Maltese First Language Acquisition of Morphosyntax:

A Focus on Nominal Inflection and Agreement

Michela Vella

A thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

To The Institute of Linguistics & Language Technology

University of Malta

2025

To all the children who participated in this study. May they continue to discover joy in life's smallest moments, and may their infectious curiosity ripple outward, brightening the world wherever they go.

Declaration of authenticity

I declare that this thesis is my own original work and that it has not been submitted, in whole or in part, for any other degree or qualification at this or any other institution. All sources, data and quotations drawn from the work of others have been appropriately cited and acknowledged.

[redacted]

Michela Vella

Acknowledgements

I would first like to express my sincere gratitude to all the children who participated in this study and to their families, who generously welcomed me into their homes on multiple occasions to document their children's language development. From accommodating sessions during the COVID-19 restrictions to finding time in their busy schedules, their kindness and wholehearted participation made this research possible.

My deepest thanks go to my main supervisor, Professor Ray Fabri, for his wisdom, calm guidance, and invaluable feedback throughout this journey. This work would not have been possible without his insightful observations and steady encouragement. I am equally grateful to my co-supervisor, Professor Alexandra Vella, for her thoughtful input, time, and stimulating discussions.

I am also indebted to colleagues and friends at the Institute of Linguistics and Language Technology, led by Dr Sarah Grech, and at the Department of Maltese, headed by Professor Adrian Grima, as well as to the staff within the Department of Early Childhood and Primary Education, the Department of Artificial Intelligence, and the Malta Centre. It has been a privilege to work within such supportive teams. Special thanks are due to Dr Jessica Nieder for her efficiency and guidance with the coding and statistical analyses, and to Dr Marc Tanti for his continued support. I extend my appreciation to Professor Michael Spagnol for always offering food for thought, and to Professor Virve Vihman and fellow PhD student Adele Vaks at the University of Tartu for making my time there both productive and memorable.

To my parents and sister, thank you for your unwavering love, patience, and encouragement, and for supporting me in every possible way. To my 3.8-legged furry friend, thank you for your quiet companionship throughout this journey. I am also grateful to my extended family for their continual care. Special remembrance goes to my late grandfather, who, as a retired professor, allowed me to tag along on his visits to the university, where I would explore the Japanese Garden and spend a few extra minutes gazing at the fish in the administrative building aquarium.

To my dear friends – Julia and Gail, the best business partners; Kirsty, Amanda, and Michael, my fellow postgraduates; and Miriana and Sam, two special sidekicks – thank you for the countless memories and motivation. Finally, I am profoundly grateful to Raffaello, my life companion, for his unfailing support, for every discussion on Maltese child language, and for loving me just the way I am (a *Bridget Jones* allusion was inevitable). To more love, laughter, and linguistics.

Above all, I thank God for the people who have accompanied me on this journey and for bringing me this far.

The research work disclosed in this thesis is funded by the Tertiary Education Scholarships Scheme (TESS).

Abstract

This study traces the development of morphosyntax in typically developing, Maltese-dominant bilingual children aged 18–60 months, examining how mean length of utterance in morphemes (MLUm), parts-of-speech (POS) category distribution, and nominal morphosyntax evolve and interact over time. To this end, a corpus comprising over 30 hours of Maltese child language data was collected. Drawing on longitudinal naturalistic data complemented by targeted elicitation tasks, and supported by statistical modelling, the study provides a detailed account of language acquisition in a bilingual environment. The results show that MLUm increases steadily with age at a rate comparable to previous studies. A higher proportion of nouns in early productions gradually gives way to increased use of verbs and function words, with mixed-effects modelling confirming a significant shift over time and in relation to MLUm. Analyses of nominal inflection reveal that concatenative plural forms are acquired earlier and with greater accuracy than templatic forms. In nominal agreement, accuracy increases with age, although earlier stabilisation of gender relative to number is not statistically confirmed. Integrative analyses indicate that the stabilisation of nominal morphosyntax coincides with the transition to multiword syntax and the attenuation of this early noun dominance. Collectively, these findings depict the development of nominal systems as gradual and interdependent, contributing new empirical evidence to cross-linguistic theories of morphosyntactic learnability and bilingual acquisition.

Keywords: Maltese first language acquisition, grammatical development, morphosyntax, bilingualism

Table of contents

List of tables

List of figures

Abbreviations

BFLA	bilingual first language acquisition
BP	broken plural
CDI	Communicative Development Inventory/ies
CDS	child-directed speech
CHAT	Codes for the Human Analysis of Transcripts
CHILDES	Child Language Data Exchange System
CI	confidence interval
CLA	child language acquisition
CLAN	Computerized Language ANalysis
CS	code-switching
CxG	Construction Grammar
ESLA	early second language acquisition
FSP	feminine sound plural
ID	individual difference
L1	first language
L2	second language
L3	third language
MAIN	Multilingual Assessment Instrument for Narratives
MD	Maltese-dominant
MFLA	Maltese first language acquisition
MLU	mean length of utterance
MLUm	mean length of utterance in morphemes
MLUw	mean length of utterances in words
MPU	morpheme per utterance
NLP	natural language processing
NP	noun phrase
OM	otitis media
P&P	Principles and Parameters
POS	parts-of-speech
POSA	Poverty of the Stimulus Argument

RQ	research question
SES	socioeconomic status
SP	sound plural
TD	typically developing
UG	Universal Grammar

Morphemic glossing

1, 2, 3	first, second, third person
COLL	collective
CMP	comparative
DEF	definite article
OBJ	object
F	feminine
IPFV	imperfective
M	masculine
NEG	negative
PL	plural
POSS	possessive
PFV	perfective

1. Introduction

Few questions in linguistics have inspired more enduring fascination than how children, under typical circumstances, acquire and master language with such extraordinary ease within the first few years of life. As Bloomfield (1933, p. 29) observes, child language acquisition constitutes "the greatest intellectual feat any of us is ever required to perform". While this seemingly effortless process has been studied extensively across numerous languages and within a range of theoretical frameworks, research on Maltese first language acquisition (MFLA) remains limited.

This thesis investigates early Maltese morphosyntactic development, focusing on nominal morphosyntax, with the dual aim of documenting an under-studied language in the field of child language acquisition (CLA) and contributing to broader models of early grammatical development. To this end, a two-year longitudinal design was employed, tracking the acquisition pathway of eight typically developing (TD), Maltese-dominant (MD) toddlers aged between 18–36 months at recruitment. Each child was followed up at four- to six-month intervals, yielding data spanning 18–60 months, thereby approximating a lengthier longitudinal trajectory.

1.1. The study in context

This section introduces the study by outlining the scope and focus of the thesis. It first elaborates on the rationale for focusing on MFLA (Section 1.1.1), before discussing the acquisition of morphosyntax more broadly (Section 1.1.2), and then elaborating on nominal inflection and agreement (Section 1.1.3).

1.1.1. Maltese first language acquisition

First language (L1) acquisition refers to the process through which children acquire their native language(s) naturally and implicitly, coming to command the phonological, lexical, morphological, syntactic and pragmatic structures used by speakers in their linguistic environment. Acquisition can be distinguished from language learning which, by contrast, is generally a conscious and explicit process that involves deliberate study and heightened metalinguistic knowledge (e.g., Krashen, 1982). While acquisition is generally associated with one's L1, the terminology sometimes overlaps in discussions in the context of second and third language (L2, L3) learning.

The acquisition of an L1(s) typically unfolds within the first five years of life, with the most rapid development occurring between 12–60 months (e.g., Clark, 2024), while refinement of linguistic knowledge continues into early adolescence. The present study encompasses what are generally considered the principal stages of early development, from 18–60 months.

Mastering the inflectional system of one's L1(s) is a major developmental achievement and understanding how this process unfolds remains one of the central challenges in language acquisition research. Despite cross-linguistic variation in inflectional complexity (e.g., Stoll, 2015) and sociocultural context (e.g., Kelly et al., 2015), children acquiring different languages display strikingly similar developmental pathways (e.g., Thornton, 2016). To explain this apparent uniformity amid diversity, child language research increasingly relies on data drawn from a typologically wide range of languages that reflect the full extent of linguistic and cultural variation.

In recent years, cross-linguistic studies have expanded beyond English and other well-documented Indo-European languages (e.g., Stoll & Bickel, 2013), highlighting the importance of studying underrepresented languages. As Kidd and Garcia (2022, p. 704) emphasise, "a comprehensive theory of any language phenomenon requires a representative sample that provides a solid foundation for theory building, and an even bigger sample for theory testing." Similarly, Bowerman (2011, p. 616) warns that, without adequate language coverage of the 7,000 or so languages of the world, researchers risk constructing theories that are based on a narrow and unrepresentative sample, thus failing to "do justice to language diversity." Collecting and analysing data from a wide range of languages is therefore essential to build theories that accurately reflect the full spectrum of CLA rather than extrapolating universal principles from a narrow empirical base.

Maltese offers a particularly valuable opportunity to contribute to this endeavour from a typological and sociolinguistic perspective. As a Semitic language that has undergone extensive contact with Romance and, more recently, English, Maltese exhibits a hybrid system combining root-based and stem-based morphology (e.g., Fabri, 2009; Spagnol, 2011). In addition, from a sociolinguistic standpoint, children acquiring Maltese are almost invariably exposed to both Maltese and English from an early age in an environment characterised by widespread bilingualism and code-switching (CS; e.g., Camilleri Grima, 2016). This unique linguistic configuration makes Maltese an ideal testing ground for exploring how children navigate morphosyntactic complexity, and how structural and sociolinguistic factors interact in the acquisition process.

Despite its potential, research on MFLA remains limited. Existing studies offer crucial insights into specific aspects of acquisition, such as lexical development, phonology, and CS (e.g., Azzopardi-Alexander, 2016; Gatt, 2017; McLeod & Crowe, 2018). Nonetheless, the developmental trajectory of Maltese nominal morphosyntax is still largely undocumented. There is limited empirical evidence on how children acquiring Maltese progress from early word combinations to complex nominal structures, or on how core morphosyntactic features emerge and stabilise. Bridging this research gap is essential for documenting MFLA, enabling its inclusion in cross-linguistic studies (e.g., Katsos et al., 2016; Sauerland et al., 2016), and for testing broader theories of language acquisition.

The present work seeks to address this lacuna by providing an account of early nominal morphosyntax in Maltese, integrating the analysis of spontaneous and elicited data to chart developmental patterns. Beyond documentation, the study contributes to theoretical debates by examining the applicability of established indices, such as mean length of utterance in morphemes (MLUm), to a bilingual and morphologically complex setting. It also traces the developmental course of nominal inflection and agreement in tandem with broader structural growth. In so doing, the study responds to calls for expanding empirical coverage and contributes data on MFLA that has thus far been largely absent from the scene of acquisition research. The findings may also inform the refinement of theoretical models of language acquisition which, as Ambridge and Wagner (2021) emphasise, must continually evolve in response to emerging empirical evidence.

Alongside its theoretical relevance, the research carries applied value. The lack of normative developmental data for Maltese results in speech and language pathologists often resorting to English-based milestones when assessing Maltese-speaking children, an approach that can be misleading given the typological differences between the two languages (Gatt, 2010). As Gatt (2010, pp. 360–361) points out, "the structural characteristics of Maltese, together with the varying degrees of language contact in young children's communicative environments, underscore the need for developmental norms" for MFLA. The longitudinal data collected for this study can therefore serve as a reference point for clinicians identifying and managing language impairment in Maltese-speaking children and for speech and language pathologists working on the rehabilitation of linguistic skills in Maltese-speaking adults, such as those recovering from aphasia.

Furthermore, insights from this study could inform early years pedagogy for both L1 and L2 learners of Maltese, offering evidence-based developmental milestones for curriculum design and language resource creation. Establishing developmental benchmarks for MFLA could also assist parents, educators, and resource developers in fostering language-rich environments for young children.

In sum, studying MFLA enriches our understanding of linguistic diversity while providing tangible benefits for education and clinical practice. Documenting how MD children come to command the complex morphosyntactic system of their language contributes both to local needs and to the global goal of achieving a genuinely cross-linguistic gauge of language acquisition.

1.1.2. Acquisition of morphosyntax

Morphosyntax encompasses two closely interdependent domains: morphology, concerned with the study of word structure and the relationship between form and meaning (e.g., Bauer, 2019), and syntax, which addresses the rules and principles that govern phrase and sentence structure, such as word order and agreement (e.g., Chomsky, 1957). The interdependence of these domains is accentuated by Svenonius (2019, p. 4), who notes that “the traditional understanding of syntactic features starts from a consideration of morphosyntactic features, features which have morphological expression and syntactically influenced distribution”. Similarly, Haspelmath (2020, p. 126) argues that “there is no clear way of delimiting morphology from syntax”. These observations shed light on the intricate interdependence of morphology and syntax, suggesting that any attempt to establish a definitive boundary between them remains conceptually and empirically contested.

Understanding how children acquire morphosyntactic knowledge offers insights into the nature of the human language faculty, the interaction between innate capacities and environmental input, and the timeline of linguistic development. Morphosyntax therefore provides a central vantage point for examining how children begin to organise linguistic structure and express grammatical relations.

1.1.3. Nominal inflection and agreement

This thesis comprises two main strands of analysis. The first examines indices of early structural growth through measures of MLUm and distribution of parts-of-speech (POS) across age groups. These complementary measures provide an overview of children's developing grammatical capacity: MLUm serves as an indicator of morphosyntactic complexity, while the evolving balance of word classes (namely nouns and verbs) in the lexicon reflects the lexical underpinnings of syntactic expansion. The second focuses on nominal inflection and agreement, specifically the development of gender and number marking. Maltese nominal morphology, drawing on both Semitic and Romance sources, presents children with a system involving multiple paradigms of inflection and patterns of agreement within the noun phrase (NP). Investigating how these features emerge and stabilise offers a view of how morphological and syntactic systems co-develop.

In addition, research indicates that nominal morphosyntax, particularly gender and number agreement, often emerges relatively early in CLA (e.g., Demuth & Weschler, 2012; Mariscal, 2009). This tendency is especially evident in Semitic languages, where pervasive morphological marking on nouns provides rich input for early learning, while verbal inflection systems require more time to develop fully (e.g., Abdalla & Crago, 2008; Berman, 2003; Slobin, 1985). The present study examines the language of children between 18–60 months, a period when nominal morphology is thought to be both frequent in the input and among the first morphosyntactic systems to stabilise (e.g., Brown, 1973; Ravid, 2019). Focusing on nominal inflection and agreement therefore provides an empirically accessible and theoretically informative window onto early grammatical development. In this sense, nouns offer an ideal empirical locus for examining how children begin to internalise and generalise patterns that underpin later morphosyntactic complexity.

By integrating overall developmental indices (MLUm and POS distribution) with analyses of specific morphosyntactic domains (nominal inflection and agreement), the study aims to capture both the breadth and depth of early nominal development. The combination of naturalistic and experimental data facilitates a nuanced understanding of the pathways through which TD children acquiring Maltese as their L1 begin to produce and generalise grammatical structures. This investigation should serve as a springboard for further research in the field, stimulating continued theoretical and applied inquiry.

1.2. Overview of the chapters

The thesis is organised as follows. Chapter 2 reviews key theoretical perspectives in the field of CLA, situating the present study within ongoing debates between nativist and constructivist approaches, the role of input and sociocultural factors, and the developmental patterns observed across languages. The chapter concludes by outlining the linguistic and sociolinguistic context of Maltese and identifying the research gap and questions that motivate the present investigation. Chapter 3 details the methodological framework, describing participant recruitment, study design, data collection, and transcription, annotation and analysis procedures. Chapters 4 and 5 present the two main empirical investigations: Chapter 4 analyses MLUm and POS category distribution, while Chapter 5 examines the acquisition of nominal inflection and agreement, and developmental timing. Finally, Chapter 6 summarises the main findings, discusses their theoretical and methodological implications, and outlines avenues for future research, aiming to provide a more comprehensive snapshot of the various facets of MFLA.

2. Child language acquisition

2.1. Theories and trends

The aim of this section is to present a general overview of the field of CLA and to shed light on recent developments and current trends. I start by tracing the origins of research on CLA (Section 2.1.1) and then outline long-standing debates and approaches which underlie the field at present (Sections 2.1.2, 2.1.3).

2.1.1. From early theories to current trends

CLA is a well-established discipline in linguistics with a long history of scholarship. According to Levelt (2013, p. vi), the "empirical roots" of psycholinguistics, and indeed, of CLA, lie in the late 1700s. The diary study published in 1787 by the German philosopher Tiedemann is regarded as the first attempt to record child language. It contains observations on the behaviour and development of Tiedemann's own son (Campell & Wales, 1970; Murchison & Langer, 1927). "Its main value is that it is a primer, the first psychological record of its sort since Héroard", who also kept a diary about a child's development, yet whose focus was not language acquisition (Levelt, 2013, p. 95).

Following these studies, numerous works were published in the 19th century and the first half of the 20th century, most of which are diary entries on the researchers' own children (e.g., Darwin, 1877; Preyer, 1882; Ronjat, 1913; Stern & Stern, 1928; Taine, 1876). In the 1800s and early 1900s, child language research mainly centred on "philosophical inquiries into human nature" (Berko Gleason &

Ratner, 2016, p. 18). In this period, while some studies (e.g., Stern & Stern, 1928) made observations which are at the core of modern-day psycholinguistics, such as memory development observations, work published in the field barely touched upon major issues in child language, with "little analysis on structure and no analysis on conversation skills" (Clark, 2024, p. 16).

