{,}700) = –14.56$, $p < 0.001$).

The main effect of target lexicality was significant ($\hat{\beta}$ = –0.06465; $t(117.9) = –5.44$, $p < 0.001$), as was the main effect of consonant contextual diversity ($\hat{\beta}$ = −0.03516; $t(134.1) = −25.51$, $p < 0.001$): Participants responded faster to real-word targets as the frequency of their consonant letters across all Maltese words increased. However, the target lexicality by consonant contextual diversity interaction was significant ($\hat{\beta}$ = 0.03248; $t(148.4) = 25.51$, $p < 0.001$), with the facilitatory advantage of consonant contextual diversity diminishing for non-word targets (Figure 3.6).

To investigate this interaction further, I split the dataset by target lexicality and re-fitted separate LMER models to the real-word and non-word data. Each model included log-transformed RT as its dependent variable. As fixed effects, each model included log consonant contextual diversity, log contextual diversity-weighted orthographic neighborhood density, log target length, participant's age, participant's trial number, participant's overall session number, and participant's session number within a given day. Each model included participant and target as random effects, and random slopes for log consonant contextual diversity by-participants: The results of likelihood ratio tests comparing these models with random intercepts models suggest that they are justified for both the real-word ($\chi^2(2) = 471.6$, $p < 0.001$) and non-word datasets ($\chi^2(2) = 584.5$, $p < 0.001$).

For real-word targets, the effects of target neighborhood density ($\hat{\beta} = -0.00259$; $t(10{,}260) = -3.38$, $p < 0.001$), target length ($\hat{\beta} = 0.12160$; $t(10{,}100) = 24.42$, $p < 0.001$), age ($\hat{\beta} = 0.00434$; $t(267.4) = 3.35$, $p < 0.001$), trial number ($\hat{\beta} = 0.00008$; $t(99{,}970) = 11.20$, $p < 0.001$), session number ($\hat{\beta} = -0.00159$; $t(91{,}000) = -10.70$, $p < 0.001$), and same-day session number were significant ($\hat{\beta} = -0.01458$; $t(100{,}500) = -7.95$, $p < 0.001$). For non-word targets, the effects of target neighborhood density ($\hat{\beta} = 0.00902$; $t(10{,}160) = 19.35$, $p < 0.001$), target length ($\hat{\beta} = 0.34600$; $t(10{,}420) = 68.55$, $p < 0.001$), age ($\hat{\beta} = 0.01031$; $t(278.0) = 7.57$, $p < 0.001$), trial number ($\hat{\beta} = -0.00003$; $t(102{,}600) = -3.88$, $p < 0.001$), session number ($\hat{\beta} = -0.00345$; $t(88{,}490) = -24.34$, $p < 0.001$), and same-day session number were significant ($\hat{\beta} = -0.02225$; $t(102{,}800) = -12.68$, $p < 0.001$). Moreover, while the consonant contextual diversity effect was significant and facilitatory for real-word targets ($\hat{\beta} = -0.03586$; $t(166.3) = -24.90$, $p < 0.001$), it was non-significant (as well as numerically inhibitory) for non-word targets ($\hat{\beta} = 0.00110$; $t(135.4) = 1.27$, *n.s.*).

3.4.2 Accuracy analysis and results

I analyzed response accuracy on all trials (228,358 responses to 21,994 unique targets, including 115,612 responses to 11,040 real-word targets and 112,746 responses to 10,954 non-word targets) by fitting a GLMER model using the binomial logit link function and the bobyqa optimizer in R (R Core Team 2021), using the lme4 package (Bates, Maechler et al. 2015). The model included response accuracy as the dependent variable (0 = incorrect, 1 = correct). As fixed effects, the model included target lexicality, log-transformed consonant contextual diversity, and the interaction of target lexicality by consonant contextual diversity, along with log-transformed orthographic neighborhood density, log-transformed target length, and participant's within-day session number as control predictors. The model included participant and target as random effects.

I included random slopes for target lexicality by-participants: The results of a likelihood ratio test indicated that including such random slopes significantly improved model fit compared to the random intercepts model ($\chi^2(7) = 1{,}966.4$, $p < 0.001$). This model also had a significantly lower AIC value/better fit (AIC = 94,833.68) than did a model that included random slopes for log-transformed consonant contextual diversity by-participants (AIC = 96,322.91; $\Delta = 1{,}489.23$, $p < 0.001$), while models with more complex random effects structures (including random slopes for target-lexicality and log-transformed consonant contextual diversity by-participants) failed to converge, and so I do not analyze models with alternate random effects structures.

All fixed effects were significant: Overall, participants responded less accurately as target neighborhood density increased ($\hat{\beta} = -0.02340$; $z = -2.52$, $p < 0.05$) but more accurately as target length increased ($\hat{\beta} = 1.03271$; $z = 12.82$, $p < 0.001$). Participants responded less accurately across sessions that were conducted within the same day ($\hat{\beta} = -0.10508$; $z = -5.14$, $p < 0.001$). The effect of target lexicality was significant ($\hat{\beta} = 3.56627$; $z = 24.96$, $p < 0.001$), with participants responding more accurately to non-words (94.3%) than to real-word targets (90.5%). The effect of consonant

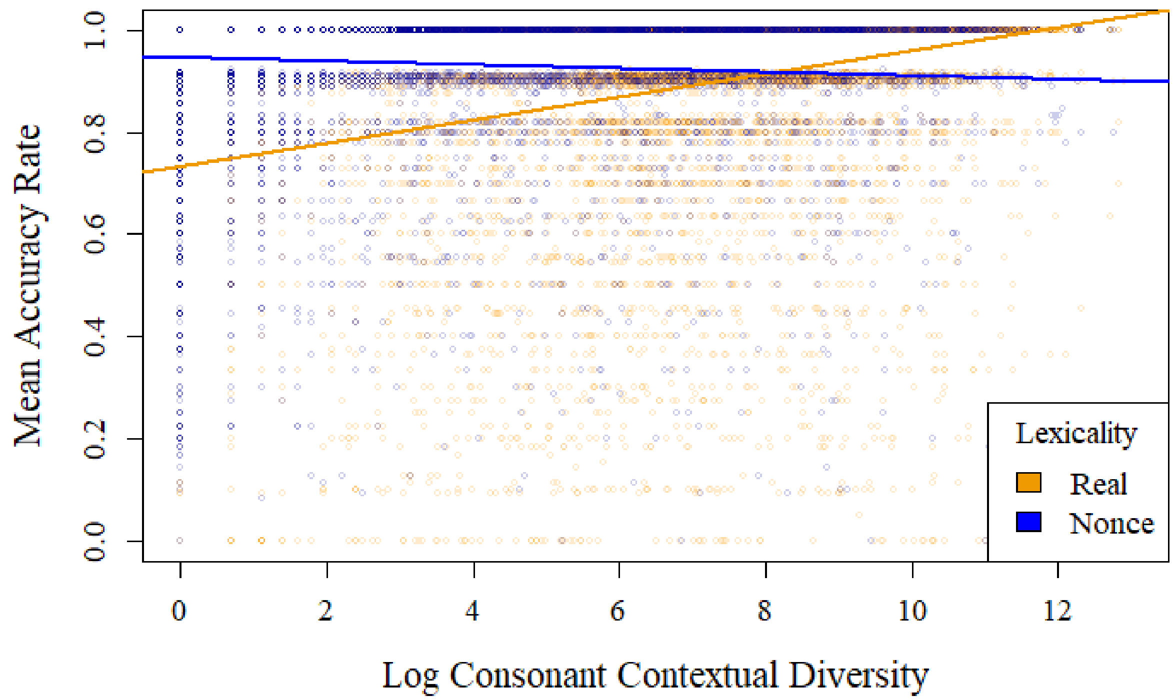

Figure 3.7. Mean accuracy by log consonant contextual diversity.

contextual diversity was likewise significant ($\hat{\beta}$ = 0.43921; z = 32.44, p < 0.001), with participants responding more accurately to real-word targets as the frequency of their consonant letters across Maltese words increased. Furthermore, the interaction of target lexicality by consonant contextual diversity was significant ($\hat{\beta}$ = –0.48227; z = –31.62, p < 0.001), with the advantage for targets with more frequent consonant letters diminishing for non-words compared to real words (Figure 3.7).

To investigate this interaction further, I split the dataset by target lexicality and re-fitted separate GLMER models to the real and non-word data. Each model included response accuracy as its dependent variable. As fixed effects, each model included log consonant contextual diversity, log orthographic neighborhood density, and log target length. I also included the participant's within-day session number as a fixed effect in the real-word model, but excluded it from the non-word model due to convergence issues. Each model also included participant and target as random effects, and random slopes for log consonant contextual diversity by-participants: Likelihood ratio tests show that including such random slopes improved model fit relative to the random intercepts model for both the real-word ($\chi^2(2) = 76.1$, $p < 0.001$) and non-word data ($\chi^2(2) = 10.9$, $p < 0.005$).

For real-word targets, the effects of orthographic neighborhood density ($\hat{\beta} = 0.11738$; $z = 7.43$, $p < 0.001$), target length ($\hat{\beta} = 3.28084$; $z = 28.89$, $p < 0.001$) and same-day session number were significant ($\hat{\beta} = -0.12464$; $z = -4.50$, $p < 0.001$): Participants responded more accurately to real-word targets as neighborhood density or length increased, but less accurately across sessions held on the same day. For non-word targets, the effects of orthographic neighborhood density ($\hat{\beta} = -0.15445$; $z = -14.94$, $p < 0.001$) and target length were significant ($\hat{\beta} = -1.80326$; $z = -16.79$, $p < 0.001$): Participants responded less accurately to non-word targets as neighborhood density or length increased. The effect of consonant contextual diversity was facilitatory for real-word targets ($\hat{\beta} = 0.42282$; $z = 22.41$, $p < 0.001$) but inhibitory for non-words ($\hat{\beta} = -0.10738$; $z = -11.29$, $p < 0.001$): That is, participants responded more accurately to real-word targets but less accurately to non-word targets as the frequency of their consonant letters across real Maltese words increased.

3.4.3 Discussion

The results of both analyses are consistent with my predictions and with the findings of prior studies that have demonstrated a role for words' consonant letters in mediating lexical access (e.g. Anderson and Geary 2018, Duñabeitia and Carreiras 2011, New et al. 2008, Peressotti and Grainger 1999): MaltLex participants responded faster and more accurately to real-word targets as their constituent consonant letters increased in frequency across Maltese words. In contrast, they responded less accurately to non-word targets as consonant-subset frequency increased (perhaps because of competition with real Maltese words that consist of the same set of consonants), while the RT advantage associated with increases in consonant-subset frequency for real-word targets diminished (and was effectively null) for non-word targets. Moreover, the effect of consonant-subset frequency was independent of the general orthographic neighborhood density effect (Wurm and Fisicaro 2014), for which I obtained expected effects for both real- and non-word targets.

3.5 Virtual experiment 8 – Analysis of consonant-subset frequency x stratum (real words)

In this analysis, I further assess the role that consonant letters play in mediating lexical access in Maltese by comparing consonant-subset frequency effects for Semitic and non-Semitic real-word targets. Further, I assess consonant-subset frequency effects in the context of models that include a measure of target frequency as a separate predictor, which I had excluded from the preceding analysis due to the inclusion of data for non-word targets (for which frequency is null, and so target frequency would have been highly correlated with target lexicality in the analysis).

3.5.1 RT analysis and results

To compare consonant-frequency effects on RTs for Maltese words of Semitic versus non-Semitic origin, I analyzed RTs to real-word targets on trials on which participants provided the correct response (103,848 lexical decision responses to 10,858 targets, including 61,198 responses to 6,445 Semitic words and 42,650 responses to 4,413 non-Semitic words) by fitting a maximum likelihood-fitted LMER model with the bobyqa optimizer in R (R Core Team 2021), using the lme4 package (Bates, Maechler et al. 2015). I assessed the significance of fixed effects by simulating Satterthwaite approximations for degrees of freedom using lmerTest (Kuznetsova et al. 2017).

The model included log-transformed RT as its dependent variable. As fixed effects, the model included the target's etymological stratum (levels: Semitic versus non-Semitic; reference: Semitic), its log-transformed consonant contextual diversity, and the interaction of target stratum by consonant contextual diversity. The model also included the following control predictors as fixed effects: log-transformed target contextual diversity, log-transformed contextual diversity-weighted target orthographic neighborhood density, log-transformed target length, participant's age, participant's trial number, participant's overall session number, and participant's session number within a given day. The model included participant and target as random effects.

I included random slopes for target stratum and for consonant contextual diversity by-participants: The results of a series of likelihood ratio tests revealed that adding random slopes for target stratum by-participants improved model fit compared to the random intercepts model ($\chi^2(2) = 220.1$, $p < 0.001$), as did adding random slopes for consonant contextual diversity by-participants ($\chi^2(2) = 472.0$, $p < 0.001$). Likewise, including random slopes for target stratum and for consonant contextual diversity by-participants improved model fit compared to the model that only included random slopes for target stratum by-participants ($\chi^2(3) = 453.8$, $p < 0.001$) and compared to the model that only included random slopes for consonant contextual diversity by-participants ($\chi^2(3) = 201.9$, $p < 0.001$). However, adding random slopes for the target stratum by consonant contextual diversity interaction failed to improve model fit compared to the current model ($\chi^2(4) = 4.3$, *n.s.*), indicating that such random slopes are not justified for this dataset (Bates, Kliegl et al. 2015). Thus, I do not consider models with more complex random effects structures here.

All fixed effects were significant and patterned in expected directions: Participants were faster to respond as targets increased in contextual diversity ($\hat{\beta} = -0.03498$; $t(11{,}450) = -34.47$, $p < 0.001$) and orthographic neighborhood density ($\hat{\beta} = -0.00150$; $t(10{,}240) = -2.03$, $p < 0.05$), but slower as targets increased in length ($\hat{\beta} = 0.16260$; $t(10{,}070) = 31.53$, $p < 0.001$). RTs increased as participants' age increased ($\hat{\beta} = 0.00437$; $t(269.8) = 3.37$, $p < 0.001$) and across the trials of a given session ($\hat{\beta} = 0.00008$; $t(99{,}770) = 11.24$, $p < 0.001$), but decreased across sessions generally ($\hat{\beta} = -0.00175$; $t(88{,}030) = -11.86$, $p < 0.001$) and across sessions that were completed within the same day ($\hat{\beta} = -0.01458$; $t(100{,}200) = -7.94$, $p < 0.001$). The effect of target stratum was significant,

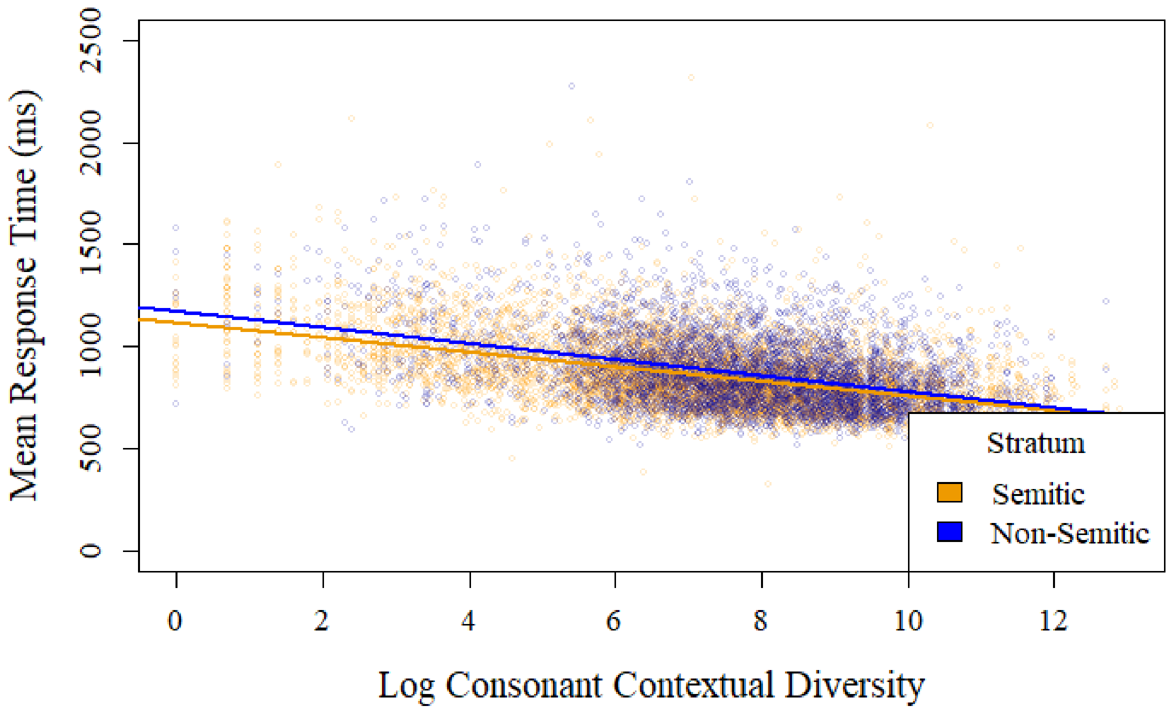

Figure 3.8. Mean RTs by log consonant contextual diversity for Semitic versus non-Semitic targets. Note that log RT was analyzed, but raw RT is presented here.

with participants responding slower to non-Semitic real-word targets compared to Semitic targets ($\hat{\beta}$ = 0.02989; $t(3{,}780) = 2.76$, $p < 0.01$). Likewise, the effect of consonant contextual diversity was significant, with participants responding faster to Semitic targets as their consonant letters became more frequent ($\hat{\beta}$ = –0.00503; $t(332.5) = -2.93$, $p < 0.005$). More importantly, the target stratum by consonant contextual diversity interaction was also significant, with the facilitatory advantage associated with increases in the frequency of a target's consonant letters being even larger for non-Semitic targets compared to Semitic targets ($\hat{\beta}$ = –0.00360; $t(10{,}760) = -2.80$, $p < 0.01$; Figure 3.8).

3.5.2 Accuracy analysis and results

To compare consonant-frequency effects on response accuracy for words of Semitic versus non-Semitic origin, I analyzed response accuracy to real-word targets on all trials on which the target was of Semitic or non-Semitic origin (114,609 lexical decision responses to 10,943 unique targets, including 68,086 responses to 6,514 Semitic targets and 46,523 responses to 4,429 targets of non-Semitic origin) by fitting a GLMER model using the binomial logit link function and the bobyqa optimizer in R (R Core Team 2021), using the lme4 package (Bates, Maechler et al. 2015).

The model included response accuracy as the dependent variable (0 = incorrect, 1 = correct). As fixed effects, the model included target's etymological stratum (levels: Semitic versus non-Semitic; reference: Semitic), its log-transformed consonant contextual diversity, and the target stratum by consonant contextual diversity interaction. The model included the following control predictors as additional fixed effects: log-transformed target contextual diversity, log-transformed target orthographic neighborhood density, log-transformed target length, and the participant's within-day session number. The model included participant and target as random effects.

I included random slopes for both target stratum and consonant contextual diversity by-participants: The results of a series of likelihood ratio tests revealed that adding random slopes for target stratum by-participants improved model fit compared to the random intercepts model ($\chi^2(2) = 185.3$, $p < 0.001$), as did adding random slopes for consonant contextual diversity by-participants ($\chi^2(2) = 55.8$, $p < 0.001$). Including random slopes for both target stratum and consonant contextual diversity by-participants improved model fit relative to the model that only included random slopes for target stratum by-participants ($\chi^2(3) = 45.8$, $p < 0.001$) and the model that only included random slopes for consonant contextual diversity ($\chi^2(3) = 175.2$, $p < 0.001$). Adding random slopes for the target stratum by consonant contextual diversity interaction did not further improve model fit ($\chi^2(4) = 3.7$, *n.s.*), and so I do not consider models with more complex random effects structures here.

All fixed effects were significant: Overall, participants responded more accurately to real-word targets as they increased in contextual diversity ($\hat{\beta} = 0.69928$; $z = 39.37$, $p < 0.001$), in neighborhood density ($\hat{\beta} = 0.06080$; $z = 4.28$, $p < 0.001$), and in length ($\hat{\beta} = 2.49749$; $z = 23.03$, $p < 0.001$), but less accurately across sessions that were conducted on the same day ($\hat{\beta} = -0.11422$;

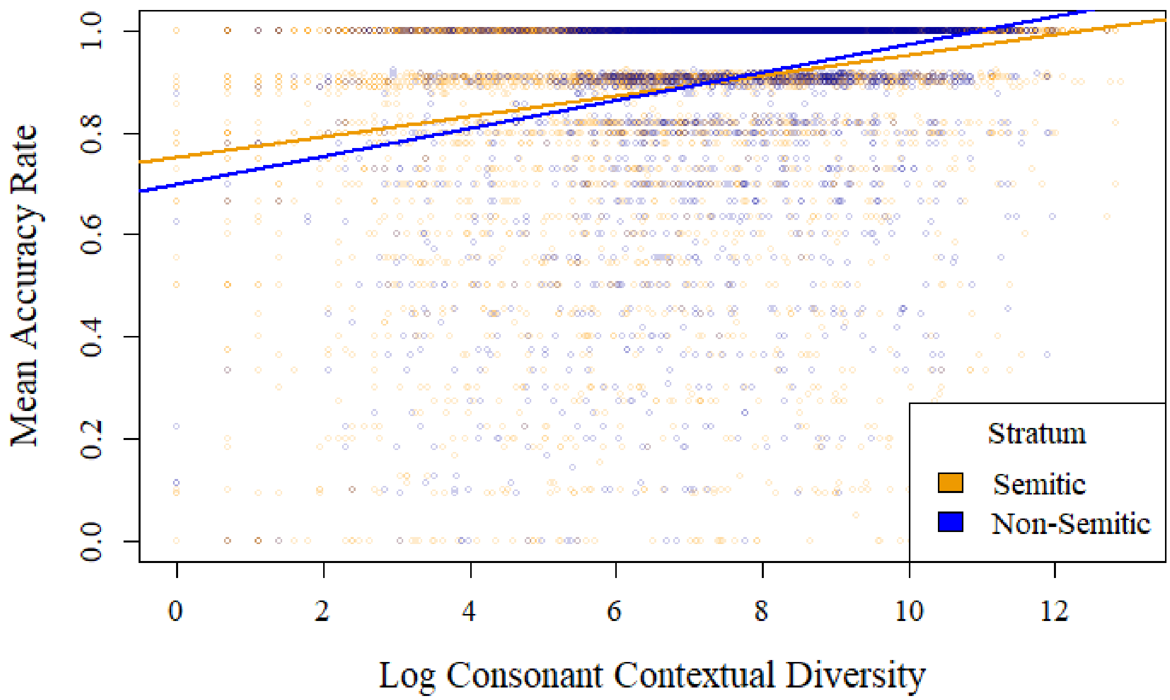

Figure 3.9. Mean accuracy by log consonant contextual diversity for Semitic and non-Semitic real-word targets. Note that this figure suggests that there is a facilitatory advantage associated with increases in consonant contextual diversity for both types of targets. However, this figure does not consider other predictors that were included in the analysis reported here, such as target contextual diversity. I discuss this in Section 3.5.3.

$z = -4.08$, $p < 0.001$). The effect of target stratum was significant, with participants responding less accurately to non-Semitic real-word targets compared to Semitic targets ($\hat{\beta} = -1.49133$; $z = -7.74$, $p < 0.001$). Likewise, the effect of consonant contextual diversity was significant, with participants responding <u>less</u> accurately to Semitic targets as their consonant letters became more frequent ($\hat{\beta} = -0.13106$; $z = -5.85$, $p < 0.001$). Further, the target stratum by consonant contextual diversity interaction was significant, with the inhibitory disadvantage associated with increases in consonant contextual diversity being smaller for non-Semitic targets compared to Semitic targets ($\hat{\beta} = 0.10609$; $z = 4.45$, $p < 0.001$; Figure 3.9). The results of this analysis regarding the effect of consonant contextual diversity are unexpected; I discuss them further in the next section.