In the mid-1950s, research on CLA underwent a radical transformation as part of the "cognitive revolution" (Miller, 2003). Noam Chomsky was among the leading scholars driving this development (Allot et al., 2021). This turning point fostered interdisciplinary work across the social sciences and reoriented research toward the processes underlying children's language development. As Berko Gleason and Ratner (2016, p. 19) describe, it was "an explosion of research," largely centred on the acquisition of syntax. The 1960s and 70s were a period of "tremendous activity in the field of language acquisition" (Goodluck, 2020, p. 3), marked in particular by Roger Brown's seminal studies. Brown and his students collected a wealth of data and offered theoretical contributions which resonate and impact research in the field to this day, as do the names Adam, Eve, and Sarah, the three subjects of Brown's longitudinal study, chronicled in his 1973 influential book *A First Language: The Early Stages*.[1]

Another pioneering 1973 publication is Bowerman's cross-linguistic paper on child language, which paved the way for cross-linguistic research in the field, an increasingly popular area and "crucial component of studies in language development" (Goodluck, 2020, p. 3). With the advent of computer science in the

[1] See Section 2.3.2 for an account about Brown's (1973) research.

1960s, studies leading up to the early 2000s were characterised by advanced scientific developments and rapidly increasing language acquisition models (e.g., Friederici, 2002; Saffran et al., 1996). This led to the emergence of new experimental techniques, such as preferential looking paradigms (e.g., Golinkoff et al., 1987), and to the introduction of corpus studies in CLA research (e.g., MacWhinney, 2000).

Cross-linguistic studies (e.g., Bergelson et al., 2023; McLeod & Crowe, 2018), experimental techniques based on cutting-edge technology (e.g., Ambridge & Rowland, 2013; Perkins & Zhang, 2024) and corpus research (e.g., Nag et al., 2024; Parisse, 2019) are all currently trending areas in CLA research. At present, there is also a focus on a) bilingual and multilingual CLA (e.g., Yip & Matthews, 2022), b) atypical child language development (e.g., Torrens, 2020), c) understudied languages in the field (e.g., Hellwig et al., 2023), and d) the impact of digital technology on language development and cognitive skills (e.g., Flewitt et al., 2024), with research in all areas contributing towards a more comprehensive understanding of the way in which children come to master language within the first few years of life.

In the field of CLA, "some of the current issues are still those that dominated debates about language acquisition after the publication of Chomsky's *Aspects of the theory of syntax* in 1965" (Clark, 2024, p. 16). The perennial debates on nature versus nurture (Section 2.1.2) and generativist versus constructivist approaches (Section 2.1.3) are still at the heart of language acquisition controversies. The aim of the following sections is to outline these debates. I do not seek to survey the entire spectrum of theories, issues, and developments in the field, as such an undertaking could constitute a separate thesis and sketches of this sort have already been covered in handbooks on L1 acquisition (e.g., Ambridge & Lieven, 2011; Bavin & Naigles, 2015; Clark, 2024; Fletcher & MacWhinney, 1995; Lust, 2006). Instead, I will provide a brief overview of Chomsky's nativist theories within the generative paradigm and then touch upon constructivism, usage-based theories and emergentist approaches.

2.1.2. Nature and nurture

The contentious nature and nurture question, sparked in the 17th century, "Do [children] start out at birth with John Locke's *tabula rasa*, or do they come with certain things already pre-wired?" (Clark, 2024, p. 2) remains a central issue though it has evolved in the light of modern research. While the strict *tabula rasa* view has been largely discredited in cognitive science and linguistics, the broader discussion concerning the respective roles of innate mechanisms and environmental input in language acquisition persists. Hence, the underlying theme of the question continues to resonate in contemporary research, as attested by theoretical and empirical papers contending for or against language innateness (e.g., Behrens, 2021; Kern et al., 2014; Snyder, 2021; Valian, 2014).

The discussion has shifted from a binary opposition (*tabula rasa* vs innate knowledge) to a more integrated view with nuanced perspectives that recognise both biological predispositions and the role of experience and environmental input. Although the nature versus nurture dichotomy continues to exist in modern day linguistics, it lives on "with significant adjustments and somewhat less polarised views – a result of an attempt to base arguments on empirically driven findings in biology as well as the cognitive and brain sciences in the last decades" (Kedar & College, 2019, p. 2).

Following the cognitive revolution (see Section 2.1.1) sparked by Chomsky's (1959) critique of Skinner's (1957) behaviourist claims about language acquisition, theories and views centring on innate endowment gained traction and caught on steadily (e.g., Miller, 2003). As stated by Kedar and College (2019), the notion of innateness in contemporary research is linked to specific behaviours and forms of knowledge believed to be present from and even pre-birth. However, behaviourists and empiricists are generally sceptical and critical of the concept of innateness, especially of Chomsky's theory of Universal Grammar (UG) and the Poverty of the Stimulus Argument (POSA), which Skidelsky (2016, p. 144) describes as "the best-known argument in favour of the innatism of certain mental structures" (see Section 2.1.3).

Even though, in the current CLA research scene, there is a general consensus that the input children receive affects language development (e.g., Hoff, 2006; Tamis-LeMonda et al., 2014) and also that "the *potential* to acquire language is [...] innate" (Ambridge & Lieven, 2011, p. 103), the extent to which children are dependent on input in order to construct the grammar of their target language remains a point of contention (e.g., Adone, 2012; Narafshan et al., 2014). Indeed, as articulated by Karmiloff-Smith (1998):

All scientists studying normal and atypical development – from the staunchest Chomskyan nativist to the most domain-general empiricist – agree that development involves contributions from both genes and environment. The gulf between the theories lies in how genes and environment are claimed to contribute to developmental outcomes. (p. 389)

Despite the schism between nature and nurture theories, a number of researchers have adopted a middle ground, arguing that the two are highly intertwined in language development and that the boundary between them is, to an extent, superficial (see, for example, Clark, 2024; Morgan, 1986). This interconnectedness has led several CLA researchers to propose *tailoring nurture to nature*, that is, optimising linguistic input to amplify the brain's capacities and compensate for its limitations (He, 2022). Such an approach seeks to augment the infant's *intake*, namely the portion of input that is absorbed and utilised in the process of language acquisition (He, 2022, p. 5).[2]

Nature and nurture theories that trended in the 21st century have been remoulded and, in some cases, replaced, through fresh scrutiny, empirical evidence and scientific advancements (Goodluck, 2020). As Clark (2024, p. 19) observes, "despite a plethora of studies since the 1960s, there are still many gaps in what we know about acquisition, even for well-studied languages, and there are still too few language-types included." Maltese, for instance, remains seldom represented in cross-linguistic research on CLA (see Section 2.5.3).

[2] See Section 2.2.2 for more on the role of input.

2.1.3. Nativist and constructivist approaches (and their next of kin)

Nativist and constructivist approaches have been in a constant struggle since the meeting held at Royaumont Abbey in 1975, widely known as the Piaget–Chomsky debate (Piattelli-Palmarini, 1979, 1994). Although, at the time, constructivist proposals were at their inception and therefore not yet fully-fledged, the main distinctions between the two positions were delineated. As stated by Pérez-Pereira (2011, p. 260), "the present debate is situated in a very different landscape, in which the knowledge of language development has evolved enormously".

Both nativism and constructivism have made great strides, grounding their theoretical propositions in empirical findings (e.g., Goldberg, 2006; Pinker, 1984; Snyder, 2007; Tomasello, 2015). Yet there still is much to be desired in the field since, as argued by Ambridge and Wagner (2021, p. 859), among others, "theoretical developments have not kept pace with the ever-increasing body of empirical work, resulting in […] a dearth of detailed, up-to-date accounts, particularly those that apply cross-linguistically".

In what follows I will provide a brief overview of the current major competing theoretical approaches in the field of CLA research. These approaches can broadly be aligned with nativism or constructivism.[3] All of the frameworks presented below strive to answer the two crucial questions at the centre of CLA theories: the definition of the initial state (What is the underlying system like? What

[3] See Ambridge and Lieven (2011) for a detailed account of contrasting CLA approaches.

do children bring to the acquisition process?) and of the end-state (How do children diverge into their target language(s)?) (Ambridge & Lieven, 2011; Socarrás, 2011).

2.1.3.1. Nativist approaches

In this section, I outline several key components of the generative framework, all of which are interconnected since they stem from the same fundamental proposition: children are pre-programmed to acquire language, that is, that certain aspects of language are innate. In the nativist context, the broad definition of *innate* is 'present from birth'. For some researchers, arguably stauncher proponents of the innate perspective, *innate* takes on the meaning of 'genetically encoded' (e.g., Dediu & Graham, 2014). Further to this issue centring on the definition of *innate* lies the debate regarding what is in fact innate: linguistic knowledge (e.g., Fodor, 2001) or cognitive computational mechanisms (e.g., Li, 2009; Skidelsky, 2013, 2016). Hence, even within nativist approaches, points of contention are present.

Chomsky's nativist theory establishes a clear-cut distinction between language acquisition and language learning and describes the way in which children are biologically endowed to acquire language (Chomsky, 1959). Chomsky started out by proposing the Language Faculty as a way of explaining the initial state. Within this framework, he focused on what he termed a Language Acquisition Device, a mechanism that, based on the input received, selects a grammar from a number of possible grammars, thus resulting in children's rapid language development (Chomsky, 1965). Upon further research Chomsky himself rejected this theory, which received criticism from both rationalists and empiricists, stating that it is an "obviously false" assumption (Chomsky, 1975, p. 119).

From the Language Acquisition Device, Chomsky moved on to the more robust theory of UG to account for children's fast-paced acquisition of language, a theory which is still influential in current CLA research. In Chomsky's own words "[UG] may be thought of as some system of principles, common to the species and available to each individual prior to experience" (1981a, p. 7). As the term implies, and unlike Chomsky's previous Language Acquisition Device theory, UG is a formal, abstract and innately specified grammar that is designed to account for all of the world's languages. UG has recently been linked to the Language Capacity (Lenneberg, 1967), "adapting a traditional notion to a new biological context" (Chomsky, 2017, p. 200). The Language Capacity is exclusive to and consistent within human populations and is considered to be biologically inherent and independent of other cognitive systems. It comprises *language of thought* and an *externalisation* process (Berwick & Chomsky, 2016; Lenneberg, 1967).

UG underlies the Principles and Parameters (P&P) framework, which consists of "a set of universal *principles* which capture what underlies language structure everywhere, and a finite set of *parameters* to account for possible linguistic variation" (Lust, 2006, p. 55). According to this view, then, the task of children acquiring a language(s) is to set the parameters correctly, depending on the incoming linguistic data. This is illustrated through the so-called Switchbox metaphor which explains the way in which children select ('switch on') a certain set of parameters (e.g., the pro-drop parameter) to yield the grammar of the language(s) they are exposed to (Chomsky, 1988).

The P&P framework, a more general description which includes Government and Binding Theory (Chomsky, 1981b) remains influential, in particular in the field of syntactic acquisition. Snyder (2021), for instance, applies (a version of) the parametric approach to the acquisition of syntax and provides scenarios in which P&P hypotheses can be tested. As one of the major advocates of the P&P framework, he claims that "the parametric approach offers a single theoretical framework encompassing not only the particulars of CLA, but also the intricate linguistic judgements of adult speakers, and the amazing patterns of cross-linguistic variation" (Snyder, 2021, p. 885).

Through three case studies, Snyder (2021) argues that changes in children's syntax are "decisive, additive and interconnected" (p. 862). *Decisive* captures the leap children make from not producing specific syntactic structures to using them productively and correctly in their target language, leading to deterministic syntactic acquisition (Berwick, 1985). This is one way of explaining the general lack of errors of commission in naturalistic child language data. *Additive* reflects the incremental manner in which children set their target-language parameters, resulting in periods when some constructions are adult-like while others are not yet evident (Snyder, 2021). Finally, *interconnected* refers to the way structures can entail syntactic information about more abstract or seemingly unrelated structures, giving rise to multiple syntactic features concurrently. Since Snyder's (2021) case studies are based on English and Spanish, two well-studied Indo-European languages, it will be valuable to examine whether these three types of syntactic change are also evident in Maltese child language data.

Closely linked to UG is the POSA. Its main premise is that the input children receive is too impoverished for them to achieve rich target-language output; hence, the linguistic knowledge required for acquisition must be internally pre-wired. This argument, described by Kedar and College (2019, p. 5) as "a central crux" in CLA research, continues to shape theoretical and empirical work and remains pivotal in nativist–constructivist debates. Supporters of the POSA highlight several properties of the primary linguistic data – the input children receive – to demonstrate that these data are insufficient and underdetermined for children to attain linguistic competence so seamlessly. Two such properties are finiteness and degeneracy. *Finiteness* refers to the fact that children can understand and produce utterances they have never encountered within their finite linguistic experience. *Degeneracy* denotes that the input includes false starts and ungrammatical utterances, yet children still manage to abstract away from these disfluencies and inconsistencies, extracting the grammatical patterns underlying them (Skidelsky, 2016).

Further to UG, the P&P approach and the POSA, Chomsky developed the Minimalist Program (Chomsky 1995, 2001) with the goal of understanding the nature of language, once more from a nativist viewpoint. Aligned with Chomsky's previous theories and frameworks, the Minimalist Program also assumes a language faculty, consisting of a cognitive system and a performance system, the former storing linguistic information and the latter accessing and making use of this information. As the name implies, this theory of language takes on a "simple design" assumption in which the interface levels of language (the articulatory-perceptual and conceptual-intentional levels) are taken to be the only levels of language (Chomsky, 1995, p. 169). In fact, "at the heart of the [Minimalist Program] approach is the assumption that the locus of language variation is in the lexicon (not the syntax), expressed in the system as feature specifications on [Lexical Items]" (Socarrás, 2011, p. 1). This, of course, has important implications for language acquisition, in particular syntactic acquisition (see Section 2.4.2. on the acquisition of syntax).

Chomsky's influential theory of generative grammar underpins all the Chomskyan frameworks outlined above. "By a generative grammar I mean simply a system of rules that in some explicit and well-defined way assigns structural descriptions to sentences" (Chomsky, 2015, p. 6). Generativist theories of CLA thus assume that children possess innate knowledge of formal grammatical rules operating on abstract linguistic categories and phrases (e.g., *noun*, *noun phrase*). The term *generative* captures the idea that grammatical rules enable speakers to produce an infinite number of novel utterances (Chomsky, 1965; Guasti, 2002), a property known as productivity.

Although sometimes used interchangeably, nativism and generativism are distinct. A framework may be *nativist* without being *generativist*, assuming that certain aspects of linguistic or cognitive architecture are innate without tying this specifically to grammar, or *generativist* without being *nativist*, proposing that formal grammatical rules (in the Chomskyan sense) are learned rather than innate (Ambridge & Lieven, 2011). The lexical-principles account of word learning exemplifies the former (Golinkoff et al., 1994). In practice, however, most nativist and generativist approaches overlap and are subsumed under UG, which posits that innate (nativist) knowledge of formal grammatical rules (generativist) underlies all human languages (Ambridge & Lieven, 2011).

The upcoming section is devoted to constructivist approaches, which are also prevalent in CLA research and, like nativist approaches, have diversified into several related frameworks.

2.1.3.2. Constructivist approaches

The approaches described in this section are non-nativist in that they assume that children do not come equipped with innate linguistic knowledge. Constructivist, usage-based, and emergentist perspectives are often used interchangeably, yet they represent overlapping but distinct frameworks.

Proponents of constructivist approaches contend that while the potential and ability to acquire language are innate, linguistic knowledge – such as knowledge of grammatical categories – is not. They argue that children depend on the linguistic input available to them in order to construct their grammar (hence *constructivists*). At the heart of this approach is the notion that children, as language learners, are driven by an inherent need to communicate. In fact, according to Bruner (1983), a leading constructivist, the language learner must not only learn the linguistic system, that is, syntax, but also the ability to refer and express meaning and communicate effectively, that is, semantics and pragmatics.

Most constructivist frameworks stem from Construction Grammar (CxG), which is based on the fundamental premise that linguistic patterns (form) and meanings (function) are the building blocks of language (see, for example, Croft & Cruse, 2004; Goldberg, 2006). Within a CxG framework, language is assumed to be "an inventory of constructions (utterance templates) of various sizes and various levels of abstraction, each of which serves some communicative or socio-pragmatic function" (Ambridge & Lieven, 2011, p. 123). Constructivist theories, and CxG, are closely related to theories of cognitive linguistics to the extent that Cognitive Grammar (Langacker, 1987) is often argued to be a branch of CxG. Both CxG and Cognitive Grammar depart from Chomsky's nativist and generativist models, framing language acquisition within a cognitive linguistic paradigm. Cognitive Grammar assumes that language is an integral part of cognition and rejects the claim that humans have a specific faculty in the brain dedicated to language.

Under a constructivist view, then, children use their environment and input in order to induce the form-function pairings of their target language. In constructivist approaches to the acquisition of language, therefore, the end state is not reached by activating UG. Instead, "[the end-state] consists of networks of constructions […] [which] are the units of grammar: they can be as small as morphemes or as large as whole utterances" (Lieven & Brandt, 2011, p. 282). Thus, apart from being non-nativist, constructivist approaches are also non-generativist since the end-state is characterised by constructions rather than a system of formal rules.