3.5.3 Discussion

The results of this pair of analyses are difficult to interpret. On the one hand, I again found that MaltLex participants responded faster to real-word targets as their constituent consonant letters increased in frequency across Maltese words. Moreover, I found that the RT advantage that is associated with increases in consonant-subset frequency was greater for Maltese words of non-Semitic origin compared to Maltese words of Semitic origin, which may be related to the fact that participants responded slower to non-Semitic targets overall compared to Semitic targets. This also raises the possibility that the difference in priming effect by consonant letter substrings that Geary and Ussishkin (2018) found for non-Semitic and Semitic Maltese targets may be related to the frequency of their constituent consonant letters, which is greater for those that comprise roots.

On the other hand, I found that participants responded less accurately to real-word targets as consonant-subset frequency increased (although this disadvantage was smaller for non-Semitic targets compared to Semitic ones), which is contrary to expectations and to the results obtained in the previous analysis (Section 3.4.2). However, it is important to remember that, unlike in the prior analysis, the models analyzed here included contextual diversity as a fixed effect. For each fixed effect included among a linear model's fixed effects structure, the model identifies the amount of variance in the dependent variable that is independently attributable to that fixed effect (i.e. that is not attributable to other fixed effects included in the same model). Whenever two factors are highly correlated, they account for much of the same variance, and so the results of a linear model that include both as fixed effects may produce anomalous results like this, since the model is assessing the variance that is independently attributable to each fixed effect (e.g. Wurm and Fisicaro 2014).

Indeed, log target contextual diversity and log consonant-subset frequency are moderately correlated in both analyses ($r_{\text{RT Analysis}} = -0.505$, $r_{\text{Accuracy Analysis}} = -0.553$), likely in part because an increase in a word's contextual diversity necessarily results in an increase in the frequency of its subset of consonant letters. When I omit log contextual diversity from these models, consonant-subset frequency actually has a facilitatory effect on both RT ($\hat{\beta} = –0.03493$; $t(199.1) = –23.22$, $p < 0.001$) and response accuracy ($\hat{\beta} = 0.43454$; $z = 20.92$, $p < 0.001$). Further, the facilitatory advantage in response accuracy that is associated with increases in consonant-subset frequency is greater for non-Semitic targets compared to Semitic targets ($\hat{\beta} = 0.10276$; $z = 3.80$, $p < 0.001$). This is not to say that omitting contextual diversity is appropriate here, given the robust effect that word frequency has on lexical processing, but it does suggest that, in the context of these models, the unexpected direction of the consonant-subset frequency effect is due to collinearity.

3.6 General discussion

The analyses reported here investigated the role that a word's consonant letters play in lexical processing in Maltese. Consistent with prior research on Indo-European languages (e.g. Anderson and Geary 2018, Duñabeitia and Carreiras 2011, Grainger et al. 2006, Peressotti and Grainger 1999), I found that readers respond faster to non-Semitic Maltese words when primed by nonce strings consisting of their consonant letters. These results are contrary to those obtained by Geary and Ussishkin (2018), who found such priming only when the target was a Semitic-origin Maltese word (for which the consonant letters comprised a root morpheme) and not a non-Semitic Maltese word (for which they were non-morphemic). My results indicate that the occurrence of priming by consonant substrings is not contingent upon the target's being a Semitic or non-Semitic Maltese word per se, given that such priming effects obtain for some non-Semitic Maltese words.

Unaware of the research on subset priming effects in Indo-European languages, Geary and Ussishkin (2018) hypothesized that the asymmetrical priming effects that they obtained for Semitic versus non-Semitic Maltese words may be due to the status of the target's consonant letters as a root morpheme in the case of the former but not the latter, i.e. that their results reflect an effect of morphological priming by root morphemes. One way to assess this is to compare priming effects for non-Semitic targets by consonant letter substrings when they comprise root morphemes versus when they are non-morphemic, and indeed I manipulated the morphemic status of the prime in a novel visual masked priming experiment (Section 3.2), though they remained non-morphemic in the targets themselves: However, I did not find that the consonant-subset priming effect differed for primes that comprised a root morpheme versus primes that were non-morphemic, and so the source of the discrepancy in my results versus those of Geary and Ussishkin remained unresolved.

Another way to investigate these differences would be to compare priming effects for non-Semitic Maltese words for which the consonant letters actually comprise a root morpheme versus monomorphemic non-Semitic Maltese words. Until now, I have assumed that all non-Semitic Maltese words do not consist of root and pattern morphemes in the traditional Semitic sense, but this is inaccurate: Many early loanwords were integrated into the Semitic root and pattern system such that their consonants comprise a root morpheme that recurs with other morphological patterns in derived forms. Mifsud (1995) describes four stages in the integration of non-Semitic loanwords, ranging from their full integration into the root and pattern system (the earliest loanwords) to full non-integration and use of concatenative derivational morphology (the most recent loanwords). For example, the consonants of the early loanword *serp* ‘snake’ (from Sicilian *serpi* ‘snake’) recur in such forms as *srejjep* ‘a small snake’, *serrep* ‘to zig-zag’, and *sserrep* ‘to meander’ (Camilleri 2013: s.v. “serp”; Mifsud 1995: 59). If Geary and Ussishkin’s failure to obtain priming for non-Semitic words reflects the non-morphemic status of such strings in their non-Semitic targets, and not some other difference, then I should find priming for integrated non-Semitic Maltese words by letter strings that comprise the letters of their root morpheme. I attempted to design a second visual masked priming experiment in which I systematically manipulated whether non-Semitic Maltese targets consisted of an integrated root morpheme or not. However, I found that most integrated non-Semitic Maltese words were now archaic or of extremely low frequency (e.g. never occurring in Korpus Malti v3.0), or otherwise that contemporary Maltese speakers did not perceive them to contain a root morpheme. Thus, I was unable to conduct this experiment as planned.[25]

I further explored the role of consonant letters in lexical processing by assessing whether the frequency with which a word's set of consonant letters influences lexical decision performance. Using data from MaltLex, I found that participants responded faster and more accurately to real-word targets as their constituent consonant letters increased in frequency across all Maltese words, but less accurately to non-word targets as consonant-subset frequency increased. This measure captures information about how orthographically similar one word is to other Maltese words, yet this effect was independent of that of another measure of orthographic similarity: orthographic neighborhood density. However, consonant-subset frequency is moderately correlated with word frequency (i.e. because when a word increases in frequency, the frequency of its consonant letters across all Maltese words also increases in frequency) and including consonant-subset frequency and a measure of word frequency, namely contextual diversity, as simultaneous predictors in the same model introduces collinearity issues that impede the interpretation of their effects. While it may be possible to refine this measure of consonant-subset frequency to mitigate this correlation with contextual diversity, I leave implementing such changes for future research.

[25] I ran a pilot study in which 18 participants judged the lexicality of 66 non-Semitic Maltese words, including 33 words that had been integrated into the root and pattern system (e.g. *sserrep 'to meander'*, root: *s-r-p*; *lesta* 'to finish, complete', root: *l-s-t*; *pinġa* 'to paint, draw', root: *p-n-ġ*), and in which targets were primed by a repetition prime, their consonant letters, or a control prime consisting of an unrelated set of letters. Because of the targets' low frequencies, however, participants responded with poor accuracy during the pilot study (14 real-word targets had accuracy rates of 50% or lower), making the results of this pilot study difficult to interpret, hence I do not analyze them here.

Chapter 4

Conclusion

4.1 Summary of results and general discussion

In this dissertation, I described the construction of the MaltLex database of visual lexical decision responses for the Semitic language Maltese. I also conducted eight "virtual experiments" using data from MaltLex, as well as one novel visual masked priming study, in order to (1) validate the use of MaltLex for studying lexical processing in Maltese by replicating established findings from the lexical decision literature and (2) demonstrate several novel findings and so contribute to our understanding of lexical processing in Maltese and in Semitic languages more generally. Some of the main results of the analyses that I reported in this dissertation are summarized in Table 4.1.

In Chapter 1, I analyzed data from MaltLex to explore how word frequency and individual differences in language use and proficiency mediate lexical processing in Maltese. Unsurprisingly, I replicated the canonical word frequency effect: Maltese readers were faster and more accurate at judging the lexicality of Maltese words as they increased in word frequency. Moreover, I found that a measure of frequency that is based on the number of documents in which a word appears in a corpus (contextual diversity) better predicts lexical processing performance in Maltese than does a more traditional measure that is based on the total number of times that a word appears in a corpus (word frequency), which is consistent with the findings of megastudy-based research for non-Semitic languages, namely English (Brysbaert and New 2009), Cantonese (Tse et al. 2017), and Dutch (Keuleers, Brysbaert, and New 2010). I did not replicate Geary and Ussishkin's (2018) finding of a general RT advantage for Semitic Maltese words over non-Semitic Maltese words. Rather, I found that processing performance for Semitic versus non-Semitic words is sensitive to

Table 4.1. Summary of major findings from virtual experiments 1–8 (MaltLex).

1. Word frequency effects (Chapter 1)
    a. Maltese readers judge the lexicality of high-frequency words faster/more accurately, but contextual diversity better predicts performance than does word frequency.
2. Cognate advantage effects (Chapter 1)
    a. More English-dominant readers exhibit less of a processing difference for Semitic versus non-Semitic Maltese words than do more Maltese-dominant readers.
3. Orthographic neighborhood density effects (Chapter 2)
    a. Maltese readers judge words with more orthographic neighbors faster and more accurately, but this advantage diminishes for high-frequency words.
    b. Maltese readers judge non-words with more neighbors slower and less accurately.
4. Consonant letter similarity effects (Chapter 2)
    a. Maltese readers judge non-words that differ from a real word by the addition or omission of a diacritic slower and less accurately.
5. Consonant letter frequency effects (Chapter 3)
    a. Maltese readers judge words whose set of constituent consonant letters occur more frequently across the Maltese lexicon faster and more accurately.
    b. Maltese readers judge non-words whose consonants occur frequently less accurately.

differences in individuals' language backgrounds: More English-dominant MaltLex participants exhibited a smaller RT advantage in judging the lexicality of Semitic Maltese words compared to non-Semitic Maltese words than did the Maltese-dominant participants. That is, the more English-dominant participants, performing lexical decision in their non-dominant language, experienced less difficulty in judging the lexicality of non-Semitic words compared to Semitic words than did their counterparts. Taking lexical strata as an approximation of cognate status (i.e. non-Semitic words often have English cognates; Semitic words often do not), these results are consistent with prior research showing an advantage for cognates in lexical processing when readers perform in their non-dominant language (e.g. Costa et al. 2000, Jared and Kroll 2001, Poort and Rodd 2017).

In Chapter 2, I continued to analyze MaltLex data to explore how the orthographic form of a word mediates lexical processing in Maltese, turning my attention to non-words for the first time. Once again, I replicated several established orthographic neighborhood density effects: Maltese readers were faster and more accurate at judging the lexicality of Maltese words as they increased in neighborhood density and thus became more word-like in terms of their orthographic form, but this advantage for high-density words diminished as they increased in frequency. On the other hand, readers were slower and less accurate at judging the lexicality of non-words as they increased in neighborhood density and so became more word-like orthographically (e.g. Andrews 1989, 1992; Coltheart et al. 1977; Forster and Shen 1996; Hendrix and Sun 2021; Yap et al. 2015). I also found that Maltese readers were even slower and less accurate at judging the lexicality of non-words that differ from an existing Maltese word except by the addition or omission of a diacritic (e.g. compare the non-word *gera* with *ġera* 'he ran, roamed, flowed'), which is consistent with participants' subjective impression that such non-words were harder to reject. This result suggests that Maltese readers process certain pairs of letters alike during visual lexical processing (e.g. "g" and "ġ"), and thus reveals a unique manner in which non-words may appear especially word-like.

In Chapter 3, I conducted a novel visual masked priming study and continued to analyze MaltLex data to explore the role that the consonant letters play in Maltese visual lexical processing. Previous visual masked priming research on non-Semitic languages like English and Spanish has found that readers more efficiently process real words when exposed to nonce strings that consist of the target's consonant letters (e.g. *csn* priming *CASINO*; Anderson and Geary 2018, Duñabeitia and Carreiras 2011). In contrast, Geary and Ussishkin (2018) obtained comparable priming effects in Maltese only when Semitic Maltese words were primed by such nonce strings, which correspond to the target's morphological root (e.g. *frx* priming *FIREX* 'to spread'; root: *frx*), but not non-Semitic words for which such strings are non-morphemic (e.g. *pnġ* priming *PINĠA* 'to paint'). Unaware of previous research on subset priming effects in non-Semitic languages, they interpreted these priming effects as stemming from activation of the root morpheme's lexical representation: Because consonant strings are non-morphemic for (most) non-Semitic Maltese words, exposure to such strings does not elicit morphological activation and thus fails to facilitate target recognition. Contrary to Geary and Ussishkin (2018) but consistent with subset priming research, however, I did obtain facilitatory priming for non-Semitic Maltese words when primed by nonce strings that consist of their consonant letters, which morphological activation cannot explain. Returning to the MaltLex data, I then show that the frequency of a word's set of consonant letters across the Maltese lexicon mediates unprimed lexical processing: Maltese readers were faster and more accurate at judging the lexicality of real Maltese words as their set of consonant letters increased in frequency, but less accurate at judging non-words as consonant-frequency increased. Consonant-frequency may mediate these differences in the priming effects that I and other researchers have obtained.

I have replicated a wide range of well-established findings from the visual lexical decision literature in Maltese using data from MaltLex, thus validating the use of MaltLex to explore other aspects of visual word recognition in Maltese and in Semitic languages more generally. I have also demonstrated several visual lexical processing phenomena for the first time in Maltese, such as how orthographic similarity to real Maltese words (e.g. as estimated using neighborhood density) inhibits the processing of non-words (e.g. due to competition from potential real-word candidates).

The real contribution of this dissertation, however, is the MaltLex dataset itself, which I make freely available to other researchers who wish to further study visual lexical processing in Maltese: I have hardly scratched the surface of the analyses that could be performed using MaltLex, which, for example, could be used to explore the contributions of other lexical factors to lexical processing in Maltese (e.g. morphological complexity, part of speech) or to study any of the variables explored here in a more nuanced way (e.g. word length), among other possible analyses. I have performed minimal analysis of the non-word data, which may prove fruitful for future research. The data may find further application, for instance, in being used to evaluate the predictions of competing models of word recognition for lexical processing in Maltese or model word recognition computationally.

MaltLex is unique compared to previous lexical decision megastudies in at least two ways that go beyond simply focusing on a new language, and that provide novel directions for future applications of the MaltLex dataset: (1) Maltese is a Semitic language, and so is characterized by use of nonconcatenative, root and pattern morphology, the discontinuous nature of which provides challenges for models of word recognition that assume that language users decompose words into their constituent morphemes during language processing (e.g. Taft 1979, 2004; Taft and Forster 1975). Over half of the real-word targets used in MaltLex exhibit nonconcatenative morphology, whether because they are Semitic-origin words or non-Semitic borrowings that were integrated into the root and pattern system (Hoberman and Aronoff 2003, Mifsud 1995a), and the MaltLex dataset may be used to further explore the role of root and pattern morphology in Maltese lexical processing (cf. Geary and Ussishkin 2018, Twist 2006). MaltLex is unique among visual lexical decision megastudies in that it alone focuses on a language that productively uses nonconcatenative morphology. (2) All MaltLex participants were bilingual speakers of Maltese and English, and I accounted for differences in relative language proficiency and use among participants by collecting language dominance scores using the BiLP (Birdsong et al. 2012). MaltLex is unique among visual lexical decision megastudies in including this information, which I have here shown to influence visual word recognition in Maltese, and researchers may use the MaltLex dataset to further explore individual differences among bilingual readers in visual lexical decision performance.

4.2 Availability and structure of the final dataset

The size and scope of the MaltLex dataset affords novel opportunities for researchers to study visual lexical processing in Maltese, a relatively understudied and underresourced language, in ways that go beyond the analyses reported here: Researchers need not commit resources to designing and running new lexical decision studies, but may instead analyze subsets of the overall MaltLex dataset to answer their research questions. At the very least, such "virtual experiments", conducted by analyzing subsets of the total dataset (Keuleers and Balota 2015, Kuperman 2015), may serve as pilot results motivating the design of subsequent factorial studies. I anticipate that MaltLex will find further uses, for instance, in evaluating competing models of word recognition through Maltese, a language whose etymologically stratified lexicon contains a large proportion of words having nonconcatenative morphology (a unique characteristic among the languages on which megastudies have focused), or in modelling word recognition in Maltese computationally. Researchers can access the MaltLex trial-by-trial results as a .csv file, which contains the following information for each trial, from my website (http://jonathangeary.github.io/maltlex.html):

Participant-level information

- **Participant,** containing the ID number for the current participant.
- **Age,** indicating the participant's self-reported age.
- **Gender,** indicating the participant's self-reported gender.

- **Hand,** indicating the participant's self-reported handedness.
- **DominanceScore,** indicating the participant's composite dominance score as assessed via the Bilingual Language Profile (Birdsong et al. 2012). Negative scores indicate greater English dominance, positive scores indicate greater Maltese dominance, and a score of "0" indicates perfect balance between English and Maltese.

Session-level information

- **List,** indicating the list to which the participant was assigned for the current session.
- **Trial,** indicating the trial number within the current session.
- **Session,** indicating the overall session number for the current participant.
- **SessionDay,** indicating the participant's session number within the current day.
- **RT,** indicating the participant's RT (ms) on the current trial.
- **Response,** indicating whether the participant provided the intended response ("RIGHT") or not ("WRONG") on the current trial.

Target-level information

- **Target,** containing the ID number for the target judged on the current trial.
- **Word,** containing the item itself which served as the target on the current trial.
- **Lexicality,** indicating whether the target was intended to be judged as a real word ("Real") or as a non-word ("Nonce").
- **WF,** indicating the target's total number of occurrences across Korpus Malti v3.0 (Gatt and Čéplö 2013), as calculated by the author.

- **CD,** indicating the total number of unique documents in which the target occurs across Korpus Malti v3.0 (Gatt and Čéplö 2013), as calculated by the author.
- **N,** indicating the total number of unique words in Korpus Malti v3.0 which can be derived from the target via the substitution, addition, or deletion of a single letter. Because Korpus Malti contains some "unique words" which represent misspellings, foreign words, etc., I also include measures of neighborhood density which are weighted according to the WF (N_WF) and CD values of the target's neighbors (N_CD), expecting that the impact of such nonce neighbors should be offset by their low WF and CD values.
- **N_WF,** indicating the summed WF values of the target's neighbors.
- **N_CD,** indicating the summed CD values of the target's neighbors.
- **Length,** indicating the target's length in number of letters.
- **Lemma,** indicating the target's lemma. Lemmas were taken from Ġabra (Camilleri 2013) and supplemented with information from Aquilina (1987–1990).
- **Gloss,** listing an English translation for the target's lemma. Glosses were taken from Ġabra (Camilleri 2013) and supplemented with information from Aquilina (1987–1990).
- **Stratum,** indicating whether the target came from the Semitic ("Semitic") or non-Semitic stratum of the Maltese lexicon ("Non-Semitic"), or that the target was consistent with two words which come from different lexical strata ("Double") or that the target's etymology was unknown ("Uncertain"). Etymologies were taken from Aquilina (1987–1990).

- **Root,** indicating the target's root morpheme (if there is one, otherwise "No Root"). Roots were taken from Ġabra (Camilleri 2013) and supplemented with information from Aquilina (1987–1990) for words for which Ġabra did not list a root morpheme.[26]

Target-PoS-level information (unless indicated otherwise, all information was taken from Ġabra (Camilleri 2013) and supplemented with information from Aquilina (1987–1990))

- **Ġabra_PoS,** indicating the target's part of speech, as taken from Ġabra (Camilleri 2013). Where Ġabra contained multiple parts of speech for a target, I used the value consistent with the target's most common part of speech in Korpus Malti v3.0 (KM_Dom_Pos).
- **KM_Dom_PoS,** indicating the target's most common part of speech tag in Korpus Malti v3.0 as tagged using the Maltese Tagset v3.0 (Gatt and Čéplö 2013).
- **NomGender,** indicating gender for nominal targets (i.e. adjectives, determiners, nouns, numerals, and pronouns).
- **NomNumber,** indicating number for nominal targets.
- **AdpPerson,** indicating the person of a fused pronoun for some adposition targets (namely, forms of the preposition *ta'* 'of', wherein the preposition combines with a pronoun to specify the possessor; e.g. *tiegħi* 'of me', *tiegħu* 'of him', *tagħha* 'of her').

[26] Semitic-origin words in Korpus Malti v3.0 (Gatt and Čéplö 2013) are tagged according to their root morpheme, and I planned to include root morpheme frequency as part of the final dataset and to analyze its effect on lexical decision for Semitic Maltese words as part of the present study (cf. in a series of auditory lexical decision experiments on Jordanian Arabic, Wray 2016 found that root frequency had a facilitatory effect on lexical decision RTs for words from productive verbal patterns, supporting that such words are decomposed into their root (and pattern) morphemes during lexical processing). However, out of 6,029 Semitic targets which occur at least once in Korpus Malti v3.0 and which I identified as containing a root morpheme, I found that 3,518 (i.e. 58.4%) were never tagged as having a root morpheme in Korpus Malti, even when other words having the same root were tagged as such. Further, I found that some words were inconsistently tagged for their root morpheme. For instance, Korpus Malti v3.0 contains 36,195 occurrences of the word *bdew* 'they began', but only 36,066 occurrences were tagged as having a root (*b-d-j*). Further, Korpus Malti v3.0 contains 977 occurrences of its negated form *bdewx* 'they did not begin', none of which have been tagged as containing any root. Because such inconsistencies will skew the estimated frequency for each root, I do not include the frequency of each target's root morpheme as part of the current version of the MaltLex dataset.

- **AdpGender,** indicating the gender of a fused pronoun for some adposition targets.
- **AdpNumber,** indicating the number of a fused pronoun for some adposition targets.
- **SubPerson,** indicating the person of the subject for verb targets.
- **SubGender,** indicating the gender of the subject for verb targets.
- **SubNumber,** indicating the number of the subject for verb targets.
- **DirPerson,** indicating the person of the direct object for verb targets.
- **DirGender,** indicating the gender of the direct object for verb targets.
- **DirNumber,** indicating the number of the direct object for verb targets.
- **IndPerson,** indicating the person of the indirect object for verb targets.
- **IndGender,** indicating the gender of the indirect object for verb targets.
- **IndNumber,** indicating the number of the indirect object for verb targets.
- **Aspect,** indicating tense-aspect (e.g. perfect, imperfect, imperative) for verb targets.
- **Polarity,** indicating whether the form is negated ("neg") or not ("pos") for verb targets.

Additionally, researchers can access (1) a .csv file containing information about the items that were used as targets in the MaltLex study and (2) a .csv file containing the details of participants' responses to the language background questionnaire (Appendix 1) from the same website.

Appendix 1 – MaltLex language background questionnaire

All MaltLex participants completed the following language background questionnaire, which was adapted from the Bilingual Language Profile (BiLP; Birdsong et al. 2012), via Google Forms after completing their initial session of MaltLex. Questions 5–31 were analyzed to produce a composite, continuous measure of participants' relative Maltese–English language dominance.