Children start to build these *units* as structure and meaning develop simultaneously, abstracted from their (ideally) linguistically rich environments. They start out by learning very specific strings which over time become schematised and productive (e.g., Dąbrowska, 2015; Langacker, 2000). At an early stage, while these units are productive, they are claimed to be low-scope since they are not fully abstract and are known as lexically specific or item-based schemas. Once similarities between these schemas emerge, children form new schemas by analogy and reach the stage of abstraction, that is, the adult end-state. Lieven and Brandt (2011) state that children's early grammar differs from the adult target-grammar in terms of *degree*. This is because, initially, "children's constructions are either fully lexically-specific or contain relatively low-scope slots […], they are also simpler with fewer parts [and] exist in a less dense network – they are more 'island-like'" (Lieven & Brandt, 2011, p. 286).

In their 2011 paper, Lieven and Brandt discuss the acquisition of inflectional marking and utterance-level constructions, taking examples from CLA literature on Cantonese, Dutch, English, German, Polish and Spanish in order to support the constructivist approach they put forth. In a similar fashion, Behrens (2021) draws examples from English, German and Dutch to provide evidence in favour of the constructivist predictions outlined in her paper. The predictions and hypotheses made in these studies, among others, can be tested on the basis of the data collected in the present study. The analysed data will be explored to confirm whether MFLA aligns with such features employed by constructivist theorists to strengthen their approach.

Usage-based approaches occupy a major branch of constructivism. Such approaches are also termed *functional* since they assume that the driving force of the language acquisition process is children's desire to communicate and understand the utterances of those around them (Ambridge & Lieven, 2011). In Tomasello's words (2000, p. 62), in a usage-based account, the inventory of constructions a person possesses "result[s] from her accumulated experience with language across the totality of usage events in her life".

From a usage-based perspective, constructions are learned through the combination of general cognitive and social skills, with no reliance on specialised innate linguistic knowledge (e.g., Ibbotson, 2020). Usage-based proponents outline four skills which they deem vital for children to be able to acquire language, namely, distributional analysis,[4] analogical reasoning, competition and intention reading (see, for example, Rowland (2014)[5] for a definition of each skill). In such models, the acquisition process is aligned with that described by constructivists (and briefly explained above), that is, learning structures from the incoming input, forming schemas by generalising over the input and abstracting away as the process progresses until finally reaching the adult end-state. Therefore, similar to constructivism, usage-based approaches "focus on the homology between the information attested in the input and the children's developing language use" (Behrens, 2021, p. 959).

Constructivist and usage-based theories emphasise the role of individual differences (IDs)[6] in language acquisition, asserting that such variability is "expected" (Dąbrowska, 2020). This expectation arises from the reliance on

[4] Distributional analysis is crucial to other constructivist theories, too, apart from usage-based models, namely, *semantic-distribution analysis* introduced by Braine (1963).
[5] Rowland (2014) defines these respectively as pattern detection in input, generalisation across forms, competition between alternatives, and inference of communicative intent.
[6] See Kidd and Donnelly (2020) for a detailed account of IDs in CLA.

domain-general learning mechanisms, rather than a pre-specified UG, combined with the fact that children receive different linguistic input and experiences. At the same time, these approaches also account for commonalities in acquisition by positing that the underlying cognitive learning processes used to extract structure from input are shared across learners. A key mechanism proposed by usage-based theories is *pattern association*, which enables children to draw analogies and generalise from lexically specific schemas.

Pattern association is at the heart of connectionist approaches to language acquisition (e.g., Elman et al., 1996; Joanisse & McClelland, 2015), which are closely aligned with usage-based frameworks. In the majority of constructivist and usage-based approaches, cognitive processes are coupled with social processes (e.g., Tomasello, 2003) so as to best explain the way in which form-function relationships are established. This combinatorial approach is described in Schmid (2020, p. 3), who states that "usage and the social and cognitive processes keep reinforcing each other […] and this allows shared conventions and individual knowledge to emerge, persist, change and embrace more or less creative innovations […] in language". Schmid's calculated choice of words here, with the use of the verb "emerge" underscores the alignment of usage-based approaches with emergentism, in contrast to Chomskyan nativism.

Emergentist models are conceptually related to constructivism and usage-based frameworks. In this context, *emergentist* is used to refer to the way in which children's inventories of constructions emerge through language exposure, experience and use, that is, from usage-events rather than pre-programmed knowledge. As claimed by Behrens (2021, p. 961), “current usage-based models see linguistic structures as an emergent property of communication and routinisation” through which language is acquired. Consistent with constructivist and usage-based accounts, emergentist approaches posit that language acquisition is driven by general cognitive and social abilities, rather than a domain-specific linguistic module.

From an emergentist viewpoint, as learners, children make use of previous language experience as the foundation on which to construct new linguistic knowledge from the incoming input. Tomasello (2000, p. 62), explains the way in which “accumulated linguistic experience undergoes processes of *entrenchment*, due to repeated uses of particular expressions across usage events”, which then leads to abstraction. The *entrenchment hypothesis* was first proposed by Braine and Brooks (1995) and has since led to various entrenchment models in CLA research, such as Schmid’s framework of entrenchment and conventionalisation (e.g., Schmid, 2020). This hypothesis seeks to provide an explanation for the fact that children normally master high-frequency constructions which they hear being used over and over again and are more likely to use low-frequency constructions ungrammatically or in an unattested way.[7] When used correctly and productively, it is clear that such items/constructions have become solidly established in the child’s inventory of constructions.

[7] Another way of explaining this phenomenon in CLA is the *pre-emption hypothesis*. See, for example, MacWhinney (2004).

Additionally, “in essence, overgeneralisation errors occur with [items] whose correct usage has not yet been firmly entrenched” (Rowland, 2014, p. 109). As has been shown in numerous studies, children are more likely to make overgeneralisation errors with low-frequency items in elicitation production tasks, and both adults and children are more likely to accept incorrect grammatical structures as correct if they involve low-frequency items rather than high-frequency items in grammaticality judgement experiments (e.g., Ambridge et al., 2008; Theakston, 2004). Note that the terms *items* and *constructions* are being used liberally, referring to units ranging from morphemes, such as the English plural morpheme -*s* and the Maltese plural morpheme -*ijiet*, to strings of words such as NPs.

Having outlined the core tenets of nativist and constructivist approaches, it is important to note that this overview has necessarily been selective. Other influential models that cannot be discussed in depth here include MacWhinney and Bates’ (1989) Competition Model, hybrid accounts such as Ninio’s (2011) integrative approach, and more recent proposals like Ambridge’s (2020) Radical Exemplar Model, which posits that “unwitnessed forms are produced and comprehended by on-the-fly analogy across multiple stored exemplars, weighted by their degree of similarity to the target with regard to the task at hand” (Ambridge, 2020, p. 511; see Section 2.4.2).

2.2. Sociocultural context and input

This section first considers language socialisation and the diverse cultural contexts in which children acquire language (Section 2.2.1) and then examines the multiple dimensions of input (Section 2.2.2). The discussion on input addresses questions of input quality, quantity, and timing, before turning to different forms of input: child-directed speech (CDS), child–child interactions, and input mediated through technology.

2.2.1. Language socialisation and cultural contexts

CLA is intrinsically interrelated to socialisation and culture. The late 1970s gave rise to a surge of research on language socialisation which has expanded both in scope and diversity to investigate how individuals, particularly young children, acquire communicative and cultural competence within their social groups. The term *language socialisation* refers to "both socialisation through language and socialisation to use language" (Ochs, 1987, p. 2). By foregrounding cultural practices and social interactions while also recognising concurrent neurodevelopmental influences, this concept serves as an interdisciplinary nexus, interconnecting anthropology, sociology, psychology and linguistics, and provides "a paradigm that assumes both nature and nurture as implicated" (Ochs & Schieffelin, 2012, p. 9). According to Ochs and Schieffelin (2017, p. 6), language socialisation involves "the language-acquirer as a child born into a lifeworld saturated with social and cultural forces, predilections, symbols, ideologies, and practices that structure language production and comprehension over developmental time".

Cross-cultural research on children's linguistic environments continues to broaden the understanding of the diverse conditions under which language socialisation occurs. For instance, cross-cultural studies on CDS[8] across a variety of communities have challenged and disproven the assumption that the characteristic of CDS in well-studied Indo-European languages are universal features of children's linguistic input (Kidd & Garcia, 2022). Similarly multimodal semiosis (joint-attention-getters), such as gaze, touch and pointing, are also non-universal components which facilitate language acquisition.

By way of example, children in Zunil communities (Guatemala) acquiring K'iche' Mayan are considered conversational partners only once they begin to speak (Pye, 1992), and San caregivers in Botswana avoid eye gaze with their infants (Takada, 2012). In Javanese communities (Indonesia), infants are typically seated facing outward on caregivers' laps and were never observed in face-to-face positions (Smith-Hefner, 1988), while Tzeltal Mayan infants (Mexico) receive as little as 3.6 minutes of directed speech per hour (Casillas et al., 2020). By contrast, on Rossel Island (Papua New Guinea, language Yélî Dnye), infants are surrounded by caregivers and passers-by who engage with and respond to them (Brown, 2012), and Zinacantec Mayan infants (Mexico) are similarly immersed in interactional scenarios, though triadic formats are preferred over dyadic interactions (de León, 2012).

[8]Also termed *motherese*, *parentese* and *infant-directed speech*. See Section 2.2.2 for more on CDS.

The examples illustrated above accentuate the importance of language being understood as an integral component of a communal system, encompassing "not only […] the speaker's utterances but also […] gesture, stance, facial expression, affective display, and any other factors that contribute to successful interaction" (Clark, 2024, p. 24). Consequently, as argued by Ochs and Schieffelin (1995), a comprehensive account of grammatical development necessitates an in-depth examination of the sociocultural environment in which language acquisition occurs. In linguistically diverse communities, analyses of grammatical development must be contextualised within the prevailing language ideologies to fully capture the sociocultural dimensions of language use.

From a young age, children exhibit an awareness of how grammatical structures within a linguistic code function as markers of social identity, selecting forms that align with their perceived identity. Thus, "culture affects grammatical development in surprising and subtle yet systematic ways" (Ochs & Schieffelin, 1995, p. 93). However, to-date, sociocultural considerations remain largely absent from most grammatical language acquisition accounts, leaving gaps in our understanding of the extent to which such factors affect acquisition and development. Ultimately, CLA is inseparable from the broader process of learning to communicate within a sociocultural framework, and hence, this facet should be explored and integrated within theoretical accounts in the field.

The framework of language socialisation shares common ground with the *social–pragmatic* account in emphasising the role of social interaction in language acquisition. Yet, the latter differs in many ways from the language socialisation framework outlined above. *Social–pragmatic theories*, grounded in emergentism, argue that children acquire language by understanding others' intentions in social contexts (e.g., Akhtar & Tomasello, 2000). Advocates of this view contend that nature and nurture both play a role in the acquisition of language (e.g., Tomasello, 2022), but that there are no genes specific to language (e.g., Elman et al., 1996). Under this view, children pick on *social cues* in their quest to acquire language (Baldwin & Moses, 2001). Consequently, the social–pragmatic account highlights the role of joint attention, intention-reading and communicative goals in language development and explains acquisition in terms of sociocognitive and interactional processes that facilitate communication.[9] Therefore, while language socialisation explores language acquisition as part of broader socialisation into cultural and linguistic practices, social–pragmatic accounts focus more on the way in which children acquire words and infer meaning from social cues.

2.2.2. Input

The role of input in CLA has been a focal point of research in the field, underscoring its significance in shaping linguistic competence (e.g., De Houwer, 2018; Gallaway & Richards, 1994; Khattab & Keren-Portnoy, 2025; Valian & Casey, 2003). Input has a multifaceted role in the language acquisition process, with its influence permeating different tiers of acquisition and development, such as phonology, syntax and lexis (e.g., Babatsouli, 2024; Behrens, 2006; Lieven, 2010). The role of input is critical to the field within all contexts of acquisition, including bilingual settings and those involving children with developmental language disorders. In addition, specific

[9] See Rowland (2014) for a brief account of the social–pragmatic view and evidence for and against this theory.

features of input, such as frequency effects, are also important since, as emphasised by Ambridge et al. (2015, p. 264), they "constitute a phenomenon that any successful account of child language acquisition must explain".

While literature in the field focuses mainly on linguistic input or interactive input, Rowe and Snow (2020) call for CLA researchers to conceptualise and analyse input on three dimensions simultaneously rather than on one single dimension, these being a) the *interactive dimension* (e.g.., responsiveness towards the child), b) the *linguistic dimension* (e.g., use of recurrent structures), and c) the *conceptual dimension* (e.g., abstract topic discussions). Under this view, "input best designed to promote language learning is interactionally supportive, linguistically adapted, and conceptually challenging for the child's age/level" (Rowe & Snow, 2020, p. 5). Naturally, all three dimensions are dependent on the sociocultural context (see Section 2.2.1).

High-quality input of the sort described above by Rowe and Snow (2020) creates meaningful and engaging contexts that support the child's language acquisition. Although it has been argued that quality carries more weight in the acquisition equation (e.g., Anderson et al., 2021; Rowe, 2012), quantity has also been proven to contribute to the trajectory and pace of acquisition. This positive correlation between quantity and acquisition outcomes is particularly evident in specific areas of acquisition, namely, statistical learning (e.g., Saffran, 2020).

Arguments in favour of the importance of input quality over quantity stress that "not all talk is equally influential" (Rowe, 2012, p. 1773). In fact, it has been shown that a) *overheard speech*, that is, speech not directed to but heard by the child (e.g., Weisleder & Fernald, 2013), and b) *non-contingent screen-speech*, that is, speech emerging from a screen which does not involve social interaction between the speaker and hearer (e.g., Roseberry et al., 2014; Zimmerman et al., 2007), are not driving forces of early language development. As stated by He (2022, p. 3), "an increase in input quantity is only meaningful when it is intake-able". This is also the case with quality. Thus, for instance, input that is overly specific or information-dense relative to the child's processing capacities is not optimal (e.g., Arunachalam et al., 2013), as it may place undue demands on developing processing mechanisms, creating a mismatch between input and intake.[10]

He (2022) also explores the concept of *input timing* in order to tailor nurture to nature.[11] Recent empirical evidence demonstrates a positive correlation between sleep and early language acquisition (e.g., Gómez et al., 2006; Horváth & Plunkett, 2016; Klinzing et al., 2019; Williams & Horst, 2014). Sleep consolidates memory of language input which in turn facilitates early language development. Hence, as argued by He (2022, p. 5), "provided the magic effect of post-learning sleep, parents may consider arranging quality input before naptime or bedtime". Furthermore, as discussed by Blum-Kulka (2017), dinnertime is also an ideal setting for such quality input and language socialisation, at least in most Western societies. This is because it enables face-to-face interaction and joint engagement of the child and the adult,

[10] For more on input, intake and frequency effects see Ambridge et al. (2015), Newman et al. (2016), and Rowe (2015), amongst others.
[11] The concept of tailoring nature to nurture is introduced in Section 2.1.2.

and incorporates a rich array of conversation genres, among others (e.g., Blum-Kulka, 2017).

CDS is one of the most notable *input tailoring* strategies across many cultures, combining prosodic exaggeration, repetition, and simplified yet well-formed structures with interactional features such as turn-taking and joint attention (e.g., Estigarribia & Clark, 2007; Hadley et al., 2011; Lester et al., 2022; Tippenhauer et al., 2020). These characteristics facilitate phonological, lexical, syntactic, and pragmatic development by enhancing the saliency of input and supporting interactional processes (e.g., Weisleder & Fernald, 2013; Hsu et al., 2017). The effectiveness of CDS depends on developmental alignment, whereby input must pace in tandem with the child's evolving linguistic abilities; as children grow, CDS typically incorporates more abstract concepts and increased linguistic complexity.[12] In this way, it supports both linguistic and cognitive development, including processes related to neural plasticity (e.g., Kuhl, 2010). While socioeconomic status (SES) may influence input quality, large-scale evidence points to more consistent effects of factors such as child age, overall adult speech exposure, maternal mean length of utterance (MLU),[13] and language-specific complexity (e.g., Bergelson et al., 2023; Ibbotson & Browne, 2024). *Babytalk lexicon*,[14] as part of CDS, reflects systematic adaptation, serving functions of simplification, clarification, and

[12] See Rowe and Snow (2020) for more on age and input.
[13] See Sections 2.3.2. and 3.6. for more on MLU.
[14] Some researchers use the terms *CDS* and *babytalk* interchangeably (e.g., Kalashnikova et al., 2017; Soderstrom, 2007) to refer to the entire register of speech directed at infants encompassing all modifications, while others reserve the term *babytalk* to refer strictly to alterations made to specific words/phrases. Under the latter view, babytalk is a subset of CDS and, therefore, not all CDS necessarily involves babytalk.

affective expression through modified, context-bound forms (e.g., Ferguson, 1964; Gutova, 2016).