I. BIOGRAPHICAL INFORMATION:

1. What is your age?
2. What is your gender?
3. Are you left-handed or right-handed?
4. Have you ever been diagnosed with a vision problem by a doctor? NOTE: If you have, but it has been corrected with glasses, contacts, surgery, etc., you should say "No".

II. LANGUAGE HISTORY. In this section, we would like you to answer some factual questions about your language history.

5. At what age did you start learning the following languages?
    a. Maltese
    b. English
6. At what age did you start to feel comfortable using the following languages?
    a. Maltese
    b. English
7. How many years of classes (grammar, history, math, etc.) have you had in the following languages (primary school through university)?
    a. Maltese
    b. English

8. How many years have you spent in a country/region where the following languages are spoken?
    a. Maltese
    b. English
9. How many years have you spent in a family where the following languages are spoken?
    a. Maltese
    b. English
10. How many years have you spent in a work environment where the following languages are spoken?
    a. Maltese
    b. English

III. LANGUAGE USE. In this section, we would like you to answer some questions about your language use. Total use for all languages in a given question should equal 100%

11. In an average week, what percentage of the time do you use the following languages with friends? (Total use for all languages should equal 100%.)
    a. Maltese
    b. English
    c. Other languages
12. In an average week, what percentage of the time do you use the following languages with family? (Total use for all languages should equal 100%.)
    a. Maltese
    b. English
    c. Other languages

13. In an average week, what percentage of the time do you use the following languages at school/work? (Total use for all languages should equal 100%.)
    a. Maltese
    b. English
    c. Other languages
14. When you talk to yourself, how often do you talk to yourself in the following languages? (Total use for all languages should equal 100%.)
    a. Maltese
    b. English
    c. Other languages
15. When you count, how often do you count in the following languages? (Total use for all languages should equal 100%.)
    a. Maltese
    b. English
    c. Other languages

IV. LANGUAGE PROFICIENCY. In this section, we would like you to rate your language proficiency by giving marks from 0 to 6. (0 = Not well at all, 6 = Very well)

16. How well do you speak Maltese?
17. How well do you speak English?
18. How well do you understand Maltese?
19. How well do you understand English?
20. How well do you read Maltese?
21. How well do you read English?
22. How well do you write Maltese?
23. How well do you write English?

V. LANGUAGE ATTITUDES. In this section, we would like you to respond to statements about language attitudes by giving marks from 0-6. (0 = Disagree, 6 = Agree)

24. I feel like myself when I speak Maltese.
25. I feel like myself when I speak English.
26. I identify with a Maltese-speaking culture.
27. I identify with an English-speaking culture.
28. It is important to me to use (or eventually use) Maltese like a native speaker.
29. It is important to me to use (or eventually use) English like a native speaker.
30. I want others to think I am a native speaker of Maltese.
31. I want others to think I am a native speaker of English.

Appendix 2 – Non-word targets with real-word counterparts that differ only in diacritics (MaltLex)

The 51 MaltLex non-word targets that resemble words listed in Ġabra (Camilleri 2013) when diacritics are added to the letters "g", "h", or "z", or when diacritics are omitted from the letters "ġ", "ħ", or "ż", which were the focus of the analyses reported in Section 3.3, are listed below with their potential real-word counterparts. Note that only 28 of these potential words occur in Korpus Malti v3.0 (Gatt and Čéplö 2013), and so only these 28 targets were treated differently from the other non-word targets in these analyses.

	Potential real-word counterpart			
Original non-word target	Word	Gloss	WF (occurrences per million words)	CD (occurrences per thousand contexts)
ga	ġa	'already'	8.935	6.156
fħimtx	fhimtx	'I/you did not understand, realize'	2.556	2.110
gera	ġera	'he ran, roamed, flowed'	2.616	2.031
zaqq	żaqq	'he fed, nourished; stomach, belly'	2.792	1.841
regax	reġax	'he did not repeat, return'	1.773	1.577
ġrafika	grafika	'graphic (feminine)'	1.858	1.367
jgibux	jġibux	'they are not bringing, calculating'	1.099	0.968
ġer	ger	'gear, gear lever'	1.007	0.585
tigbru	tiġbru	'you all are collecting, assembling'	0.538	0.466
vaġun	vagun	'wagon, carriage (of a train)'	0.686	0.281
jinhakmu	jinħakmu	'they are being ruled, colonized'	0.277	0.261
hawwar	ħawwar	'he flavored something, he elaborated'	0.148	0.146
seġa	sega	'saw'	0.169	0.122
jingarrux	jinġarrux	'they are not being carried, moved'	0.104	0.095
tahbu	taħbu	'you all are hiding, keeping secret'	0.116	0.087
ngemmgħu	nġemmgħu	'we are collecting'	0.064	0.051
mgarra	mġarra	'hand-driven cart'	0.052	0.047
jxenġlu	jxenglu	'they are swaying, leaning'	0.04	0.040
seġi	segi	'saws'	0.056	0.024

jithallu	jitħallu	'they are going sour'	0.016	0.016
ġażi	gażi	'telling on somebody, accusing somebody'	0.008	0.008
haru	ħaru	'they made an effort'	0.008	0.008
hwar	ħwar	'seasoning, spices and herbs'	0.032	0.008
sirħa	sirha	'you all are being cooked her'	0.008	0.008
tħeddet	theddet	'she was threatened'	0.008	0.008
mażżu	mazzu	'they shuffled cards'	0.004	0.004
migi	miġi	'coming, arrival, advent'	0.004	0.004
riega	rieġa	'to support'	0.004	0.004
fiehu	fieħu	'they had a sweet smell'	0	0
gammar	ġammar	'he burned or glowed with heat'	0	0
gergħa	ġergħa	'a swallowing (feminine)'	0	0
ġerres	gerres	'he upset'	0	0
għażżiet	għazziet	'moles (count plural)'	0	0
hann	ħann	'to have mercy'	0	0
jhemmx	jħemmx	'he is not clearing his throat'	0	0
jilħbu	jilhbu	'they are longing for, craving'	0	0
jinġass	jingass	'he is being kept under surveillance'	0	0
jisirgu	jisirġu	'they are shining, dazzling'	0	0
jitgħarrgu	jitgħarrġu	'they are starting to limp'	0	0
jmerrhuna	jmerrħuna	'they are yielding fruit us'	0	0
jpejjeż	jpejjez	'he is chirping'	0	0
leħmet	lehmet	'she was inspired'	0	0
nhara	nħara	'he was defecated'	0	0
nħellu	nhellu	'we are praising, lauding'	0	0
raħbuhom	rahbuhom	'they joined them a religious order'	0	0
raħbux	rahbux	'they did not join a religious order'	0	0
tirwih	tirwiħ	'act of fanning oneself; ventilation'	0	0
tkabbaż	tkabbaz	'he wrapped himself'	0	0
torħob	torhob	'you are joining a religious order'	0	0
żiek	ziek	'he made insulting remarks'	0	0
zwal	żwal	'trash'	0	0

Appendix 3 – Real-word stimuli used in Experiment 1

The real-word items that were used as targets in Experiment 1 are listed below, along with the three possible primes that they occurred with during the experiment (repetition, root-letter, and control) and whether their consonant letters/root-letter prime occurred as a root morpheme in Semitic Maltese words. Items marked with an asterisk “*” were omitted from data analysis for low accuracy.

			Primes			CD (occurrences per thousand contexts)		
Target	Gloss	Sem. root?	Repetition	Root-Letter	Control	Target CD	Root CD	CD-weighted neigh. density
BANJU	‘bathtub’	Yes	banju	bnj	dsk	8.689	140.476	0.944
FATTUR	‘factor’	Yes	fattur	ftr	ċng	17.474	0.225	31.92
XMARA	‘river’	Yes	xmara	xmr	pżl	3.41	0.071	80.947
RIKATT	‘blackmail’	Yes	rikatt	rkt	xfb	1.312	0.004	0.403
ISSOKTA	‘to continue’	Yes	issokta	skt	mlf	0.889	9.503	7.543
BRODU	‘broth’	Yes	brodu	brd	tvl	0.644	1.387	0.095
TABELLA	‘table’	Yes	tabella	tbl	fnr	24.851	0	9.902
KUTRA	‘blanket’	Yes	kutra	ktr	ġsf	0.494	5.931	64.951
SKEMA	‘scheme’	Yes	skema	skm	hbt	20.931	0	75.758
MUTUR	‘motor, engine’	Yes	mutur	mtr	kxs	16.695	0.043	69.863
ASSOLUT	‘absolute, utter’	Yes	assolut	slt	mrp	9.452	1.592	16.833
KOMDU	‘convenient’	Yes	komdu	kmd	tfr	10.333	0	15.743
FATALI	‘fatal, fateful’	Yes	fatali	ftl	żds	8.29	0.036	2.343
ORDNA	‘to order, give orders’	Yes	ordna	rdn	kżf	9.871	0.024	90.399
BORMA	‘cooking pot’	Yes	borma	brm	lsk	3.872	1.174	101.791
KLIMA	‘climate’	Yes	klima	klm	rsf	19.627	74.573	2.142
FORKA	‘gallows’	Yes	forka	frk	ġnl	0.628	6.105	83.203
* MILSA	‘spleen’	Yes	milsa	mls	dnt	0.217	0.126	6.346
TEORIJA	‘theory’	Yes	teorija	trj	nsk	6.496	0.032	2.363
SETTUR	‘sector’	Yes	settur	str	hdf	79.876	0.281	41.582
ABBILTÀ	‘ability, skill’	Yes	abbiltà	blt	pks	2.584	107.003	1.96
SOFRA	‘to suffer’	Yes	sofra	sfr	mzp	18.335	3.126	18.149
FRODI	‘fraud’	Yes	frodi	frd	nlm	10.448	3.592	2.45
ENFASI	‘emphasis’	Yes	enfasi	nfs	vċk	14.711	6.935	3.201

LIBERU	‘free, vacant’	Yes	liberu	lbr	kvm	10.894	0.004	8.152
ORBITA	‘orbit, limit’	Yes	orbita	rbt	kżs	0.237	83.685	0.249
METALL	‘metal’	Yes	metall	mtl	hdr	3.971	0.008	2.96
AĊĊERTA	‘to make certain’	Yes	aċċerta	ċrt	vfm	1.122	0.897	25.341
KONNESS	‘connected’	Yes	konness	kns	blż	5.884	0.178	15.142
TARIFFA	‘tariff, fare’	Yes	tariffa	trf	knd	8.78	0.695	25.294
ESEBIT	‘to exhibit’	Yes	esebit	sbt	lnp	1.134	0.269	5.2
BIDILLU	‘caretaker’	Yes	bidillu	bdl	swx	0.186	122.386	0.281
KATUBA	‘bass drum’	Yes	katuba	ktb	srl	0.162	111.531	0.929
SILLABA	‘syllable’	Yes	sillaba	slb	xrm	0.067	0.34	3.126
ANOMALU	‘anomalous’	Yes	anomalu	nml	kżt	0.146	0	0.929
BAROKK	‘baroque’	Yes	barokk	brk	psn	0.751	7.504	5.078
SALM	‘psalm’	Yes	salm	slm	bnż	0.395	0.288	51.729
ŻEBRA	‘zebra’	Yes	żebra	żbr	lsf	0.194	0.099	4.75
FOSSILI	‘fossil’	Yes	fossili	fsl	vżx	0.415	16.865	0.68
POTER	‘power’	Yes	poter	ptr	ċfk	36.856	3.948	64.931
FORZA	‘force’	No	forza	frz	dht	14.834	0	101.933
FIGURA	‘figure’	No	figura	fgr	snt	15.529	0	42.617
TINDA	‘tent’	No	tinda	tnd	wsk	5.212	0	76.363
STILLA	‘star’	No	stilla	stl	mkx	1.513	0	8.84
BERITTA	‘cap’	No	beritta	brt	psk	0.952	0	0.162
KAWZA	‘cause’	No	kawza	kwz	lsm	0.988	0	53.713
ANALIŻI	‘analysis’	No	analiżi	nlż	bjk	25.337	0	6.575
REĊTA	‘recital’	No	reċta	rċt	lnz	0.553	0	2.746
GWIDA	‘guide’	No	gwida	gwd	mbn	22.788	0	5.686
KAŻIN	‘casino, club’	No	każin	kżn	vdm	15.83	0	144.076
PIXXINA	‘swimming pool’	No	pixxina	pxn	rġf	5.335	0	0.415
REĠIM	‘regime’	No	reġim	rġm	skż	9.195	0	1.608
ĠENERU	‘gender’	No	ġeneru	ġnr	skp	4.987	0	4.228
LITRU	‘liter’	No	litru	ltr	snf	5.074	0	3.386
TENUR	‘tenor’	No	tenur	tnr	ċvk	3.541	0	3.118
ĊIFRA	‘price, quantity’	No	ċifra	ċfr	khż	18.082	0	23.82
ĊORMA	‘crew’	No	ċorma	ċrm	bnk	1.043	0	66.302
KARUS	‘piggybank’	No	karus	krs	fxn	0.478	0	0.818
ĊAVETTA	‘key’	No	ċavetta	ċvt	lnk	5.832	0	0.237
RAĠUNI	‘reason’	No	raġuni	rġn	mbt	76.173	0	34.987

LIMITU	‘limit’	No	limitu	lmt	kżx	2.193	0	31.604
NUTAR	‘notary’	No	nutar	ntr	ġfz	6.595	0	3.477
BORŻA	‘paper bag, pouch’	No	borża	brż	lsx	8.132	0	7.856
REGOLA	‘rule’	No	regola	rgl	kbn	14.415	0	59.672
MUNITA	‘money’	No	munita	mnt	rġt	11.108	0	5.252
PUDINA	‘pudding’	No	pudina	pdn	khl	0.486	0	0.964
SPALLA	‘shoulder’	No	spalla	spl	trz	3.849	0	1.391
MASSIV	‘massive’	No	massiv	msv	drt	1.162	0	3.521
KAPPELL	‘hat’	No	kappell	kpl	rnt	3.224	0	49.204
MISSIER	‘father’	No	missier	msr	bfl	24.092	0	18.375
STOFFA	‘material, stuff’	No	stoffa	stf	kżm	1.581	0	0.466
SIGARRU	‘cigar’	No	sigarru	sgr	bfp	0.186	0	0.411
EDUKAT	‘polite, educated’	No	edukat	dkt	mrn	0.454	0	3.675
LANDA	‘tin, can’	No	landa	lnd	sbġ	2.197	0	60.415
LIXKA	‘bait’	No	lixka	lxk	rġb	1.739	0	5.979
AĊĊENT	‘accent’	No	aċċent	ċnt	dżk	1.174	0	5.038
XABLA	‘sword’	No	xabla	xbl	wnr	1.28	0	0.838
KAZZOLA	‘saucepan, trowel’	No	kazzola	kzl	nfb	0.526	0	0.213
KAMERA	‘camera’	No	kamera	kmr	plt	2.651	0	95.453
FESTA	‘feast’	No	festa	fst	żld	31.632	0	57.301

Appendix 4 – Non-word stimuli used in Experiment 1

The non-word items that were used as targets in Experiment 1 are listed below, along with the three primes that they occurred with during the experiment (repetition, root-letter, and control).

	Primes			CD-weighted neighborhood density
Target	Repetition	Root-Letter	Control	(occurrences per thousand contexts)
KAŻLU	każlu	kżl	msp	1.833
FASSUR	fassur	fsr	xtk	0.17
SWARA	swara	swr	hdt	545.909
MISATT	misatt	mst	nlf	0.004
ISSOLFA	issolfa	slf	kżp	0.087
SFODU	sfodu	sfd	pkn	0.036
ĠABELLA	ġabella	ġbl	msf	24.946
SUKRA	sukra	skr	vċx	10.001
SKEFA	skefa	skf	lnċ	34.394
FUDUR	fudur	fdr	nlk	61.197
AŻŻOMUT	ażżomut	żmt	kvx	0
FOMDU	fomdu	fmd	swk	12.044
NARALI	narali	nrl	xtd	6.089
ORSMA	orsma	rsm	hdk	50.018
FOLMA	folma	flm	srt	79.726
KRIGA	kriga	krg	vfd	50.599
BOSKA	boska	bsk	xrt	74.225
XIRSA	xirsa	xrs	bnf	4.299
BEOTIJA	beotija	btj	ġnk	1.225
PEŻŻUT	peżżut	pżt	knr	0.822
ABBIXKA	abbixka	bxk	ċfd	0
ROŻTA	rożta	rżt	lnf	28.644
KLODI	klodi	kld	psr	0.68
ENWALI	enwali	nwl	kżd	1.087
ĊIBERU	ċiberu	ċbr	vżt	10.906
ORPINA	orpina	rpn	ċfm	0.095
RESALL	resall	rsl	mkp	0.458
AĊĊETMA	aċċetma	ċtm	xrġ	13.317
BORRESS	borress	brs	kjl	0.012
LARIFFA	lariffa	lrf	psm	8.784
EKEBIT	ekebit	kbt	rġw	1.178
TIFILLU	tifillu	tfl	bnr	0.126
LAKUBA	lakuba	lkb	rġx	1.051
TIŻŻANA	tiżżana	tżn	wsb	1.913
AFONALU	afonalu	fnl	mkf	0
FAXOKK	faxokk	fxk	blm	0.008
DARM	darm	drm	bjr	194.908
TEMRA	temra	tmr	ċfż	28.72
BOTTILI	bottili	btl	kjp	2.221
MOKER	moker	mkr	lsp	1.541
XORZA	xorza	xrz	snk	109.923
ĊIŻURA	ċiżura	ċżr	dhf	26.688

MINDA	minda	mnd	blr	66.808
PRILLA	prilla	prl	vżm	0.059
ĊEFITTA	ċefitta	ċft	mbx	0
KAJBA	kajba	kjb	lng	95.919
AVARIŻI	avariżi	vrż	bjx	0.087
ŻEKTA	żekta	żkt	lsr	2.075
GLIKA	glika	glk	mdċ	4.276
BAŻIT	bażit	bżt	snp	108.03
BITTINA	bittina	btn	psl	0.47
FEŻIT	feżit	fżt	mkn	0.126
ĊEMERU	ċemeru	ċmr	wsd	0.83
NISRU	nisru	nsr	vdk	14.798
TERUM	terum	trm	skd	10.831
DIMRA	dimra	dmr	ġfk	0.403
FOSMA	fosma	fsm	khd	61.726
FALUS	falus	fls	rġd	120.398
LAKETTA	laketta	lkt	mrġ	0.103
FAŻURI	fażuri	fżr	khs	1.189
SIMITU	simitu	smt	bfk	21.702
KUPAR	kupar	kpr	żnt	0.304
BOLFA	bolfa	blf	rġs	14.451
REKORA	rekora	rkr	żlx	21.848
FURITA	furita	frt	ġkl	6.654
KUMINA	kumina	kmn	dżb	0.719
STAŻŻA	stażża	stż	lnr	0
ĠASSIV	ġassiv	ġsv	wnt	3.098
BASSELL	bassell	bsl	nfk	0.087
NIŻŻIEX	niżżiex	nżx	pld	0.119
GROFFA	groffa	grf	lsb	1.869
VIGARRU	vigarru	vgr	bft	1.877
ERUVAT	eruvat	rvt	hsm	0.972
PAŻNA	pażna	pżn	sbd	40.238
TILKA	tilka	tlk	mrx	6.35
APPENT	appent	pnt	ċsr	6.556
MASLA	masla	msl	thk	25.685
MAŻŻONA	mażżona	mżn	ġrl	0.012
XAGERA	xagera	xgr	dsm	11.918
ĠESTA	ġesta	ġst	dżr	85.463

References

Adelman, J. S., Brown, G. D. A., and Quesada, J. F. (2006). Contextual diversity, not word frequency, determines word-naming and lexical decision times. *Psychological Science* 17: 814-823. https://doi.org/10.1111/j.1467-9280.2006.01787.x.

Adelman, J. S., and Brown, G. D. A. (2008). Modeling lexical decision: The form of frequency and diversity effects. *Psychological Review* 115: 214-229. https://doi.org/10.1037/0033-295x.115.1.214.

Adelman, J. S., Johnson, R. L., McCormick, S. F., McKague, M., Kinoshita, S., Bowers, J. S., Perry, J. R., Lupker, S. J., Forster, K. I., Cortese, M. J., Scaltritti, M., Aschenbrenner, A. J., Coane, J. H., White, L., Yap, M. J., Davis, C., Kim, J., and Davis, C. J. (2014). A behavioral database for masked form priming. *Behavior Research Methods* 46: 1052-1067. https://doi.org/10.3758/s13428-013-0442-y.

Agius, D. A. (1996). *Siculo Arabic*. London: Kegan Paul International.

Aguasvivas, J. A., Carreiras, M., Brysbaert, M., Mandera, P., Keuleers, E., and Duñabeitia, J. A. (2018). SPALEX: A Spanish lexical decision database from a massive online data collection. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2018.02156.

Aguasvivas, J., Carreiras, M., Brysbaert, M., Mandera, P., Keuleers, E., and Duñabeitia, J. A. (2020). How do Spanish speakers read words? Insights from a crowdsourced lexical decision megastudy. *Behavior Research Methods*. https://doi.org/10.3758/s13428-020-01357-9.

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In Petrov, B. N., and Csaki, B. F. (eds.), *Second international symposium of information theory*, 267-281. Budapest: Academiai Kiado.

Anderson, S. J., and Geary, J A. (2019). Form priming by discontinuous consonant letter strings in visual masked priming. *Proceedings of the 11th International Conference on the Mental Lexicon* 1: e060. https://doi.org/10.7939/r3-pvan-0083.

Andrews, S. (1989). Frequency and neighborhood effects on lexical access: Activation or search? *Journal of Experimental Psychology: Learning, Memory, and Cognition* 15: 802−814. https://doi.org/10.1037/0278-7393.15.5.802.

Andrews, S. (1992). Frequency and neighborhood effects on lexical access: Lexical similarity or orthographic redundancy? *Journal of Experimental Psychology: Learning, Memory, and Cognition 18*: 234−254. https://doi.org/10.1037/0278-7393.18.2.234.

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin & Review* 4: 439−461. https://doi.org/10.3758/BF03214334.

Andrews, S. (2012). Individual differences in skilled visual word recognition and reading: The role of lexical quality. In Adelman, J. S. (ed.), *Visual word recognition volume 2: Meaning and context, individuals and development*, 151−172. New York, NY: Psychology Press. https://doi.org/10.4324/9780203106976-15.

Andrews, S., and Hersch, J. (2010). Lexical precision in skilled readers: Individual differences in masked neighbor priming. *Journal of Experimental Psychology: General* 139: 299−318. https://doi.org/10.1037/a0018366.

Andrews, S., and Lo, S. (2012). Not all skilled readers have cracked the code: individual differences in masked form priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 38: 152−163. https://doi.org/10.1037/a0024953.

Aquilina, J. 1987–1990. *Maltese-English-Maltese dictionary*. Malta: Midsea Books.

Baayen, R. H., Milin, P., Đurđević, D. F., Hendrix, P., and Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review* 118: 438-481. https://doi.org/10.1037/a0023851.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., and Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods* 39: 445-459. https://doi.org/10.3758/bf03193014.

Bates, D., Kliegl, R., Vasishth, S., and Baayen, H. (2015). Parsimonious mixed models. *arXiv:* 1506.04967v1. https://arxiv.org/abs/1506.04967v1.

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67: 1-48. https://doi.org/10.18637/jss.v067.i01.

Berman, R. (1978). *Modern Hebrew structure*. Tel-Aviv: University Publishing Projects.