While parental speech has been widely studied, research on sibling exposure remains limited (e.g., Tsinivits & Unsworth, 2021). Older siblings may provide additional models, scaffolding, and exposure, particularly in bilingual contexts, although outcomes depend on interaction quality and sibling characteristics, such as age gap and sex (e.g., Sorenson Duncan & Paradis, 2020; Taylor & Kan, 2018). In addition, the arrival of a younger sibling alters family dynamics and interaction patterns (e.g., Volling et al., 2017). While there is wide variation in how children respond to the birth of a younger sibling, it has been found that older siblings often assume roles of caregiver and instructor, while younger siblings act as learners and observers (e.g., Howe et al., 2016, 2022; Tavassoli et al., 2019).

Children's input environments, however, extend beyond the family to include peer interactions in childcare and preschool settings, which also contribute to language development (e.g., Henry & Rickman, 2007; Theobald et al., 2017). Within these environments, educators further shape acquisition through the quality of input and interactional support (e.g., Wasik et al., 2006), with language growth closely linked to responsive, structured, and cognitively stimulating practices, particularly in multilingual contexts (e.g., Hamre & Pianta, 2007; Zheng et al., 2021).

In an era of highly technologised environments, traditional input sources coexist with digital media, which form part of many children's daily lives, particularly in the Western world. While certain interactive applications may support aspects of language learning, especially when embedded in social interaction, non-contingent or excessive screen exposure has been associated with weaker outcomes (e.g., Massaroni et al., 2024). The effects of digital input are contingent on factors such as content, duration, and caregiver involvement, with co-engagement helping to render screen-based input more interactionally meaningful (e.g., Vescovi et al., 2024). At the same time, device use may disrupt caregiver–child interaction, a phenomenon termed *technoference* (McDaniel & Radesky, 2018), though similar effects have been observed with non-digital distractions, pointing more broadly to the importance of sustained attention in interaction (e.g., Chamam et al., 2024). Overall, current research suggests that the impact of digital technology on language development is mediated less by its presence per se than by its effects on the quality, contingency, and richness of children's input environments.

2.3. Patterns and stages of development

"Infants don't produce their first words until age one or later, but by three or four they can talk quite fluently about some topics" (Clark, 2024, p. 14). This section describes the way in which language is acquired and mastered so naturally within the first few years of life. Section 2.3.1 reviews the main developmental milestones in L1 acquisition, followed by Section 2.3.2, which examines the pioneering work of Roger Brown and his students and its lasting impact on the field. Section 2.3.3 focuses on early expressive vocabulary, while Section 2.3.4 addresses bilingual CLA, an area central to this study given that all participants are bilingual to varying degrees.

2.3.1. From babbling to narratives

L1 acquisition follows a series of developmental milestones that are generally relatively consistent across TD children. "Milestones represent norms, 'standards', for children's expected language development" (De Houwer, 2024a, p. 14). Such milestones, which include babbling, understanding and producing words, and forming phrases, reflect the progression from initial sound recognition to complex language use, typically achieved within the first few years of life. In CLA literature, these milestones are categorised into stages encompassing various aspects of language development, commonly referred to as a) the pre-linguistic stage (0–1;0)[15], b) the one-word/holophrastic stage (1;0–1;6), c) the two-word stage (1;6–2;0), d) the early multiword stage (2;0–2;6), e) the later multiword stage (2;6+), and e) adult-like grammar (4+).

In the first year of life, infants progress from using solely reflexive vocalisations, such as crying and sneezing, to canonical babbling. Between 0;2–0;4, cooing, chuckling and laughter emerge (e.g., Cattell, 2000), with infants experimenting with the production of vowel-like sounds and consonant sequences. Building on these earlier vocalisations, vocal play at around 0;4–0;6, during which infants explore pitch, amplitude and intonation, among others (e.g., Stark, 1980). Between 0;6–0;10, infants proceed to canonical babbling (e.g., *da-da*), that is, the repetition of consonant–vowel combinations (e.g., Gervain, 2020; Oller, 2000). Canonical babbling is often considered as the precursor to speech by CLA researchers (e.g., Majorano et al., 2014); however, this view is not shared by those

[15] Age in *years;months*.

who argue for a discontinuity between babbling and first words (e.g., Jakobson, 1968).

During this prelinguistic phase, it has been observed that "the most frequent occurrence of canonical babbling appears to correspond to the much more rare occasion of vocal turn-taking" (Long et al., 2022, p. 23). Interestingly, even in deaf children exposed to sign language, analogous manual babbling emerges between 0;6–0;10 (Petitto & Marentette, 1991). At this stage, infants as young as 0;8 employ statistical relationships to successfully segment words from fluent speech (e.g., Pelucchi et al., 2009), and at around 0;9 word comprehension is underway (e.g., Syrnyk & Meints, 2017).

Clark (2024, p. 101) also notes that "towards the end of their first year, infants begin to use two broad classes of gestures: pointing and reaching", apart from the use of word sequences. The use of gestures typically emerges between 0;8–1;0, with the onset depending on the type of gesture, such that deictic gestures normally feature prior to the repertoire of iconic gestures, although the emergence and use of gestures depend also on a host of other factors, like SES (e.g., Goldin-Meadow et al., 2014). Gesturing is vital in the language acquisition trajectory since, for instance, early gestures predict vocabulary development and size, and word–gesture combinations often act as precursors to the two-word stage (Goldin-Meadow, 2015). In Goldin-Meadow's own words (2015, p. 202), gesturing "offers us what may be the clearest view onto the skills that children themselves bring to language learning".

Around their first birthday children enter the one-word/holophrastic stage during which single lexical items are used to express entire propositions or desires (e.g., *ball= I want to play with the ball*). As stated by Clark, by the time infants start talking, “they have already had about twelve months of perceptual and conceptual development [. …] setting up representations of what they see and know” (2024, p. 7). Early vocabularies tend to be dominated by nouns, particularly those referring to salient objects and social agents and are confined to a vocabulary size of around 50 words (e.g., Dromi, 1987; Nelson, 1973).

At this stage, words are normally uttered in familiar settings, not necessarily fully intelligible and often simplified/form part of babytalk lexicon (e.g., *ballie = ball*). While over- and under-extensions have traditionally been reported to extend well into the second year of life (e.g., Rescorla, 1980), such as referring to all spherical objects as *ball* (over-extension), their status as developmental phenomena has been questioned (e.g., Ambridge & Lieven, 2011). In addition, while children continue to produce babbled sequences and nonwords, understood here as child-produced forms that do not yet correspond to existing lexical items, the ratio of real words to nonwords is around 3:1 during this period (Peters, 1983).

The one-word stage also includes the production of multi-word phrases learnt as one formulaic sequence, such as *clean up* (e.g., Clark, 1993). The child treats these phrases as unanalysed wholes, i.e., fixed expressions stored and functioning as single items (e.g., Peters, 1983). By way of example, in July of the year 2000, my family went on a trip to Italy for the Catholic Church’s Great Jubilee, and it is reported that at the time, at the age of 0;11, I regularly uttered the phrase *viva l-*

papa ('long live the pope'), which I presumably produced as a sequence involving rote-learning as a single unit.

By 1;6–2;0, toddlers typically begin combining words, producing two-word utterances to convey meaning (e.g., Berk & Lillo-Martin, 2012). Children start demonstrating awareness of syntactic relationships, in so doing marking the onset of grammar acquisition (e.g., Brown, 1973; Gervain, 2020). Although a *vocabulary spurt* has often been proposed for this stage (e.g., Bloom, 1970), this claim has been challenged, with evidence suggesting more gradual and variable lexical development (e.g., Bloom, 2004; Clark, 2024). Commonly, at this stage children make use of telegraphic speech, that is, they omit function words and grammatical morphemes, a feature also often observed in CDS, especially that directed to children developing atypically (Lorang et al., 2020).

As they embark on the second year of life, children enter the early multiword stage, and later that year, at around 2;6, they progress to the later multiword stage. The emergence of negation, questions, coordination and subordination are typical around the age of 3;0 (e.g., Brown, 1973). From this point onwards, children produce increasingly complex utterances, incorporating grammatical morphemes, auxiliaries, and function words, showcasing an understanding of morphosyntactic rules. At this age children produce both correct and overgeneralised forms, such as attaching the English *-ed* suffix to regular and irregular verbs, producing forms like *buyed* instead of *bought* (e.g., Dressler, 2011).

By around 3;0, children's vocabularies expand rapidly; for English-speaking children, this growth typically reaches approximately 8,000 words (Berko Gleason, 1985). As children's vocabulary size expands, they make inferences and use a number of strategies to acquire meaning along with new words and, according to Clark (2024, p. 169), they "rely critically on pragmatic knowledge to interpret the intentions of others". Apart from acquiring basic grammatical structures and semantic and pragmatic knowledge, children also gain a good grasp of their language's sound system, giving importance to articulatory routines, intonation and fluency, among others (e.g., Hoff, 2013).

By 4;0–5;0, most children possess an adult-like command of their L1's core systems, although fine-tuning continues throughout childhood and into early adolescence across the board (e.g., Berman, 2016; de Villiers & de Villiers, 1973). For instance, as stated by Hoff (2013, p. 134), "the full repertoire of speech sounds develops only gradually, with children typically sounding adult-like by about 7 years of age". Furthermore, Ravid and Avidor's (1998) production study on derivational morphology in Hebrew, which, like Maltese, is a derivationally complex language, reveals that the acquisition of derivational nominals takes place between the ages of 8–15 years. Children's ability to use language in socially appropriate ways, together with increased metalinguistic awareness, also continues to develop as they grow older. Hence, while mastery of many different levels of language is achieved by age five, an age at which, for example, overgeneralised forms gradually disappear (e.g., Maslen et al., 2004), the developmental trajectory continues to progress to adolescence.

Identifying developmental milestones through empirically driven research allows for a better understanding of typical language development and aids in recognising and addressing atypical trajectories (e.g., Kover et al., 2016; Schmitt et al., 2017). Yet, these milestones are approximate and vary across individual children, languages, and contexts, such as bi/multilingual acquisition settings (e.g., De Houwer, 2024a; Thordardottir, 2015; Tulloch & Hoff, 2023). IDs are not arbitrary but reflect a combination of intrinsic child characteristics and environmental influences that shape language learning outcomes (Paradis, 2023). Thus, while milestones provide a useful general framework, they can obscure the nuanced variability of language development. Researchers therefore advocate for a more individualised approach to CLA, emphasising that "language acquisition is a process characterised by stable and meaningful IDs" (Kidd & Donnelly, 2020, p. 332; see also Lieven, 1994).

2.3.2. Brown 1973 and beyond

Roger Brown's pioneering 1973 work, *A First Language: The Early Stages*, is one of the most influential contributions to the study of CLA to date. His longitudinal research on three English-speaking children, Adam, Eve, and Sarah, provides a detailed and systematic account of early morphosyntactic development in naturalistic settings. Brown's work is a cornerstone in CLA research due to its innovative methodologies and analyses, through which he introduced standardised approaches to understanding how children acquire language, focusing on the sequence of acquisition and morpheme inventories, among others (MacWhinney, 2008). His work therefore laid the foundation for subsequent research by providing a robust framework for analysing children's spontaneous speech and testing claims about the morphosyntactic development.

Through his research, Brown emphasises the importance of using child-adult interactions, naturalistic observations and longitudinal data to test theoretical claims, advocating for a systematic methodology in collecting and analysing child language data. The concepts of *obligatory context* and the *90% criterion of correct use in three successive speech samples* are examples of Brown's methodological innovations. In his research, he employs these concepts to determine when a child has acquired a particular morpheme, that is, under this view, mastery is reached when a child correctly uses a morpheme in 90% or more of contexts in which it is required, hence *obligatory* (Brown, 1973).

Brown's most enduring contribution is his identification of five developmental stages in the acquisition of fourteen grammatical morphemes in English, which he ranked according to the order in which they typically appear, not by age, but by MLU, which measures the average number of words/morphemes per utterance. The use of MLU as a developmental metric offers a more sensitive measure of linguistic development than chronological age, which can vary significantly among children. As put by Brown himself (1973, p. 53), "[MLU] is an excellent simple index of grammatical development because almost every new kind of knowledge increases length". This approach is now widely established in language development research as it enables more precise tracking of linguistic progress, and it is adopted in the present study (see Section 3.6). MLU can be calculated either in words (MLUw) or morphemes (MLUm).[16]

Brown found that children in his sample acquire grammatical morphemes in a relatively consistent order, largely determined by the semantic and grammatical complexity of these morphemes, despite individual variation in timing. For example, he shows that the progressive -*ing* tends to be acquired earlier than the third person singular -*s*. His findings demonstrate that grammatical development proceeds in predictable sequences, shaped by cognitive, linguistic, and input-related factors, rather than through simple imitation or reinforcement. Interestingly, Cazden's (1968) U-shaped development curve in the acquisition of grammatical morphemes, which also depicts a consistent order of acquisition, uses the data of Brown's study, published later in 1973.

The classic U-shaped trajectory, documented in earlier work particularly in the domain of verb morphology (see Ervin & Miller, 1963), refers to the pattern observed when the proportion of correct forms is plotted against time or developmental stage: accuracy is initially high, then declines as children begin to overgeneralise, before rising again once rule-based knowledge is established. This pattern illustrates how children may first produce correct forms through rote learning, then apply rules too broadly as they start to identify patterns and finally stabilise their use through rule-based understanding. Numerous critiques have

[16] In this thesis, MLU is used as the general term, while MLUw and MLUm are used when specifically referring to MLU calculated in words and morphemes, respectively.

been raised regarding such rigid U-shaped trajectories; errors occur at various stages of development, evolve over time, and are generally infrequent overall (Behrens, 2015).

In a similar fashion to Cazden (1968), other researchers also drew on the data collected by Brown and his students. Moerk (1980), for instance, reanalysed the data to show a correlation between parental input frequency and language acquisition, which Brown initially downplayed, instigating further research on the role of environmental factors in CLA. Furthermore, Brown's methodologies have been applied to:

- other child language studies of English-speaking children to validate his findings (e.g., Balason & Dollaghan, 2002; Roy et al., 2013),
- CLA in languages other than English to examine the universality of the language acquisition patterns described in his work (e.g., Allen & Crago,1992; Tomas & Dorofeeva, 2019),
- research on children with communicative disorders to compare typical and atypical trajectories (e.g., Tek et al., 2014; Werfel, 2018) and,
- bi/multilingual acquisition research, taking into consideration that Brown’s participants were three monolingual speakers of English (e.g., Davison & Hammer, 2012; Padilla, 1978).

These studies have influenced both theoretical models of syntax acquisition and practical approaches to language development in educational and clinical contexts.

One specific analytical method taken up widely across the field is MLU, nowadays considered a standard measure in CLA (see Section 3.6). MLU has been applied to several languages (e.g., Georgian, an agglutinating language (Tchintcharauli et al., 2024)) and across both monolingual and bilingual acquisition contexts (e.g., a Polish–English bilingual CLA study (Otwinowska et al., 2022)), including Maltese (Portelli, 2005; see Section 3.6.2). Additionally, over the years, measures encapsulating a similar notion to that of MLU, that is, to compare cohorts of children based on their linguistic development rather than chronological age, have been documented. These include, but are not limited to, the Index of Productive Syntax (Scarborough, 1990), Vocabulary Diversity (Malvern et al., 2004) and the First of Regular Uses (Snyder, 2007), each focusing on a slightly different aspect/approach of morphosyntactic and lexical development and which can also be used in tandem.

In many languages, measures of morphosyntactic development can be computed automatically using tagged corpora. For instance, the Computerized Language Analysis (CLAN) program provides automatic morphosyntactic analysis for data transcribed in the Codes for the Human Analysis of Transcripts (CHAT) format and supports a wide range of languages, offering measures of linguistic development such as MLU, Index of Productive Syntax, and Vocabulary Diversity (MacWhinney, 2008). This functionality is not yet available for Maltese child language data, which currently requires manual annotation or the development of a morphological analyser to enable compatibility with CLAN. Work on computational morphology for Maltese has been undertaken (e.g., Borg & Gatt, 2017); however, the tools developed have not yet been adapted for the automatic analysis of Maltese child language datasets (see Section 3.7).

While MLUm has long been used as a core measure of morphosyntactic development, it is important to acknowledge that its calculation is not theory neutral, as it presupposes that utterances can be segmented into discrete morphemes that are psychologically real for the child. Usage-based and emergentist approaches to CLA have questioned whether morphological structure is necessarily represented in term of discrete combinatorial units in early language (e.g., Baayen et al., 2011). Under this view, forms such as English *ducks* may be represented and processed holistically rather than as *duck + -s*, particularly in early stages of acquisition. These considerations are especially relevant in morphologically rich and typologically diverse languages, where the identification of morpheme boundaries is not always straightforward. In Semitic languages, including Maltese, templatic morphology, phonological alternations, and cliticisation further complicate segmentation decisions (see Section 3.6).

2.3.3. Early expressive vocabularies and the noun–verb bias debate

Research on early expressive vocabularies has examined both the pace of lexical growth and the make-up of children's first words. The latter addresses stylistic variation, that is, IDs in the acquisition of POS, and developmental variation, namely, shifts in the distribution of POS over time (Bates et al., 1994). A central issue within this line of research concerns the extent to which early vocabularies are characterised by a noun bias, and how this pattern evolves over development.