Birdsong, D., Gertken, L. M., and Amengual, M. (2012). *Bilingual Language Profile: An easy-to-use instrument to assess bilingualism*. COERLL, University of Texas at Austin. https://sites.la.utexas.edu/bilingual/.

Bodner, G. E., and Masson, M. E. J. (1997). Masked repetition priming of words and nonwords: Evidence for a nonlexical basis for priming. *Journal of Memory and Language* 37: 268–293. https://doi.org/10.1006/jmla.1996.2507.

Borg, A. (1994). Some evolutionary parallels and divergences in Cypriot Arabic and Maltese. *Mediterranean Language Review* 8: 41-67. https://www.jstor.org/stable/10.13173/medilangrevi.8.1994.0041.

Borg, A. J., and Azzopardi-Alexander, M. (1997). *Maltese*. London: Routledge.

Borg, C., Fabri, R., Gatt, A., and Rosner, M. (2012). Korpus Malti: A corpus of contemporary Maltese, v.2.0. http://mlrs.research.um.edu.mt/.

Boudelaa, S., and Marslen-Wilson, W. D. (2001). The time-course of morphological, phonological, and semantic processes in reading Modern Standard Arabic. In *Proceedings of the Twenty Third Annual Meeting of the Cognitive Science Society*, 138-143. Edinburgh, Scotland. https://escholarship.org/uc/item/33n880t7.

Boudelaa, S., and Marslen-Wilson, W. D. (2004). Allomorphic variation in Arabic: Implications for lexical processing and representation. *Brain and Language* 90: 106-116. https://doi.org/10.1016/s0093-934x(03)00424-3.

Bovingdon, R. and Dalli, A. (2006). Statistical analysis of the source origin of Maltese. In Wilson, A., Archer, D., and Rayson, P. (eds.), *Corpus linguistics around the world*, 63-76. New York: Rodopi. https://doi.org/10.1163/9789401202213_006.

Brincat, J. (2011). *Maltese and other languages: A linguistic history of Malta*. Sta Venera, Malta: Midsea Books.

Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bölte, J., and Böhl, A. (2011). The word frequency effect: A review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology* 58: 412-424. https://doi.org/10.1027/1618-3169/a000123.

Brysbaert, M., Mandera, P., and Keuleers, E. (2017). The word frequency effect in word processing: An updated review. *Current Directions in Psychological Science* 27: 45-50. https://doi.org/10.1177/0963721417727521.

Brysbaert, M., and New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods* 41: 977-990. http://doi.org/10.3758/BRM.41.4.977.

Brysbaert, M., Stevens, M., Mandera, P., and Keuleers, E. (2016). The impact of word prevalence on lexical decision times: Evidence from the Dutch Lexicon Project 2. *Journal of Experimental Psychology: Human Perception and Performance* 42: 441-458. https://doi.org/10.1037/xhp0000159.

Burnham, K. P., and Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research* 33: 261-304. http://doi.org/10.1177/0049124104268644.

Camilleri, J. J. (2013). *A computational grammar and lexicon for Maltese* (MSc Thesis). Chalmers University of Technology, Gothenburg, Sweden. https://mlrs.research.um.edu.mt/resources/gabra/.

Carreiras, M., Perea, M., and Grainger, J. (1997). Effects of orthographic neighborhood in visual word recognition: Cross-task comparisons. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 23: 857−871. https://doi.org/10.1037/0278-7393.23.4.857.

Chetail, F., and Boursain, E. (2019). Shared or separated representations for letters with diacritics? *Psychonomic Bulletin & Review* 26: 347−352. https://doi.org/10.3758/s13423-018-1503-0.

Coltheart, M., Davelaar, E., Jonasson, J. T., and Besner, D. (1977). Access to the internal lexicon. In Dornick, S. (ed.), *Attention and performance* VI: 535−556. Hillsdale, NJ: Erlbaum.

Comrie, B. (1991). Towards a history of Arabic Maltese. In Kaye, A. S. (ed.), *Semitic studies in honor of Wolf Leslau on the occasion of his eighty-fifth birthday*, 234-244. Wiesbaden: Otto Harrassowitz.

Comrie, B., and Spagnol, M. (2016). Maltese loanword typology. In Puech, G., and Saade, B. (eds.), *Shifts and patterns in Maltese*, 315-329. De Gruyter. https://doi.org/10.1515/9783110496376-012.

Connine, C. M., Mullennix, J., Shernoff, E., and Yelen, J. (1990). Word familiarity and frequency in visual and auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 16: 1084-1096. https://doi.org/10.1037//0278-7393.16.6.1084.

Costa, A., Caramazza, A., and Sebastian-Galles, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26: 1283-1296. https://doi.org/10.1037/0278-7393.26.5.1283.

Cuetos, F., Rodríguez-Ferreiro, J., Sage, K., and Ellis, A. W. (2012). A fresh look at the predictors of naming accuracy and errors in Alzheimer's disease. *Journal of Neurophysiology* 6: 242-256. https://doi.org/10.1111/j.1748-6653.2011.02025.x.

Deutsch, A., Frost, R., and Forster, K. I. (1998). Verbs and nouns are organized and accessed differently in the mental lexicon: Evidence from Hebrew. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 24: 1238-1255. https://doi.org/10.1037/0278-7393.24.5.1238.

Diependaele, K., Brysbaert, M., and Neri, P. (2012). How noisy is lexical decision? *Frontiers in Psychology* 3: 348. http://doi.org/10.3389/fpsyg.2012.00348.

Dijkstra, T., Grainger, J., and van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language* 41: 496-518. https://doi.org/10.1006/jmla.1999.2654.

Domínguez, A., and Cuetos, F. (2018). The contrastive value of lexical stress in visual word recognition: Evidence from Spanish. *Psicothema* 30: 276−282. https://doi.org/10.7334/psicothema2018.30.

Duñabeitia, J. A., and Carreiras, M. (2011). The relative position priming effect depends on whether letters are vowels or consonants. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 17: 1143-1163. https://doi.org/10.1037/a0023577.

Ernestus, M., and Cutler, A. (2015). BALDEY: A database of auditory lexical decisions. *The Quarterly Journal of Experimental Psychology* 68: 1469-1488. https://doi.org/10.1080/17470218.2014.984730.

Feldman, L. B., and Bentin, S. (1994). Morphological analysis of disrupted morphemes: Evidence from Hebrew. *The Quarterly Journal of Experimental Psychology* 47A: 407-435. https://doi.org/10.1080/14640749408401118.

Ferrand, L., Brysbaert, M., Keuleers, E., New, B., Bonin, P., Méot, A., Augustinova, A., and Pallier, C. (2011). Comparing word processing times in naming, lexical decision, and progressive demasking: Evidence from Chronolex. *Frontiers in Psychology* 2: 306. https://doi.org/10.3389/fpsyg.2011.00306.

Ferrand, L., Méot, A., Spinelli, E., New, B., Pallier, C., Bonin, P., Dufau, S., Mathôt, S., and Grainger, J. (2018). MEGALEX: A megastudy of visual and auditory word recognition. *Behavior Research Methods* 50: 1285-1307. https://doi.org/10.3758/s13428-017-0943-1.

Ferrand, L., New, B., Brysbaert, M., Keuleers, E., Bonin, P., Méot, A., Augustinova, M., and Pallier, C. (2010). The French lexicon project: Lexical decision data for 38,840 French words and 38,840 pseudowords. *Behavior Research Methods* 42: 488-496. https://doi.org/10.3758/brm.42.2.488.

Forster, K. I. (1976). Accessing the mental lexicon. In Wales, R. J., and Walker, E. (eds.), *New Approaches to Language Mechanisms*, 257-287. Amsterdam: North-Holland.

Forster, K. I. (1992). Lexical effects in masked form-priming. Paper presented at the 33rd Annual Meeting of the Psychonomics Society, St. Louis, MO.

Forster, K. I. (1998). The pros and cons of masked priming. *Journal of Psycholinguistic Research* 27: 203−233. https://doi.org/10.1023/a:1023202116609.

Forster, K. I., and Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior* 12: 627-635. https://doi.org/10.1016/S0022-5371(73)80042-8.

Forster, K. I., and Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 10: 680−698. https://doi.org/10.1037/0278-7393.10.4.680.

Forster, K. I., and Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers* 35: 116-124. https://doi.org/10.3758/BF03195503.

Forster, K. I., Mohan, K., and Hector, J. (2003). The mechanics of masked priming. In Kinoshita, S., and Lupker, S. J. (eds.), *Masked priming: The state of the art*, p.3−37. https://doi.org/10.4324/9780203502846.

Forster, K. I., and Shen, D. (1996). No enemies in the neighborhood: Absence of inhibitory effects in lexical decision and categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22: 696−713. https://doi.org/10.1037//0278-7393.22.3.696.

Frost, R., Deutsch, A., and Forster, K. I. (2000). Decomposing morphologically complex words in a nonlinear morphology. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26: 751-765. https://doi.org/10.1037//0278-7393.26.3.751.

Frost, R., Forster, K. I., and Deutsch, A. (1997). What can we learn from the morphology of Hebrew? A masked-priming investigation of morphological representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 23: 829-856. https://doi.org/10.1037/0278-7393.23.4.829.

Frost, R., Kugler, T., Deutsch, A., and Forster, K. I. (2005). Orthographic structure versus morphological structure: Principles of lexical organization in a given language. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 31: 1293-1326. https://doi.org/10.1037/0278-7393.31.6.1293.

Gardani, F. (2008). *Borrowing of inflectional morphemes in language contact*. Peter Lang.

Gatt, A., and Čéplö, S. (2013). Digital corpora and other electronic resources for Maltese. In *Proceedings of the international conference on corpus linguistics*. University of Lancaster.

Gatt, A., and Fabri, R. (2018). Borrowed affixes and morphological productivity: A case study of two Maltese nominalisations. In Paggio, P., and Gatt, A. (eds.), *The languages of Malta*, 143-169. Berlin: Language Science Press. http://dx.doi.org/10.5281/zenodo.1181793.

Geary, J. A., and Ussishkin, A. (2018). Root-letter priming in Maltese visual word recognition. *The Mental Lexicon* 13: 1-25. https://doi.org/10.1075/ml.18001.gea.

Geary, J. A., and Ussishkin, A. (2019). Morphological priming without semantic relationship in Hebrew spoken word recognition. *Proceedings of the Linguistics Society of America* 4. https://doi.org/10.3765/plsa.v4i1.4509.

González-Nosti, M., Barbón, A., Rodríguez-Ferreiro, J., and Cuetos, F. (2014). Effects of the psycholinguistic variables on the lexical decision task in Spanish: A study with 2,765 words. *Behavior Research Methods* 46: 517-525. https://doi.org/10.3758/s13428-013-0383-5.

Grainger, J., Granier, J.-P., Farioli, F., Van Assche, E., and van Heuven, W. J. B. (2006). Letter position information and printed word perception: The relative-position priming constraint. *Journal of Experimental Psychology: Human Perception and Performance* 32: 865−884. https://doi.org/10.1037/0096-1523.32.4.865.

Hendrix, P., and Sun, C. C. (2021). A word or two about nonwords: Frequency, semantic neighborhood density, and orthography-to-semantics consistency effects for nonwords in the lexical decision task. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 47: 157−183. http://dx.doi.org/10.1037/xlm0000819.

Hoberman, R. D. (2007). Maltese morphology. In Kaye, A. S. (ed.), *Morphologies of Asia and Africa, volume 1*, 257-281. Winona Lake, Indiana: Eisenbrauns.

Hoberman, R. D., and Aronoff, M. (2003). The verbal morphology of Maltese: From Semitic to Romance. In Shimron, J. (ed.), *Language processing and acquisition in languages of Semitic, root-based, morphology*, 61-78. Amsterdam: John Benjamins.

Jared, D., and Kroll, J. F. (2001). Do bilinguals activate phonological representations in one or both of their languages when naming words? *Journal of Memory and Language* 44: 2-31. https://doi.org/10.1006/jmla.2000.2747.

Jones, M. N., Johns, B. T., and Recchia, G. (2012). The role of semantic diversity in lexical organization. *Canadian Journal of Experimental Psychology* 66: 115-124. https://doi.org/10.1037/a0026727.

Keuleers, E., and Balota, D. A. (2015). Megastudies, crowdsourcing, and large datasets in psycholinguistics: An overview of recent developments. *The Quarterly Journal of Experimental Psychology* 68: 1457-1468. http://doi.org/10.1080/17470218.2015.1051065.

Keuleers, E., Brysbaert, M., and New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods* 42: 643-650. https://doi.org/10.3758/BRM.42.3.643.

Keuleers, E., Diependaele, K., and Brysbaert, M. (2010). Practice effects in large-scale visual word recognition studies: A lexical decision study on 14,000 Dutch mono- and disyllabic words and nonwords. *Frontiers in Psychology* 1: http://doi.org/10.3389/fpsyg.2010.00174.

Keuleers, E., Lacey, P., Rastle, K., and Brysbaert, M. (2012). The British lexicon project: Lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behavior Research Methods* 44: 287-304. https://doi.org/10.3758/s13428-011-0118-4.

Kinoshita, S., Robidoux, S., Mills, L., and Norris, D. (2013). Visual similarity effects on masked priming. *Memory & Cognition* 42: 821−833. https://doi.org/10.3758/s13421-013-0388-4.

Kliegl, R., Masson, M. E. J., and Richter, E. M. (2010). A linear mixed model analysis of masked repetition priming. *Visual Cognition* 18: 655−681. https://doi.org/10.1080/13506280902986058.

Kuperman, V. (2015). Virtual experiments in megastudies: A case study of language and emotion. *The Quarterly Journal of Experimental Psychology* 68: 1693-1710. http://dx.doi.org/10.1080/17470218.2014.989865.

Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software* 82: 1-26. https://doi.org/10.18637/jss.v082.i13.

Lim, S. W. H. (2016). The influence of orthographic neighborhood density and word frequency on visual word recognition: Insights from RT distributional analyses. *Frontiers in Psychology* 7: 401. https://doi.org/10.3389/fpsyg.2016.00401.

Mandera, P., Keuleers, E., and Brysbaert, M. (2020). Recognition times for 62 thousand English words: Data from the English Crowdsourcing Project. *Behavior Research Methods* 52: 741-760. https://doi.org/10.3758/s13428-019-01272-8.

Marcet, A., Ghukasyan. H., Fernández-López, M., and Perea, M. (2020). Jalapeno or jalapeño: Do diacritics in consonant letters modulate visual similarity effects during word recognition? *Applied Psycholinguistics* 41: 579−593. https://doi.org/10.1017/S0142716420000090.

Marcet, A., and Perea, M. (2017). Is nevtral NEUTRAL? Visual similarity effects in the early phases of written-word recognition. *Psychonomic Bulletin and Review* 24: 1180−1185. https://doi.org/10.3758/s13423-016-1180-9.

Marcet, A., and Perea, M. (2018). Can I order a burger at rnacdonalds.com? Visual similarity effects of multi-letter combinations at the early stages of word recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition* 44: 699−706. https://doi.org/10.1037/xlm0000477.

Masson, M. E. J., and Isaak, M. I. (1999). Masked priming of words and nonwords in a naming task: Further evidence for a nonlexical basis for priming. *Memory & Cognition* 27: 399−412. https://doi.org/10.3758/BF03211536.

Mayer, T., Spagnol, M., and Schönhuber, F. (2013). Fixing the Maltese broken plural. In Borg, A., Caruana, S., and Vella, A. (eds.), *Perspectives on Maltese Linguistics*, 129-158. Berlin: Akademie Verlag. https://doi.org/10.1524/9783050065144.129.

McCarthy, J. J. (1979). *Formal problems in Semitic phonology and morphology* (PhD Thesis). Massachusetts Institute of Technology, Cambridge, MA.

Mifsud, J., and Mitterer, M. (2020). Language variation influences the sensitivity to visual detail in word recognition. Poster presented at Architectures and Mechanisms for Language Processing 2020 (AMLaP 2020). https://amlap2020.github.io/a/76.pdf.

Mifsud, M. (1995a). *Loan verbs in Maltese: A descriptive and comparative study*. Brill.

Mifsud, M. (1995b). The productivity of Arabic in Maltese. *Proceedings of the 2nd International Conference of AIDA*, 151-160. Cambridge University Press.

New, B., Araújo, V., and Nazzi, T. (2008). Differential processing of consonants and vowels in lexical access through reading. *Psychological Science* 19: 1223–1227. https://doi.org/10.1111/j.1467-9280.2008.02228.x.

Perea, M., Abu Mallouh, R., Mohammed, A., Khalifa, B., and Carreiras, M. (2016). Do diacritical marks play a role at the early stages of word recognition in Arabic? *Frontiers in Psychology* 7: 1255. https://doi.org/10.3389/fpsyg.2016.01255.

Perea, M., Abu Mallouh, R., Mohammed, A., Khalifa, B., and Carreiras, M. (2018). Does visual letter similarity modulate masked form priming in young readers of Arabic? *Journal of Experimental Child Psychology* 169: 110–117. https://doi.org/10.1016/j.jecp.2017.12.004.

Perea, M., Gatt, A., Moret-Tatay, C., and Fabri, R. (2012). Are all Semitic languages immune to letter transpositions? The case of Maltese. *Psychonomic Bulletin & Review* 19: 942-947. https://doi.org/10.3758/s13423-012-0273-3.

Peressotti, F., and Grainger, J. (1999). The role of letter identity and letter position in orthographic priming. *Perception & Psychophysics* 61: 691–706. https://doi.org/10.3758/bf03205539.

Perfetti, C. A. (1992). The representation problem in reading acquisition. In Gough, P. B., Ehri, L. C., and Treiman, R. (eds.), *Reading acquisition*, 145–174. Hillsdale, NJ: Erlbaum.

Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading* 11: 357−383. https://doi.org/10.1080/10888430701530730.

Poort, E. D., and Rodd, J. M. (2017). The cognate facilitation effect in bilingual lexical decision is influenced by stimulus list composition. *Acta Psychologica* 180: 52-63. https://doi.org/10.1016/j.actpsy.2017.08.008.

Preston, K. A. (1935). The speed of word perception and its relation to reading ability. *Journal of General Psychology* 13: 199-203. https://doi.org/10.1080/00221309.1935.9917878.

R Core Team. (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Schluter, K. (2013). *Hearing words without structure: Subliminal speech priming and the organization of the Moroccan Arabic lexicon*. PhD Dissertation. University of Arizona, Tucson, AZ. http://hdl.handle.net/10150/301752.

Schröter, P., and Schroeder, S. (2017). The Developmental Lexicon Project: A behavioral database to investigate visual word recognition across the lifespan. *Behavior Research Methods* 49: 2183-2203. https://doi.org/10.3758/s13428-016-0851-9.

Sears, C. R., Hino., Y., and Lupker, S. J. (1995). Neighborhood size and neighborhood frequency effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance* 21: 876−900. https://doi.org/10.1037/0096-1523.21.4.876.

Soares, A. P., Lages, A., Silva, A., Comesaña, M., Sousa, I., Pinheiro, A. P., and Perea, M. (2019). Psycholinguistic variables in visual word recognition and pronunciation of European Portuguese words: A mega-study approach. *Language, Cognition and Neuroscience* 34: 689-719. https://doi.org/10.1080/23273798.2019.1578395.

Soares, A. P., Perea, M., and Comesaña, M. (2014). Tracking the emergence of the consonant bias in visual-word recognition: Evidence with developing readers. *PLoS ONE* 9: e88580. https://doi.org/10.1371/journal.pone.0088580.

Taft, M. (1979). Recognition of affixed words and the word frequency effect. *Memory & Cognition* 74: 263-272. https://doi.org/10.3758/BF03197599.

Taft, M. (2004). Morphological decomposition and the reverse base frequency effect. *The Quarterly Journal of Experimental Psychology* 57A: 745-765. https://doi.org/10.1080/02724980343000477.

Taft, M., and Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior* 14: 638-647. https://doi.org/10.1016/S0022-5371(75)80051-X.

Taft, M., and Hambly, G. (1986). Exploring the Cohort Model of spoken word recognition. *Cognition* 22: 259-282. https://doi.org/10.1016/0010-0277(86)90017-X.

Tsang, Y.-K., Huang, J., Lui, M., Xue, M., Chan, Y.-W. F., Wang, S., and Chen, H.-C. (2018). MELD-SCH: A megastudy of lexical decision in simplified Chinese. *Behavior Research Methods* 50: 1763-1777. https://doi.org/10.3758/s13428-017-0944-0.

Tse, C.-S., Yap, M. J., Chan, Y.-L., Sze, W. P., Shaoul, C., and Lin, D. (2017). The Chinese Lexicon Project: A megastudy of lexical decision performance for 25,000+ traditional Chinese two-character compound words. *Behavior Research Methods* 49: 1503-1519. https://doi.org/10.3758/s13428-016-0810-5.

Tucker, B. V., Brenner, D., Danielson, D. K., Kelley, M. C., Nenadić, F., and Sims, M. (2018). The Massive Auditory Lexical Decision (MALD) database. *Behavior Research Methods*. https://doi.org/10.3758/s13428-018-1056-1.

Twist, A. E. (2006). *A psycholinguistic investigation of the verbal morphology of Maltese* (PhD Thesis). University of Arizona, Tucson, AZ. http://hdl.handle.net/10150/194996.

Ussishkin, A., Dawson, C. R., Wedel, A., and Schluter, K. (2015). Auditory masked priming in Maltese spoken word recognition. *Language, Cognition, and Neuroscience* 30: 1096–1115. https://doi.org/10.1080/23273798.2015.1005635.

Ussishkin, A., Velan, H., Spagnol, M., Wedel, A., and Wray, S. (2016). Maltese root priming is morphological, not phonological. Poster presented at the 22nd AMLaP (Architecture and Mechanisms of Language Processing) Conference, Basque Center for Brain and Language, Bilbao, Spain, September 2, 2016.

Velan, H., and Frost, R. (2007). Cambridge University versus Hebrew University: The impact of letter transposition on reading English and Hebrew. *Psychonomic Bulletin & Review* 14: 913-918. https://doi.org/10.3758/BF03194121.

Velan, H., and Frost, R. (2009). Letter-transposition effects are not universal: The impact of transposing letters in Hebrew. *Journal of Memory and Language* 61: 285-302. http://doi.org/10.1016/j.jml.2009.05.003.

Velan, H., and Frost, R. (2011). Words with and without internal structure: What determines the nature of orthographic and morphological processing? *Cognition* 118: 141-156. https://doi.org/10.1016/j.cognition.2010.11.013.

Witzel, J., Cornelius, S., Witzel, N., Forster, K. I., and Forster, J. C. (2013). Testing the viability of webDMDX for masked priming experiments. *The Mental Lexicon* 8: 372-400. https://doi.org/10.1075/ml.8.3.07wit.

Wurm. L. H., and Fisicaro, S. A. (2014). What residualizing predictors in regression analyses does (and what it does not do). *Journal of Memory and Language* 72: 37–48. https://doi.org/10.1016/j.jml.2013.12.003.

Wray, S. C. (2016). Decomposability and the effects of morpheme frequency in lexical access (PhD Thesis). University of Arizona, Tucson, AZ. http://hdl.handle.net/10150/621751.

Yap, M. J., Rickard Liow, S. J., Jalil, S. B., and Faizal, S. S. B. (2010). The Malay Lexicon Project: A database of lexical statistics for 9,592 words. *Behavior Research Methods* 42: 992-1003. https://doi.org/10.3758/BRM.42.4.992.

Yap, M. J., Sibley, D. E., Balota, D. A., Ratcliff, R., and Rueckl, J. (2015). Responding to nonwords in the lexical decision task: Insights from the English Lexicon Project. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 41: 597–613. https://doi.org/10.1037/xlm0000064.