Gentner's (1982) influential cross-linguistic study of six languages galvanised the field, prompting monolingual, bilingual, atypical, and cross-linguistic investigations (e.g., Checa et al., 2016; Xuan & Dollaghan, 2013). Centred on the *natural partition hypothesis*, her account predicts that concrete nouns are acquired before verbs, leading to the proposal of a universal noun bias. This claim sparked debates about whether the pattern is driven primarily by conceptual salience or by linguistic input (e.g., Gentner & Boroditsky, 2001), as well as challenges to universality from studies of early vocabularies in Mandarin Chinese and Korean (e.g., Choi & Gopnik, 1995; Tardif, 1996), among others. A persistent difficulty in examining these issues is the limited comparability of findings across studies due to heterogeneous data collection methods and analytical approaches.

Many investigations of early expressive vocabulary rely on the MacArthur-Bates Communicative Development Inventories (CDIs; Fenson et al., 1993, 2007), a parent-report instrument. Analyses of large CDI datasets converge on a characteristic developmental profile: a) children's first 100 words or so are dominated by common nouns, b) between about 100 and 400 words, the proportion of verbs and other predicates increases, and c) with the onset of multiword speech, closed-class/function words rise sharply. Bates et al. (1994), using data from 1,803 children, articulated this three-wave progression, linking lexical composition to the emergence of grammar.

Cross-linguistic comparisons reveal both regularities and meaningful variation. For instance, Bornstein et al. (2004) found that, across seven languages, nouns generally occupy a larger share than other classes by around 1;8.

Contrastingly, languages such as Mandarin show a smaller noun advantage, in line with caregiver input profiles (Tardif, 1996). Moreover, recent Wordbank analyses pooling tens of thousands of CDI observations across roughly 30 languages show both consistency and systematic variation across languages and categories as vocabularies grow (Braginsky et al., 2019; Frank et al., 2021).

Notwithstanding their strengths for large-scale and comparable cross-linguistic work, CDIs have been criticised for over-representing nouns in checklist items, raising concerns that parent reporting and language sampling may tap different facets of lexical composition. In fact, several researchers argue that parent reports may favour a noun bias (e.g., Pine et al., 1996). To address such limitations, investigators have developed CDI adaptations and approaches that integrate parent reports with naturalistic language sampling. In the Maltese context, Gatt (2010) employed this combined approach to examine early expressive vocabulary (see Section 2.5.3).

In the present study, the notion of a noun bias is examined empirically through changes in the proportion of noun and verb tokens over time. The term is used descriptively to refer to a higher proportion of nouns in early productions, rather than as evidence of a universal or categorical developmental stage.

2.3.4. Bilingual child language acquisition

The acquisition of two or more languages from birth or in the early years of life has been researched extensively in the field of CLA and constitutes a key area of the present study. In what follows, I first describe bilingualism more broadly (Section 2.3.4.1), then turn attention to early bilingual acquisition (Section 2.3.4.2) and close off the section with an account of CS in both CDS and child speech (Section 2.3.4.3).

2.3.4.1. Bilingualism

The term *bilingualism* is here loosely used to refer to individuals learning or using two or more languages, or varieties, hence also encompassing multilinguals (Grosjean, 2012). Recent statistics indicate that bilingualism is a prevalent phenomenon worldwide, with more than half of the world's children growing up bilingual or in a bilingual community (Grosjean, 2010). For example, it is reported that 79% of young Europeans (ages 15–24) can have a conversation in a language other than their mother tongue (European Commission, 2024); in the United States, by contrast, only 22% of the nationwide population aged between 5;0–17 is considered bilingual (Kids Count Data Centre, 2023).

In this study, the approach to bilingualism follows that proposed by Baetens Beardsmore (1986), who states that bilingualism:

> must be able to account for the presence of at least two languages within one and the same speaker, remembering that ability in these two languages may or may not be equal, and that the way the two/more languages are used plays a highly significant role (p. 3).

This broad view encapsulates the multidimensional and complex nature of bilingualism, putting emphasis on *language fluency* and *language use*, which are intrinsically related and subject to change over time.[17] *Language fluency* refers to the degree of mastery in each language, which is many a time far from balanced (e.g., Cantone et al., 2008) and not restricted to productive competence, that is, an individual could be a receptive or productive user of the languages to be considered bilingual (e.g., Grosjean, 2012). *Language use* denotes the contextual and functional distribution of languages (e.g., Wei, 2007), for example Maltese in domestic settings and English in professional domains, or Maltese for oral production and English for written production. This bilingual's language use across domains is termed the *Complementarity Principle* by Grosjean (1997). It is different from, yet can be linked to, *diglossia* which is about a community's structured, social division of languages and language varieties (Ferguson, 1959).

2.3.4.2. Early childhood bilingualism

In the field of early childhood bilingualism, a distinction is made between bilingual first language acquisition (BFLA) and early second language acquisition (ESLA) (De Houwer, 2024b). The former encompasses the acquisition of two or more languages from birth or very early on (by around age 3;0), resulting in two L1s most likely acquired in a home environment. The latter refers to acquisition of two or more languages in the early years of life; however, as opposed to BFLA, in this case the child is exposed to the L2 after the L1 has started developing (by around 7 years of age). In ESLA, the context of acquisition is typically a childcare centre or an

[17] Hence, for instance, the body of research on *language attrition*. See, for example, Schmid (2011).

educational setting (e.g., De Houwer, 2024b; Serratrice, 2019). Therefore, BFLA is referred to as simultaneous bilingualism and ESLA is considered sequential or successive bilingualism (e.g., Wei, 2007). The delineation between BFLA and ESLA is not clearly defined and remains ambiguous, with disagreement among developmental psycholinguists regarding the precise criteria that define these boundaries (e.g., De Houwer, 2009; McLaughlin, 1985; Paradis et al., 2011).

Monolingual and bilingual infants can discriminate acoustically between different languages early on (e.g., Sundara, 2022; Vihman et al., 2007), and it has also been found that while newborns of monolingual mothers show preference towards the one language they are exposed to, newborns of bilingual mothers distinguish the two languages, yet exhibit no particular preference (Byers-Heinlein et al., 2010). As BFLA infants continue on their acquisition trajectory, this phonetic differentiation is reflected in their speech and, as some argue, in their babbling, too (e.g., Maneva & Genesee, 2001); however, research in this domain is still in its infancy.

De Houwer (2024a) describes the similar trajectories guiding bilingual and monolingual infants and points out two additional milestones specific to bilinguals. These major language development milestones, both those acquired by all children and those specific to bilinguals, are listed in Table 2.1 below and discussed in Section 2.3.1.

Table 2.1 *The Main Language Development Milestones and their Timing as Described in De Houwer (2024a)*

Babbling	**Monolingual and bilingual infants**
Understanding words	
First words	
Saying a total of 50 different words	
Combining words	
Understanding words in two languages	**Bilingual infants**
Showing evidence of separate grammatical systems	

Although the developmental trajectories of bilingual and monolingual children share many universal features, such as the stages of babbling, first words, and syntactic development, there are notable differences in pace, variability, and language-specific outcomes. While comparisons between bilingual and monolingual infants bring out these universal features and prevalent differences, CLA researchers should not fall back on "monolingualism as the default benchmark of comparison for bilingualism" (De Houwer, 2023, p. 354) and should be aware of the dangers which come with these comparisons, and which can lead to misconceptions about BFLA.

There are several misconceptions surrounding the BFLA trajectory, most of which are outlined in Genesee (2015). These misconceptions may have arisen from the monolingual bias in CLA research (e.g., Genesee, 2022; Kidd & Garcia, 2022) and have been debunked over the years. One such mistaken belief is that bilingualism in early childhood delays the infants' acquisition trajectory in comparison to their monolingual peers and results in incomplete competence in one or both languages.

As emphasised by Hoff (2015, p. 491), "it is important to note that bilingual children are not delayed in the rate at which they acquire linguistic knowledge in total". This is because bilingual children typically acquire a smaller vocabulary in each language when measured separately, especially in the early stages of development, compared to monolingual peers (e.g., Hoff et al., 2012; Pearson et al., 1993). However, when the total conceptual vocabulary across both languages is considered, bilinguals generally demonstrate comparable or greater lexical breadth than monolinguals. This indicates that bilingual infants distribute their vocabulary learning across two linguistic systems rather than showing an overall delay in language acquisition.

De Houwer (2024a, p. 17) goes a step further in stating that "in fact, monolinguals can be seen as being delayed in comparison to bilinguals when the total number of words understood is considered". While a slight lag has been observed between bilingual and monolingual infants when measured on each language independently, this lag diminishes with age (e.g., Hoff, 2015). In addition, several studies comparing bilingual and monolingual infants, including longitudinal studies centred on the age at which milestones are reached, found no difference at all between the two cohorts in achieving most language milestones (e.g., Höhle, 2020; Muszyńska et al., 2025).

Grammatical development in bilingual infants often parallels that of monolinguals, with variation depending on input quality and quantity (e.g., Thordardottir, 2015), among others. Studies have shown that input frequency and consistency[18] play a crucial role in shaping the acquisition of morphosyntactic features in bilingual infants (e.g., Paradis & Genesee, 1996). Additionally, cross-linguistic influence, whereby the two languages being acquired affect one another within a bilingual individual's linguistic repertoire, also plays a crucial role in BFLA of morphosyntax (e.g., Serratrice, 2013; van Dijk, 2022).[19]

Grammatical development in BFLA has been employed as a "source of evidence relevant to the one- versus two-system controversy" (Clark, 2024, p. 392), a deep-rooted debate on whether bilingual children have one system per language or one system encompassing both languages, with most studies in this domain centring on the acquisition of a) grammatical morphemes, and b) negation. Nowadays, as stated by Hoff (2015, p. 485), "the current consensus in the field is that children exposed to two languages build two systems from the start".

Furthermore, substantial IDs and variation are present among monolingual and bilingual infants alike; however, this variability tends to be more marked in bilingual populations, largely due to the heterogeneity of bilingual environments (e.g., Meir, 2023; Paradis, 2023). For example, in BFLA asymmetrical language development is frequently observed, whereby one language advances more rapidly

[18] See Section 2.2.2 for more on input.
[19] The interplay between typological similarity and cross-linguistic influence is often discussed in terms of *positive and negative transfer* and *psychotypology*, that is, the learner's perception of similarity between the two languages (Kellerman, 1983).

than the other across various domains of use. In some cases, this disparity is so pronounced that the child primarily communicates in only one language (e.g., De Hower, 2024b). Factors such as language dominance, sociolinguistic context, input balance, and exposure onset introduce variation in how bilingual infants acquire and use their languages (e.g., De Houwer, 2007; Unsworth, 2013). Paradis (2023) distributes these factors into two groups: child-internal and child-external factors, such as socioemotional wellbeing and parental proficiency in the language(s) addressed to the child, respectively. "Little is known about how different factors interact and collectively shape language acquisition outcomes in bilingual children with typical and atypical language development" and hence, more "multifaceted and interdisciplinary approaches attempting to model bilingual language development" are necessary (Meir, 2023, p. 828).

In conclusion, BFLA does not lead to long-term cognitive or linguistic disadvantages. Employing ID approaches to bilingual development accounts in the field "could shift us from a *bilingual deficit* to a *bilingual difference* ideology" (Paradis, 2023, p. 811). This will enable the field to reorient itself away from the pervasive monolingual bias and move beyond the "deficit notion" that bilingual children are two monolinguals in one (De Hower, 2023, p. 354), crucial not only from a theoretical standpoint, but also for parents, educators and clinicians in bilingual settings.

2.3.4.3. Code-switching in input and production

CS, that is, the alternation between two or more languages/language varieties within or across utterances is widespread among bilingual populations. Since the 1990s, the significant role of CS in BFLA, both in the input received and in the child's own speech, has garnered significant attention. This includes research on inter-sentential (across utterances) and intra-sentential (within utterances, often referred to as code-mixing) switches. Initially, research in the field of CLA interpreted CS as indicative of an inability to differentiate between languages, a transitory phase to be surpassed (e.g., Volterra & Taeschner, 1978). However, contemporary research highlights the complex multifaceted nature of BFLA through phenomena like CS (Lanza & Wei, 2023). This dynamic process is far from arbitrary and is, in fact, a strategic tool influenced by various factors, such as the social context in which children are immersed.

Literature on the effects of code-switched input on bilingual development is mixed, with different studies finding a) no association between the two (e.g., Place & Hoff, 2016), b) a negative correlation (e.g., Lipsky, 2013), and c) a positive correlation (e.g., Kremin et al., 2022). The studies observing a correlation between code-switched input and development often attribute specific features of CS to this correlation, be it positive or negative. For instance, Bail et al. (2015) found intra-sentential code-switched input specifically to be positively correlated to the vocabulary size of the 1;5–2;0 English–Spanish bilinguals tested. Kaushanskaya and Crespo (2019) found both a negative and a positive correlation, with the dependent variable being working memory, that is, the cohort with low levels of working memory had reduced language scores due to CS exposure and vice versa.

Kaushanskaya and Crespo (2019, p. 714) therefore suggest that the results provide a "possible reconciliation for the discrepant finding in the literature".

"Bilingual speakers can code-switch in multiple ways, to varying degrees, and for different reasons" (Bail et al., 2015, p. 1074). Both in CDS and child speech, the reasons for CS are numerous. Kremin et al.'s (2022) study on CS in the speech of parents to French–English bilingual infants reveals that the two major motivations for CS are to enhance the infants' comprehension and facilitate vocabulary acquisition. Contextual and situational factors, cultural identity and emotional expression, such as switching language depending on the formality of the setting or to express affection, also play a role in CDS CS practices.

The reasons for CS in child speech are also various and, while some are interrelated to CDS CS factors, others are communicative tools indicative of a bilingual developmental trajectory. CS in bilingual infants can be attributed to their developmental trajectory, and, in fact, research indicates that CS diminishes as children grow older. Studies show that CS is particularly high among 2–3-year-olds and that it decreases beyond this age (e.g., Gaskins et al., 2019; Redlinger & Park, 1980).

Many a time, the ability to code-switch allows bilingual children to convey meaning with greater precision and navigate lexical and syntactic gaps in either language (e.g., Nicoladis & Secco, 2000); gaps which do not necessarily arise from the unavailability of items but rather from insufficient entrenchment. Therefore, CS "is often the result of the faster activation of the other-language element, be it a word or a partially schematic construction, itself presumably the result of a higher degree of entrenchment for that element" (Quick et al., 2018, p. 405). Quick et al. (2018) calculated the MLU for three German–English bilingual infants between 2;0–4;0 and found that their MLU and syntactic complexity were highest in mixed utterances, allowing for successful communication, complex utterances and few fragments. These findings are corroborated by similar studies, such as Baird's (2024) longitudinal research on an Estonian–English bilingual child aged 2;0–5;0, which also reveals mixed utterances as the longest.

Another reason for CS in child speech is the exposure to switching itself although studies in this area show mixed results. From an early age, children develop an acute awareness of which languages are spoken by different individuals in their surroundings and are attuned to the statistical patterns of language use in their environments (Clark, 2024). Hence, CS in bilingual children often mirrors the linguistic input they receive (e.g., Comeau et al., 2003). The social context significantly influences children's language choices during interaction, as they tend to adopt similar switching patterns to those used by their interlocutors. (e.g., Byers-Heinlein, 2013; Yip & Matthews, 2016). For instance, Quick et al. (2019) found that 74% of the mixed utterances of their German–English bilingual infant cohort were based on patterns from parental input, highlighting the role of input in shaping CS. However, this is not always the case (e.g., Mishina-Mori, 2011). Gaskins et al. (2022) investigated bilingual children who are not exposed to switching in their input and pinpoint one aspect of input which drives the production of mixed utterances, this being the children's own prior speech. They argue for the importance of applying a bottom-up usage-based approach to research on CS in child speech.

Lastly, we turn our focus to the morphosyntactic structure of code-switched utterances in the speech of bilingual children, which is crucial to understanding how young learners manage and integrate multiple linguistic systems. Research demonstrates that bilingual children often adhere to the grammatical constraints of both languages when switching. For example, Paradis et al.'s (2000) study involving French–English bilingual toddlers reveals that their CS patterns conform to the syntactic rules of both languages, indicating an early ability to differentiate and apply distinct grammatical rules. Similarly, analyses of Swedish–Italian and Swedish–French bilingual children show that their mixed utterances maintained grammatical integrity, further supporting the notion of systematic CS (Bernardini & Schlyter, 2004).

Conflicting evidence, showing morpheme order and syntactic structure violation has also been put forth. One such study is Vihman's research (2018, p. 39) that focuses on the CS of two Estonian–English bilingual children and reveals that "the children's languages affect one another in more fundamental ways than lexical insertion, with examples of structural transfer across various domains". As stated by Vihman (2018), this complex and contradictory evidence calls for more detailed, methodological and representative data covering a diverse selection of language pairs being acquired simultaneously.

The pressing need for more nuanced and integrated theoretical approaches stemming from systematic and representative data is palpable in all areas of CLA (e.g., Kelly et al., 2015; Kidd & Garcia, 2022). It is hoped that this study on Maltese child-language will contribute to the broadening of languages and contexts under study, including CS research, since practically all Maltese children are brought up in bilingual environments. Against this backdrop, we now focus on the acquisition of morphological and syntactic structure.

2.4. Morphology and syntax in acquisition

"Languages vary to an extreme degree in their structures, and one of the major tasks of language acquisition research is to explain how children can learn any of these structures" (Stoll, 2015, p. 351). The following sections examine the strategies children employ and the developmental stages they traverse in acquiring the morphological and syntactic structures of their L1 (Sections 2.4.1 and 2.4.2, respectively). In both cases, the NP is foregrounded due to its central relevance to the present study (Sections 2.4.1.1 and 2.4.2.1). Finally, Section 2.4.3 outlines a selection of the myriads of methodologies for data collection and analysis commonly used in research on morphosyntactic development.

2.4.1. Acquiring grammatical morphemes

The acquisition of grammatical morphemes is a foundational aspect of early language development.[20] Brown's (1973) seminal work established an order of acquisition for English grammatical morphemes (e.g., present progressive *-ing*, plural *-s*, past tense *-ed*), suggesting that children follow relatively consistent developmental sequences. These sequences, as well as the pace of acquisition, are influenced by several factors such as the learning strategies of the child acquirer, input frequencies and typological differences (e.g., Slobin, 1985). For instance, L1

[20] For a detailed account of L1 acquisition of morphology and its developmental trajectory, see Ravid (2019).

learners of languages with rich morphological systems often show earlier and more accurate morpheme use (Berman, 1985). More work is necessary in this area as research on the acquisition of grammatical morphemes remains skewed towards English. Although studies on highly inflectional languages (e.g., Forshaw, 2021 on Murrinhpatha; Savičiūtė et al., 2018 on Lithuanian) and cross-linguistic studies (e.g., Dressler et al., 2019 on Estonian, Finnish, German, Greek, Lithuanian, Russian and Saami; Engelmann et al., 2019 on Finnish and Polish) are on the rise, they are still rather limited (e.g., Granlund et al., 2019).

Two competing routes dominate research on the acquisition of inflectional morphology: *single-route* and *dual-route* models.[21] Both mechanisms were initially based on English; however, over the years they have been substantiated by cross-linguistic data and are equally backed by ample evidence. The *single-route model* assumes that regular and irregular forms are acquired by one single underlying mechanism. Emergentist and usage-based language acquisition approaches gave rise to this model, also termed *connectionist model* since "learning consists of strengthening or weakening connections between pairs of units as a result of the changing nature of linguistic input" (Saiegh-Haddad et al., 2012, p. 1082). According to this theory, both monomorphemic and multimorphemic words are stored as wholes, and differences in frequency of tokens then lead to processing and, over the course of development, productivity (Bybee, 2001). Once children begin to acquire phonological schemas, they extend these schemas to new tokens by mapping the

[21] See Nakisa et al. (2000) for more on these competing theories.

structure of similar words already stored in their mental lexicon to novel words (e.g., Daelemans, 2002).

This model accounts for overgeneralisations of regular forms by arguing that, since acquisition is based on analogy and type-frequencies, irregular forms, most of which are of a lower type-frequency, are overgeneralised, therefore leading to productions such as *mouses* rather than *mice*, taking an example of English (ablaut) irregular plurals (e.g., Marchman, 1997; Racz et al., 2015). Frequency, together with other factors such as similarity, saliency, and consistency drive acquisition in single-route models as children gradually proceed to target-like productions (e.g., McClelland & Patterson, 2002).

Contrastingly, the *dual-route model* contends for a split in the processing mechanism by proposing that regular and irregular forms are processed separately by two distinct systems (e.g., Pinker, 1991). This model is also known as the *Words and Rules model* since it assumes that command of both the lexicon and grammar is necessary for the acquisition of inflectional morphology. Proponents of this model ascertain that regular forms are acquired by computing and applying inflectional rules to morphemes (stems and affixes) which are stored separately and form larger syntactic units by combining together. In contrast, irregular forms are memorised as wholes in the lexicon, with morphemes such as plural suffixes already fused in these lexical entries. The former process depends on procedural memory while the latter requires declarative associative memory (Ullman, 2002). According to this model, children make use of underlying rules or memorised forms (regular vs irregular morphology) rather than phonological schemas, as in single-route models.

The dual-route model also accounts for overgeneralisations by assuming that children extend underlying rules of regular forms to irregular forms too.

The issue that emerges in relation to the dual-route mechanism is that, while it is heavily based on the distinction between regular and irregular morphology, this distinction is in fact not as clear-cut as one might imagine,[22] even in a language with a rather straightforward morphology such as English (e.g., Albright & Hayes, 2003). As Nieder et al. (2020) argue, with reference to Maltese, in hybrid languages that combine concatenative and non-concatenative processes, the regular–irregular divide cannot be straightforwardly mapped onto morphological structure. Hence the question arises: How are concatenative and non-concatenative forms stored and computed by children acquiring Maltese inflectional morphology? Nieder et al.'s (2020) novel-word[23] plural production study with Maltese-speaking adults provides evidence for the single-route model rather than the dual-route model, while evidence from Arabic[24] acquisition studies is mixed (e.g., Masrai, 2016; Ravid & Farah, 2009).

Numerous additional theoretical frameworks, stemming from nativist and constructivist traditions, and beyond, have been proposed to account for morphological processing and development. One such framework is the *pre-/protomorphological approach*, addressing the gradual understanding and use of morphology, starting before the recognition of morphemes as distinct, meaningful

[22] See Seidenberg and McClelland (1989) on *quasiregularity*.
[23] Novel-word experiments (e.g., Zapf & Smith, 2007) are also known as nonce/nonsense-word tasks. See Section 2.4.3 for more on production studies and novel-word tasks.
[24] Arabic inflectional morphology is also characterised by concatenative and non-concatenative forms.

units (e.g., Bittner et al., 2003; Dressler & Karpf, 1995). The *pre-/protomorphological approach* draws on constructivist, cognitive and usage-based theories and emphasises gradual abstraction from input. Other approaches on morphological L1 processing, beyond nativism and constructivism, include Statistical Learning Models (e.g., Romberg & Saffran, 2010; Saffran et al., 1996), Optimality Theory approaches (e.g., Prince & Smolensky, 2004; Xu, 2011) and Social–Interactionist frameworks (e.g., Bruner, 1981; Lytle & Kuhl, 2017).

2.4.1.1. Nominal inflection

The acquisition of nominal categories, specifically number, gender, and case, has long been a central focus in morphological research within the field of CLA. In what follows, attention is given to the L1 acquisition of number and gender inflection on nouns. Case is excluded from this account since Maltese nouns do not inflect for case.

The acquisition of the plural

A substantial body of CLA research has concentrated on the development of plural morphology, with a significant portion focused specifically on the acquisition of English plural allomorphs (/s/, /z/, /ɪz/). Findings suggest that these allomorphs are acquired incrementally and that this variation appears to pose minimal difficulties for comprehension beyond 3;0 (e.g., Davies et al., 2017, 2019). One of the most influential studies in the domain of English plural acquisition is Berko Gleasons's experimental research (Berko, 1958), widely recognised for its use of the novel item 'wug'. The study's findings reveal that children under 5;0 produced the target inflected form 'wugs' /wʌgz/ at a rate of 76%, while children aged 5;0 and older achieved a success rate of 97%. These results provide early evidence of children's productive control over plural morphology. Further support for early sensitivity to morphological markers comes from studies involving younger infants. For instance, Kim and Sundara (2021) report that infants exposed to English demonstrate sensitivity to morphological suffixes as early as 0;6.

In contrast to the relatively fast-paced acquisition of plural morphology in English-speaking children, findings from classifier languages, where number is not obligatorily marked, paint a different picture. For example, children acquiring Japanese typically do not demonstrate consistent plural marking until around 6;0 (Nakano et al., 2009). However, this slower trajectory does not apply to the acquisition of all classifier languages. By way of example, Ahn and Snedeker (2022) report that Korean-speaking children exhibit adult-like plural marking by 4;0, despite Korean also being a classifier language. The authors attribute this earlier acquisition to two key factors: a) the plural marker in Korean appears more frequently in CDS, and b) plural marking in Korean is more consistent compared to languages such as Mandarin and Japanese.

Blount (1988), in a study of Luo-speaking children in Kenya aged 3;0–14, examined the acquisition of plural morphology using both real nouns and novel forms. Owing to the high degree of morphological complexity in the Luo plural system, even the oldest participants achieved only an 86% accuracy rate on the novel-noun task. Adult speakers, however, demonstrated full mastery, producing correct forms in 100% of instances. These findings contrast sharply with those of Berko Gleason (Berko, 1958), where English-speaking children showed high rates of correct pluralisation at much younger ages, as discussed above. A different picture emerges in Croatian: research on nominal agreement in this morphologically rich and highly inflected language suggests that "rather than being an obstacle, morphological complexity proves to be a trigger for the early development of inflection" (Kovačević et al., 2009, p. 174).

Semitic languages offer additional insight into the acquisition of plural morphology due to the coexistence of concatenative and non-concatenative processes. This dual system provides a useful testing ground for examining how children acquire and generalise morphologically distinct patterns that differ in transparency, complexity, and productivity. A growing body of research across Arabic dialects has explored this issue, with a particular focus on the developmental relationship between sound and broken plurals (SP; BP).

Across studies, a consistent asymmetry emerges in which SPs, especially the feminine sound plural (FSP), are acquired earlier and used more productively than BPs. For example, Saiegh-Haddad et al. (2012) investigated plural acquisition in Palestinian Arabic using repetition, structured production, and semi-natural production tasks with 36 children aged 3;0–8. Their findings show that the FSP is acquired earlier and follows a shorter developmental trajectory than BPs. Children frequently extended the FSP beyond its target contexts, indicating its role as a default strategy, while BPs remained less stable and more variable. Importantly, performance on BPs was sensitive to both lexical familiarity and the frequency of plural patterns, suggesting a stronger reliance on item-based knowledge for non-concatenative morphology.

Similar patterns have been reported in other Arabic varieties. In a picture-based elicited production study with 20 Jordanian Arabic-speaking children aged 3;0–5;0, Mashaqba et al. (2020) found that children relied heavily on the FSP as a default form, often overgeneralising it to contexts requiring BPs. Notably, this tendency emerged despite the high frequency of BPs in the input, with children's productions instead shaped more strongly by productivity and predictability than by token frequency alone. In addition, Albirini (2015) examined plural acquisition in 60 Jordanian Arabic-speaking children aged 3;0–8 using both real-word and novel-word pluralisation tasks. The results confirm that SPs are acquired earlier and generalised more readily, particularly in novel-word contexts, supporting their status as productive morphological markers. In contrast, BPs showed greater variability and were more dependent on frequency and lexical familiarity, with predictability playing an increasingly important role at later stages of development.

Evidence from Kuwaiti Arabic further supports the prominent role of SPs in early acquisition. Abdalla et al. (2013) examined plural production in TD children, children with specific language impairment, and adults using both real and novel-word stimuli. TD children showed higher overall accuracy but displayed a strong tendency to overgeneralise the FSP across contexts, including those requiring masculine SPs or BPs. This pattern of overgeneralisation was also observed in novel-word tasks, providing clear evidence of the productivity of the SP system.

Converging evidence also comes from studies examining different learner populations and the broader structure of the Arabic plural system. Albirini and Benmamoun (2014), comparing L1, L2, and heritage speakers using picture naming and plural formation tasks, found that SPs are generally acquired earlier and processed more successfully than BPs, with difficulties in BP morphology persisting even in more advanced speakers. Work addressing the learnability of the Arabic plural system further supports these acquisition findings. Using computational modelling, Dawdy-Hesterberg and Pierrehumbert (2014) showed that Arabic plural formation depends heavily on abstract phonological templates and analogical generalisation, highlighting the structural complexity of the BP system and the absence of a fully predictive rule.

Findings from studies on the acquisition of the Maltese plural system, which displays SP and BP forms, align with those described above for different varieties of Arabic (e.g., Cachia, 1982; Zammit, 2020). Maltese presents a slightly different case to the Arabic languages mentioned above because of its gender-neutral plural morphology (see Section 2.5.1). Nevertheless, studies involving both preschool and school-aged children have shown a strong preference for overgeneralising SP suffixes to nouns that typically take BP patterns and to nouns that conventionally take a different SP suffix (e.g., Vella, 2025). Research on error patterns in the acquisition of the plural indicates that overgeneralisation is a common developmental phase. Notably, nouns with low input frequency tend to exhibit higher rates of overgeneralisation (e.g., Maslen et al., 2004).

A particularly intriguing dimension of plural acquisition arises in languages that mark dual number, such as Hebrew, Scottish and Warlpiri (Australia), and, in a reduced form, also Maltese (see Section 2.5.1). Like Maltese, these languages feature a dedicated morphological form to express the notion of 'two of an object', alongside singular and plural categories. Almoammer et al. (2013) found that children acquiring Slovenian and Saudi Arabic grasp the concept of 'two' more rapidly than their peers learning languages without a dual form, such as English. These findings suggest that morphological cues embedded in the input language facilitate the acquisition of specific numerical concepts.

The acquisition of gender inflection

Gender often plays a fundamental role in grammatical agreement both within and beyond the clause and therefore presents the child with a complex system of morphosyntactic dependencies that must be mastered for target-like proficiency. Cross-linguistic studies have shown that children acquiring gendered languages are sensitive to gender distinctions from an early age, although the timing and accuracy of acquisition can vary considerably, depending on factors such as the transparency of the gender system, the consistency of gender cues, and the morphological richness of the language (e.g., Stephany & Voeikova, 2009).

In Romance languages, such as Spanish, French, and Italian, grammatical gender is typically binary (masculine vs feminine) and is largely marked on determiners, nouns, and adjectives. These languages tend to offer relatively transparent gender assignment rules based on phonological cues (e.g., Spanish: *-o* for masculine, *-a* for feminine). Work on the L1 acquisition of inflection in such languages reveals that children produce target-like forms of grammatical features as early as 1;6, including the use of nouns and determiners in both masculine and feminine (Parisse, 2023). Research shows that children acquiring Spanish, for instance, begin to use gendered determiners correctly by 2;0, with overgeneralisation errors decreasing steadily with age (e.g., Pérez-Pereira, 1991; Socarrás, 2011; Schnell de Acedo, 1994). Similarly, French-speaking children have been shown to use gender-marked articles appropriately in the second year of life, although errors persist in contexts involving phonological ambiguity (e.g., Clark, 1985).

German presents a more complex system, with a three-gender distinction (masculine, feminine, neuter) and less predictable assignment rules. Despite these complexities, children acquiring German show sensitivity to gender distinctions by 2;0; however, full mastery of gender agreement, particularly with less frequent or irregular nouns, may not occur until age 6;0 or later (e.g., Mills, 1986; Szagun et al., 2007). Error patterns in German acquisition often involve defaulting to the most frequent gender, typically masculine, reflecting both input frequency effects and morphological salience. Research on Icelandic, also a Germanic language with a complex gender and case system, reveals that children's knowledge of grammatical gender facilitates their understanding of the plural formation, and that productivity in these two grammatical features is correlated, highlighting the interdependence of morphological categories (Björnsdóttir, 2023). In Norwegian, also Germanic, the feminine gender appears to be more vulnerable to attrition and change, both in language acquisition and in dialectal variation, with research in this area suggesting that children may acquire masculine and neuter genders more robustly than feminine gender (Lohndal & Westergaard, 2021).

Slavic languages, including Russian, Croatian, and Polish, feature rich gender systems and extensive nominal inflection. In Croatian, for example, gender is intertwined with case and number, yielding complex paradigms. Nevertheless, CLA research indicates that Croatian-speaking children acquire gender distinctions relatively early, with gender-inflected forms emerging during the second year of life and showing high levels of accuracy by age 3;0 (Kovačević et al., 2009). Longitudinal L1 research on Polish, reveals that noun inflection is in place by 2;0 and that, despite the arbitrary distribution of endings of masculine genitive singular nouns, Polish-speaking children perform few errors in this domain (Dąbrowska, 2001).

In Semitic languages such as Hebrew and Modern Standard Arabic, gender is overtly marked on verbs, pronouns, and numerals, in addition to nouns, resulting in a dense morphosyntactic environment. Studies on Hebrew L1 acquisition (e.g., Berman, 1985) show that children as young as 2;0 produce gender-appropriate verb forms, although mastery of agreement in more complex syntactic constructions continues to develop well into the preschool years. In addition, research on Egyptian Arabic, for instance, suggests that children also acquire the core gender distinctions early while continuing to exhibit variability in less frequent or irregular contexts (e.g., Omar, 2007). The limited research on L1 acquisition of gender inflection and agreement in Maltese includes Coppini (2002), who emphasises that these phenomena are present at age 3;0, the youngest age tested, and suggests that the masculine form of adjectives (generally considered 'unmarked') is mastered before the feminine singular and plural forms.

A growing body of research underscores the influence of input frequency and morphological transparency on the acquisition of gender and number inflection. For instance, overall, it seems that languages with consistent and phonologically salient gender markers tend to support earlier and more accurate acquisition (e.g., Tsimpli & Hulk, 2013). Conversely, languages with opaque or arbitrary gender systems, such as Russian or German, often show a prolonged developmental trajectory. Nonetheless, children appear to use distributional cues in the input to bootstrap their acquisition of gender and number categories (see Höhle (2009) for an overview on bootstrapping in CLA). Furthermore, error patterns across languages indicate a reliance on default strategies in the face of ambiguity. For example, as stated above, defaulting to the masculine gender has been widely reported cross-linguistically, reflecting both the statistical dominance of the masculine form in the input and its cognitive function as an unmarked category in the grammatical system (e.g., Levy, 1983; Mills, 1986).

Another flourishing area of research concerns morphological patterns in bilingual L1 acquisition and the cross-linguistic influences that bilingual children experience, particularly in gender and number assignment. For example, English-speaking children acquiring Maltese either as their L2 or simultaneously with English may transfer plural formation strategies from one language to the other, influencing both their processing and production in each language (e.g., Vella, 2020). CLA research in this area has focused on error analysis (e.g., Vaks & Vihman, 2025), the impact of age of acquisition (e.g., Bosch et al., 2019), the effect of short-term versus working memory (e.g., Blom et al., 2021), and various theoretical approaches to morphological development, such as the whole-word approach (e.g., Vihman, 1982). Collectively, these findings emphasise that bilingual acquisition of inflection is neither uniformly delayed nor monolithic; rather, it follows a complex and nuanced developmental trajectory shaped by factors such as language-specific input, structural characteristics and cognitive resources.[25]

While children across language backgrounds demonstrate early sensitivity to gender and number inflection, the path to full mastery varies depending on several features touched upon in this section. Importantly, cross-linguistic research reveals that rather than hindering acquisition, morphological complexity can often act as a catalyst, providing children with abundant and salient cues from which to extract grammatical rules and generalisations. As highlighted by Björnsdóttir (2021) and Stoll (2015), among others, more cross-linguistic research is necessary to better understand L1 comprehension, productivity and proficiency of inflection, and morphology more generally, in linguistically diverse systems through "a more comparative stance", focusing on both naturalistic and experimental data collection methods (Stoll, 2015, p. 373).

2.4.2. Acquiring syntactic structures

Mastery of syntax is essential for understanding the hierarchical and relational organisation of linguistic elements and allows language users to navigate and generate complex linguistic constructions with precision.[26] Beyond its linguistic function, recent empirical research suggests that syntactic competence is closely tied to broader cognitive development; interwoven with capacities such as abstract

[25] For more on bilingual language acquisition, refer to Section 2.3.4.
[26] For a detailed account of L1 acquisition of syntax, see Thornton (2016).

thinking and problem-solving. For instance, White et al. (2017) found that preschoolers' syntactic skills are tightly coupled with working memory and inhibitory control. Similarly, Thothathiri et al. (2025) demonstrate that executive functioning promotes successful comprehension and acquisition of non-canonical structures, substantiating the view that syntactic competence supports and is supported by domain-general cognitive processes.

Eisenbeiß (2015, pp. 1792–3) groups the controversies surrounding the acquisition of syntax into three main issues: "(i) the role of nature and nurture […]; (ii) the time course of syntactic development; (iii) age effects and differences between monolingual and bilingual acquisition". The *nature versus nurture dichotomy*, mainly pivoted around nativist and constructivist approaches, has been discussed in Sections 2.1.2. and 2.1.3. While nativist approaches emphasise innate syntactic knowledge, constructivist and usage-based theories highlight the role of input and learning mechanisms. These approaches, and others, such as connectionist models and Optimality Theory, strive to explain the way in which children come to string words together within the first few years of life, taking into account the various factors involved.

The *time course* of syntactic acquisition refers to the developmental trajectory outlined in Section 2.3.1. Children typically begin producing recognisable syntactic structures during the second year of life, with early two-word combinations emerging at around 1;6–2;0 (e.g., Brown, 1973). These combinations often display target-appropriate word order even in the absence of function words, indicating early sensitivity to linear relations. By 3;0, most children can produce complex utterances, although the acquisition of certain constructions – such as the English passive – extends into the school years (e.g., de Villiers, 1985; Diessel & Tomasello, 2005). As Ambridge (2020, p. 527) notes, "if there is one thing on which all mainstream accounts of child language acquisition agree, it is that by – at the latest – around 3;0 (Tomasello, 2000, 2003) children have abstract knowledge of word order." However, Ambridge (2020, p. 513) challenges this assumption, proposing "a pure exemplar-learning based account" in which language acquisition is grounded in specific usage experiences, with grammatical knowledge arising through analogy, memory, and pattern recognition across rich, context-dependent exemplars.

Lastly, the issue of *age effects and differences between monolingual and bilingual acquisition* focuses on the interdependence of age and language acquisition, both L1 (monolingual/bilingual) and L2 acquisition. Research into syntactic acquisition in bilingual contexts and L2 acquisition constitutes a well-established and rapidly growing field within applied linguistics. The effect of CS and mixing and the way in which the two or more systems are handled are often at the heart of empirical and theoretical investigation (see Section 2.3.4.3). In addition, one of the main arguments which regularly surfaces in debates about the effect of age on acquisition is the *critical period hypothesis*, which states that there is a biologically determined optimal period for acquiring language, particularly grammar, and which predicts that full native-like language mastery is unlikely to be achieved when acquisition occurs later (Lenneberg, 1967). Despite its influence, this hypothesis has not gone unchallenged, with various studies offering considerable critique and counterevidence (e.g., Scovel, 2000).

Cross-linguistic research on the L1 acquisition of syntax reveals both universals and variations across different contexts, in line with other domains of acquisition. The specific trajectories and difficulties encountered vary depending on several factors, both language specific features, such as pro-drop, case marking and word order flexibility, as well as factors linked to input, cognitive abilities and sociocultural context (e.g., Guasti, 2002; Kidd & Arciuli, 2016; Vasilyeva et al., 2008). Findings emerging from cross-linguistic studies in this area highlight the importance of integrating typological diversity and IDs into theories of syntactic development. Syntactic acquisition studies also feature in the realm of neurocognition which attempts to identify the neural correlates of syntactic processing. Event-related potential studies show that infants as young as 1;0 exhibit neural responses to syntactic violations, suggesting early sensitivity to structural regularities (see, for example, Conboy et al., 2008; Friederici, 2005). Such cross-linguistic and neurocognitive findings continue to refine our understanding on how children acquire syntax, suggesting a nuanced interplay between universality and diversity in language development.

2.4.2.1. Nominal agreement

Nominal agreement, where features such as gender, number, or case are shared across elements within the NP, is a core component of morphosyntax in many languages. Agreement in the context of morphosyntactic acquisition has often been analysed within frameworks such as UG, minimalism, or Statistical Learning Models. As stated by Demuth and Weschler (2012, p. 68), "unfortunately, there has been relatively little research on agreement [within acquisition] per se, making it difficult to draw definitive conclusions". Although research has expanded over the past decade or so, this statement still holds. The subsequent paragraphs consider cross-linguistic evidence from the literature on L1 acquisition of nominal agreement.

In the domain of gender assignment and agreement, particularly in languages that offer both semantic and phonological cues, some scholars argue that children initially acquire gender through nouns that denote entities with natural gender, later generalising these patterns to nouns where gender is not semantically motivated. For example, Mulford (1985) found that Icelandic-speaking children aged 4;0–8 initially rely on semantic cues to assign grammatical gender to unfamiliar referents. Other researchers argue that children are sensitive to phonological cues when acquiring grammatical gender, even in languages where these cues are not consistently transparent. Supporting this view, Culbertson et al. (2019), in a study using an artificial language learning paradigm with 6–7-year-olds, found that children exhibited a bias toward phonological cues.

Like Maltese, Spanish is a particularly suitable language for investigating nominal agreement as the majority of nouns and their modifiers display explicit morphological markings for both gender and number. Socarrás (2011) conducted a cross-sectional study with three monolingual Spanish-speaking children between 2;0–3;0, focusing on gender and number agreement. Within a Minimalist Framework, she documented progression of the acquisition of determiner phrases from bare nominals to proto-determiners to full determiner phrases, showing graded mastery of gender and number marking (Socarrás, 2011). Research on French, which exhibits a nominal agreement system similar to that of Spanish, indicates that by 1;6, French-speaking children are already sensitive to number agreement patterns (e.g., Nazzi et al., 2011), and by 2;6, they demonstrate a more thorough understanding of gender assignment and agreement (e.g., Cyr & Shi, 2013). Typically, gender agreement on determiners is acquired earlier, while agreement on adjectives develops more gradually, with errors persisting until around 5;0 (e.g., Royle & Valois, 2010). Notably, in the case of certain written gender markers, omissions may persist until approximately age 15 (e.g., Bosse et al., 2021).

Research on Polish-speaking children indicates that grammatical gender acquisition is typically mastered in the second year of life (e.g., Smoczyńska, 1985). Similarly, studies on German L1 learners show a notable reduction in gender assignment errors by 3;0 (e.g., Szagun et al., 2007). Longitudinal investigations into NP acquisition in a German–Polish bilingual child between the ages of 1;6–4;0 reveal evidence of cross-linguistic influence, overgeneralisation, and differing rates of morphosyntactic development across the two languages (Jachimek, 2022).

More recently, Długosz and Brown-Bousfield (2025) examined the acquisition of grammatical gender assignment and agreement in German as an additional language among L1 Polish and L1 English speakers. Their findings show that the L1 Polish cohort, whose native language features grammatical gender (unlike English), outperformed the L1 English cohort on gender-related tasks in German. The advantage observed in the L1 Polish cohort is attributed to the presence of grammatical gender in their native language, which eases the acquisition of gender in subsequent languages. Such findings are particularly relevant to the present study, as all participants are bilingual speakers of Maltese and English; languages that differ in terms of gender marking, with Maltese encoding grammatical gender and English lacking such a system.

Demuth and Weschler (2012) conducted research on the L1 acquisition of Sesotho nominal agreement, a Bantu language with a complex gender system.[27] Upon examination of nominal agreement in the spontaneous speech of Sesotho-speaking children aged between 2;0–3;0, they found that participants have a strong command of the system by 2;6. The findings also indicate that a large part of the nominal agreement errors detected stem from phonological underspecification rather than morphosyntactic issues. Demuth and Weschler (2012, p. 86) therefore state that "although Bantu languages have many more nominal gender classes than most other languages, the phonological transparency, regularity and pervasiveness of the system appear to facilitate the early acquisition of nominal agreement in these languages".

The studies discussed here, along with numerous others, underscore nominal agreement as a locus of morphosyntactic complexity, making it amenable to analysis from diverse theoretical perspectives, including input-driven Statistical Learning Models and the Minimalist Framework. Comparative studies of monolingual/bilingual L1 (e.g., Molnar et al., 2021) and L2 learners (e.g., Cheng et al., 2022) of nominal agreement systems show that while monolingual L1 learners typically acquire core agreement by around age 3;0, bilingual and L2 learners often display systematic difficulties, resolved over time. Despite growing interest,

[27] Unlike Indo-European languages that have two/three genders, Bantu languages have up to 23 (Demuth & Weschler, 2012).

research on nominal agreement remains relatively limited (e.g., Ayoun, 2022; Demuth & Weschler, 2012). There is therefore a clear need for further investigation, particularly through cross-linguistic, longitudinal studies involving both monolingual and bilingual L1 speakers of typologically diverse languages, as well as L2 learners in varied learning contexts. Additionally, experimental approaches, such as eye-tracking studies using novel items, offer supplementary avenues for probing the underlying mechanisms of agreement acquisition. Such approaches used to study morphosyntactic acquisition are described in the following section.

2.4.3. Approaches to data collection and analysis

Collecting L1 morphosyntactic data

Morphosyntactic research in CLA employs a rich methodological toolkit from corpus analysis to cutting-edge neurocognitive experimentation, underpinned by robust qualitative and quantitative frameworks. In what follows, I survey a few methods of data collection adopted in the field and key analytical approaches in relation to the L1 acquisition of morphosyntax.

As noted by Eisenbeiß (2015, p. 5), "acquisition researchers have developed a broad range of methods for (i) naturalistic sampling, (ii) experiments and (iii) semi-structured elicitation". In *naturalistic sampling*, participants are typically recorded, either knowingly or unknowingly, while interacting with familiar interlocutors such as primary caregivers and siblings, and engaging in routine activities like pretend play or assisting with household tasks. The role of the researcher(s) can vary depending on the design of the study: in some cases, they act as interlocutors; in others, they monitor recording equipment or are absent altogether, having previously instructed adult participants on how to conduct and record the sessions.

Naturalistic sampling often yields child language corpora, which are central to L1 acquisition research. One of the most widely used resources is the Child Language Data Exchange System (CHILDES), which provides access to corpora across numerous languages, including both longitudinal and cross-sectional datasets. These corpora are frequently accompanied by tools for automatic analysis, such as morphosyntactic tagging and frequency profiling.

While naturalistic methods are valued for their high ecological validity (e.g., Eisenbeiß, 2010) and are also cost-effective since they can be used multiple times to study different aspects of acquisition (e.g., Ambridge & Rowland, 2013), they also present logistical challenges. Chief among these is the need for time-intensive manual transcription and annotation, as current automatic speech recognition technologies for child language, particularly in under-resourced languages, remain in early stages of development. Furthermore, although naturalistic studies are considered "an ideal starting point" for language acquisition research (Eisenbeiß, 2010, p. 12), they are not always optimal for capturing the full range of morphosyntactic features because certain linguistic constructions may simply not occur in the recorded data. Moreover, the variability inherent in spontaneous interaction can hinder comparability across participants (e.g., Brown, 1973). For studies aiming to investigate specific grammatical phenomena in a controlled and replicable manner, *experiments* and *semi-structured elicitation* methods are more appropriate.

Experimental approaches encompass a range of methodologies, including production tasks, grammaticality judgement tests, comprehension assessments, and online processing measures such as eye-tracking (e.g., Baron et al., 2023; Kenanidis et al., 2021; Valleau et al., 2018).[28] In addition, experiments also include neurocognitive techniques, such as event-related potential studies and functional magnetic resonance imaging, which have been employed to examine real-time brain responses to morphosyntactic stimuli and to map neural correlates of linguistic processing (e.g., Rispens & Krikhaar, 2010; Vannest et al., 2009). These methodologies offer insight into implicit processing dynamics and can reveal neurocognitive mechanisms underlying language acquisition.

Importantly, many experimental designs, particularly comprehension-based tasks and eye-tracking paradigms, are suitable for use with preverbal infants, thereby allowing researchers to investigate early stages of development before the onset of productive language (e.g., Syrnyk & Meints, 2017). In studies involving verbal infants, production tasks are commonly employed. Production tasks have been used in a multitude of languages to study various CLA phenomena, such as the acquisition of the plural in both well-studied languages like English (e.g., Clark &

[28] For detailed overviews on experimental methods refer to Blom and Unsworth (2010), and Ambridge and Rowland (2013).

Nikitina, 2009) and German (e.g., Laaha et al., 2006), and relatively understudied languages like Jordanian Arabic (e.g., Mashaqba et al., 2020).

Novel-word tasks are a type of production study frequently utilised with both children and adults to investigate various morphosyntactic features. Such tasks became popular following Berko Gleason's (1958) pivotal work and have remained central in the field of CLA ever since, with such experiments being carried out in a host of languages, varieties and dialects such as Australian-English (e.g., Davies et al., 2020), Kuwaiti-Arabic (e.g., Abdalla et al., 2013), Japanese (e.g., Matsuo et al., 2012), Estonian (e.g., Vihman et al., 2021) and Hungarian (e.g., Lukács et al., 2010). Both receptive and production tasks, including a novel-word task, are incorporated in this study (see Section 3.3.2).[29]

In semi-structured elicitation methods "the communicative situation is as natural as possible, but games, pictures, videos, etc. are used to encourage the production of rich and comparable samples of spoken speech" (Eisenbeiß, 2009, p. 2). As such, these methods represent a compromise between naturalistic data collection and fully controlled experimental designs. It is a common practice in CLA research to adapt classic games in such a way as to elicit particular constructions, especially when a) working with very young children, b) aiming for as natural as possible elicited-production settings, and c) developing elicitation strategies in longitudinal studies (Eisenbeiß, 2009).[30] In the present study a bingo game is used

[29] Receptive tasks are defined as those requiring behavioural responses to verbal prompts, while production tasks require overt verbal output.
[30] See, for example, Saiegh-Haddad et al.'s (2012) use of a Lotto game in a study on the acquisition of noun plurals in Palestinian Arabic and Eisenbeiß's (2009) description of a picture-pairing task similar to the classic Memory® game, employed as a semi-structured elicitation task with children as young as 2;0.

with the older child participants as well as a spot-the-difference activity (see Section 3.3.2).

Furthermore, narrative visual stimuli constitute a well-established elicitation technique in CLA and are also employed in the present study (see Section 3.3.2). Some studies opt for popular picture books (e.g., Berman & Slobin's (1994) seminal work with the *Frog Story*), while others develop their own image sequences (e.g., Kauschke et al., 2024) or books (Khan et al., 2016). Two notable resources employed extensively in the field are Mayer's (1969) *Frog Story* (see below) and the Multilingual Assessment Instrument for Narratives (MAIN; Gagarina et al., 2012, 2019).[31]

Additionally, semi-structured elicitation tasks often employ child-appropriate fictional characters and toys. Puppets, for instance, have a long history of association with CLA studies, specifically in elicitation settings (see, Ambridge & Rowland, 2013, among others). The incorporation of puppets is known to effectively enable a more playful and engaging atmosphere, facilitate active participation, promote socio-emotional communication and enhance compliance. In addition, "puppetry reduces the asymmetry between children and adult experimenters, boosting children's self-confidence" (Stengelin et al., 2023, p. 1119). Countless studies in the field have utilised puppets to examine various aspects of children's language development, from morphology (e.g., Keydeniers et al., 2023; Meyer,

[31] MAIN is designed to evaluate narrative comprehension and production in children learning one/more languages, using 4 picture-based narrative tasks centred around animals. This narrative tool has been adapted to over 90 languages (access website: https://main.leibniz-zas.de/en/worldwide-network/), including Maltese (Gagarina et al., 2020; Gatt & Borg Cutajar, 2023). It would be interesting to implement this tool in future MFLA studies investigating narrative skills.

2019) and syntax (Socarrás, 2011) to pragmatics (e.g., Özçalışkan, 2007) and sociolinguistic skills (e.g., Anderson, 1990). The present research employs a dog puppet throughout the recording sessions to engage participants and facilitate language production (see Section 3.3.2).

Both experimental methods and semi-structured elicitation methods are not without limitations. Experimental methods, in particular neurocognitive approaches, often require complex and costly equipment, and their implementation can be challenging when working with young children, who may not consistently comply with task demands. Moreover, while these methods provide high levels of experimental control, they may lack the ecological validity associated with naturalistic observation. Semi-structured elicitation tasks, while more ecologically valid, still lack the spontaneity of naturalistic samples and may be vulnerable to interviewer influence. Also, both experimental and semi-structured elicitation studies necessitate rigorous research designs and extensive preparatory work, including training experimenters, prior to the commencement of fieldwork.

Analysing L1 morphosyntactic data

Both quantitative and qualitative methods are vital in contemporary research on morphosyntactic development in CLA, each contributing in distinct yet complementary ways to the study of linguistic complexity, developmental variation, and contextual sensitivity. Quantitative approaches, particularly those based on corpus analysis, allow for large-scale, systematic investigation of morphosyntactic patterns. For instance, the CHILDES database provides both richly annotated corpora of naturalistic child–caregiver interactions and analytical tools such as CLAN, which facilitates morphosyntactic tagging and the extraction of quantitative metrics, including MLU (see Section 3.6). Liu and MacWhinney (2024) have recently incorporated the Universal Dependencies Framework across CHILDES corpora in 27 languages, offering a cross-linguistically consistent approach to morphosyntactic annotation and enabling more rigorous comparative research. Tools of this sort have yet to be adapted for Maltese before one can hope to be able to chart morphosyntactic acquisition trajectories in novel empirical settings.

The field has also benefited from the integration of natural language processing (NLP) techniques. Although automated tools such as POS taggers have been present for some time (e.g., Parisse & Le Normand, 2000, in their study of early French morphosyntax), recent advances in artificial intelligence and machine learning have led to the development of state-of-the-art NLP models capable of processing vast amounts of both CDS and child speech data (Nikolaus et al., 2024). These models hold promise for significantly scaling up morphosyntactic analysis as well as other areas of CLA research.

In parallel, advanced statistical modelling techniques have become increasingly central to morphosyntactic acquisition research. Mixed-effects models are now standard for analysing language acquisition data, as they accommodate individual-level variation and account for repeated measures (e.g., Baayen et al. 2008). For example, Blom et al. (2013) employed mixed-effects logistic regression to model Dutch children's verb morphology production, showing how lexical frequency interacts with age while accounting for child-specific variability.

Similarly, Bayesian models, such as those used by Frank et al. (2009) in studies of probabilistic word learning, allow researchers to formally model uncertainty and incorporate prior knowledge. The integration of both frameworks (i.e., mixed-effects models and Bayesian models) in Bayesian mixed-effects models has proven particularly useful in cross-linguistic and developmental research where data complexity and sparsity often challenge traditional approaches. These methods enhance the statistical power, flexibility, and interpretive depth of morphosyntactic analysis.

Crucially, while quantitative approaches offer generalisability and replicability, qualitative methods remain indispensable for capturing rich contextual and functional dimensions of morphosyntactic development (e.g., Dörnyei, 2007). Through detailed transcription and close analysis of elicited and spontaneous speech, qualitative research sheds light on the discourse functions of grammatical forms, children's strategic use of syntactic constructions and error analysis, among others. Such approaches are particularly valuable in under-documented languages for which existing annotation tools and statistical baselines are lacking.

As stressed by Ambridge and Rowland (2013), the field of CLA should strive towards methodological and analytical sophistication, enhancing the reliability and interpretive power of findings related to the mechanisms underpinning language acquisition trajectories. The future of CLA research lies in the integration of approaches referred to above, combining the rigour and scale of quantitative methods with the interpretive value of qualitative analysis. This methodological and analytical pluralism enhances the reliability, ecological validity, and theoretical richness of findings concerning the developmental mechanisms underlying L1 acquisition of morphosyntax.

2.5. Mind the gap

This section situates this study within its linguistic context beginning with an overview of the Maltese language (Section 2.5.1). The aim is not to offer an exhaustive description of the language but to highlight the features most relevant to the analyses reported on in Chapter 5, focusing on inflectional marking on nouns and adjectives (Section 2.5.1.1) and NP-internal agreement (Section 2.5.1.2),[32] with a final note on system complexity and its relevance to acquisition (Section 2.5.1.3). This is followed by a description of the sociolinguistic dynamics in Malta (Section 2.5.2), considering both home (Section 2.5.2.1) and childcare (Section 2.5.2.2) environments. Section 2.5.3 reviews existing research on Maltese child language, leading to the identification of the research gap outlined in Section 2.5.4, while Section 2.5.5 formulates the research questions that underpin this study.

2.5.1. Maltese

Maltese is the national language of Malta[33] and one of the three official languages on the Islands, alongside English and Maltese Sign Language. It is the only Semitic language written in Latin script, belonging to the South Arabic branch of Central

[32] Predicative adjectives are also included in the analyses, primarily due to the overall scarcity of adjectives and adjectival-function words in the naturalistic data, and because their agreement behaviour is parallel to that of attributive adjectives (see Section 5.2).
[33] Unless otherwise stated, throughout the study Malta is being used to refer to the Maltese archipelago, consisting of Malta, Gozo and Comino.

Semitic (Fabri, 2010). Maltese evolved from Siculo-Arabic, brought to Malta by Arabic invaders in 1048AD (e.g., Brincat, 2011; Spagnol, 2011) and exhibits similar traits to Maghrebi Arabic (particularly in its morphosyntax) and Levantine (Eastern) dialects (particularly in its phonology and lexis) (Vella, 2012).

Over the years, Maltese underwent a process of Latinisation, being extensively influenced by Sicilian and Italian, and was later also influenced by English. These influences infiltrated Maltese at the lexical, morphosyntactic and phonological levels, to different degrees, resulting in a language constructed through "the melding of different elements from different sources [...] into a separate entity" (Vella, 2012, p. 532). Consequently, Arabic is described as the stratum, Sicilian and Italian the superstratum and English the adstratum (Fabri, 2010). Comrie's (2009) typological study, comparing Maltese to Cairene Arabic and Spanish, confirms the closeness of Maltese to Semitic through shared characteristics. Comrie (2009, p. 10) states that "original genealogical affiliation still wins over the effects of rather intense language contact, although these latter effects are nonetheless clearly discernible".

Although Malta gained its independence in 1964 and is not subject anymore to language imposition by foreign rulers, it is still undergoing rapid changes in particular at the lexical level through borrowing, especially from English. Comrie and Spagnol (2016) classify Maltese as a *high borrower* with a total of 35.1% of loanwords. Apart from new loanwords which continue to enter, neologisms also expand and enrich the Maltese lexicon. Such neologisms include examples like *stessu* ('selfie') and *borrinu* ('snowman'). Another testament to the dynamic nature of

Maltese is the evolution of the /r/ sound, traditionally pronounced as an alveolar trill, similar to its pronunciation in Italian and Arabic, shifting to the English-style approximant in younger generations possibly due to the infiltration of English in various sectors of society (Mitterer et al., 2025).

As a result of the merging of elements described above, Maltese has been labelled as having a "double-sided morphology" (Spagnol, 2011, p. 12), consisting of a root-and-pattern system, also known as non-concatenative or templatic, typical of Semitic languages, and a stem-based system, commonly referred to as concatenative or affixational, derived from Romance and English structures. It should be noted that while the terms *non-concatenative* and *templatic* are often used interchangeably, they are not synonymous: *non-concatenative* refers broadly to morphology that is not formed by linear affixation, whereas *templatic* specifically denotes systems in which roots are interwoven with vocalic patterns according to fixed prosodic templates.

Table 2.2 below lists a select set of Maltese morphological and syntactic features of relevance to this research. This list is by no means exhaustive and excludes features from domains that intersect with morphology and/or syntax such as phonology. For more comprehensive accounts on Maltese and detailed descriptions, refer to Borg and Azzopardi-Alexander (1997), Comrie et al. (2009), Fabri (2010, 2014) and Spagnol and Fabri (2023), among others.

Table 2.2 *A Select Set of Maltese Morphological and Syntactic Features*

<table>
<tr><th colspan="2">Feature</th><th>Attributes</th></tr>
<tr><td rowspan="3">Inflection</td><td>Nouns</td><td>Number (singular, dual, collective, plural – suppletion, sound, broken)
Gender (masculine, feminine – plural forms are undefined for gender)
State (construct, absolute)</td></tr>
<tr><td>Verbs</td><td>Person (1st, 2nd, 3rd)
Number (singular, plural – for each person)
Gender (masculine, feminine in 3rd person singular)
Tense-aspect (perfective, imperfective)
Mood (indicative, imperative)</td></tr>
<tr><td>Adjectives</td><td>Number (singular, plural)
Gender (masculine, feminine – plural forms are undefined for gender)</td></tr>
<tr><td rowspan="2">Agreement</td><td>NP-internal</td><td>Number, gender
E.g., noun-adjective</td></tr>
<tr><td>NP-external</td><td>Person, number, gender (3rd person singular)
E.g., subject–verb</td></tr>
<tr><td colspan="2">Constituent order</td><td>Unmarked: subject–verb-object (can vary for topicalization, emphasis and stylistic reasons)</td></tr>
<tr><td colspan="2">Definiteness</td><td>Definiteness marked by il- (can also occur on the adjective; assimilates phonologically)
Indefiniteness is generally unmarked</td></tr>
<tr><td colspan="2">Pronouns and clitics</td><td>Pro-drop (subject, object)</td></tr>
</table>

	Independent pronouns and bound forms (clitics attach to verbs, nouns, prepositions)
Adjective degree	Positive, comparative, superlative (synthetic, analytic)
Diminutive and augmentative	Concatenative, non-concatenative
Negation	Anaphoric (*le* 'no'), verbal (bipartite construction *ma …-x*, negators), constituent (negator *mhux*, negative prefixes)
Question structure	Yes/no questions (marked by intonation), *wh*-elements typically clause-initial (other positions possible)

2.5.1.1. Nominal inflection in Maltese

Nominal inflection in Maltese primarily encodes number (singular, plural, collective, dual) and gender in the singular. Two grammatical genders, masculine and feminine, are distinguished, with various factors interacting to determine gender assignment, such as semantic and morphophonological ones (e.g., Farrugia, 2015).[34]

Inanimate nouns are typically assigned gender according to declensional class: nouns ending in *-u* or a consonant are usually masculine (e.g., *siġġu* 'chair.SGM'), whereas nouns ending in *-a* are generally feminine (e.g., *mejda* 'table.SGF').[35] Exceptions exist in both groups (e.g., *dar* 'house.SGF', *sema* 'sky.SGM'), particularly among loanwords and opaque forms (e.g., Farrugia, 2015, 2018; Stolz, 2009). Animate nouns, by contrast, usually reflect the referent's sex. Many form masculine–feminine pairs by suffixing *-a* to the masculine base (e.g., *kelb* 'dog.SGM' –

[34] See Farrugia (2018) for in-depth research on grammatical gender in Maltese.
[35] Other less common noun endings include *-i*, *-o*, and *-e*.

kelba 'dog.SGF'). Others are suppletive (e.g., *raġel* 'man.SGM' – *mara* 'woman.SGF'), draw on Italian-derived suffixes such as *-essa* (e.g., *prinċep* 'prince.SGM' – *prinċipessa* 'princess.SGF'), or are underspecified for gender (e.g., *pilota* 'pilot.SGM/F'; Fabri, 2009). Gender in Maltese is therefore morphologically salient but only partially predictable, requiring children to acquire both rule-based patterns (phonologically conditioned allomorphy) and lexical knowledge (lexically conditioned allomorphy).

Number inflection is realised through plural formation.[36] The plural system reflects the hybrid character of Maltese: Arabic in origin but shaped by contact with Sicilian, Italian, and English (e.g., Mifsud, 1995). Pluralisation relies on two major strategies: affix-based SPs and templatic BPs. The exact number of suffixes and templates varies by classification, with Nieder et al. (2020) identifying 12 SP suffixes and 11 BP patterns, while other accounts report smaller or larger sets (e.g., Aquilina, 1959; Borg & Azzopardi-Alexander, 1997; Marinaccio, 2024; Schembri, 2012).

SP forms are more common in non-Semitic nouns (e.g., *mutur* – *muturi* 'motors'), whereas BP forms are typical of Semitic-origin nouns (e.g., *kelb* – *klieb* 'dogs'). Some nouns allow both strategies (e.g., *kaxxa* – *kaxxi* (SP)/*kaxex* (BP) 'boxes'), others involve stem allomorphy (e.g., *lapes* – *lapsijiet* 'pencils'), suppletion (e.g., *tifla* – *bniet* 'girls'), or contrasts between the collective/type-noun[37] and determinate plurals (e.g., *siġra* – *siġar* – *siġriet* 'trees') and plural-of-plural forms (e.g., *elf* – *eluf* – *elufijiet* 'thousands'). Maltese also retains a dual suffix *-ejn/-ajn* with

[36] See Vella (2025) for a study on L1 acquisition of the plural, using a subset of the data collected for this research.
[37] See Mifsud (1996) for a distinction between the collective and the type-noun. Here, the term 'collective' is being used as a cover term for ease of reference.

a limited set of nouns (e.g., *jum* – *jumejn* 'days'). The coexistence of suffixal and templatic pluralisation makes the Maltese number system morphologically complex and a particularly interesting case for acquisition.

Beyond plural marking, Maltese nouns may also participate in a collective–unit distinction. The collective represents the 'unmarked form' of the noun, i.e., masculine singular and refers to the substance or type as a whole (Mifsud, 1996). For those nouns that display this distinction, the singular is morphologically derived through the addition of the suffix *-a* to the collective, yielding a feminine singular form. A select set of nouns of unity can also take the masculine singular suffix *-u* (SGM) (e.g., *nemla* (SGF), *nemlu* (SGM), *nemel* 'ant, ants').

Collectives can form part of complex paradigms that distinguish between a collective/type-noun, a noun of unity (derived singular), a determinate plural (indicating 2–10) and, in some cases, a plural of the type-noun (e.g., *dubbien* – collective/type-noun (SGM), *dubbiena* – noun of unity (SGF), *dubbiniet* – determinate plural, *dbieben* – plural of type-noun 'fly, flies'). Although the collective is morphologically masculine singular, it occurs in contexts where it refers to multiple objects and, in this sense, it can overlap semantically with the plural, as can be seen in (2).

(1)	*Kil-t*	*tuffieħ-a*	*velenuż-a.*
	ate-1SG.PFV	apple-SGF	poisonous-SGF
	'I ate a poisonous apple.'		(noun of unity)

(2)	*Għand-i*	*ħafna*	*tuffieħ*	*aħdar.*	
	have-1SG	a lot	apple.SGM.COLL	green.SGM	
	'I have a lot of green apples.'			(collective/type-noun)	

(3)	*Xtraj-t*	*żewġ*	*tuffiħ-at*	*żgħar.*	
	bought-1SG.PFV	two	apple-PL	small.PL	
	'I bought two small apples.'			(determinate plural noun)	

Adjectives inflect to agree with nouns in both gender and number. Masculine singular forms are generally unmarked, while feminine singular forms typically end in *-a* (e.g., *mutur ġdid* 'new.SGM motor.SGM' vs *karozza ġdida* 'new.SGF car.SGF'). Plural adjectival forms are created either through suffixation (often *-in*; e.g., *tajjeb* (SGM) – *tajba* (SGF) – *tajbin* (PL) 'good') or non-concatenative processes (e.g., *sabiħ* (SGM) – *sabiħa* (SGF) – *sbieħ* (PL) 'nice'). Participles follow similar patterns (e.g., *maħmuġ* (SGM) – *maħmuġa* (SGF) – *maħmuġin* (PL) 'dirty'). In addition, adjectives inflect for the comparative and superlative: comparatives may be synthetic (via Arabic templates) or periphrastic (with *aktar/iktar/iżjed* 'more'), while superlatives are formed by prefixing the definite article (*l-*). Although not directly relevant to gender and number inflection, such forms may appear in child language as base adjective forms. Hence their inclusion in this section.

In acquisition, children are expected to first stabilise transparent inflectional patterns (e.g., plural suffixes, animate gender alternations) before extending agreement to adjectives and, later, to more complex forms such as templatic plurals. Additional challenges arise with invariant nouns and opaque gender assignments, among other areas.

2.5.1.2. NP-internal agreement in Maltese

Agreement in Maltese operates both within the NP (NP-internal) and across the clause (NP-external).[38] The present discussion is limited to NP-internal agreement, which involves a) demonstratives and nouns, b) nouns and adjectives, and c) numeral/indefinite w-ħ-d ('one/a certain') and nouns (Fabri, 2023). The demonstratives in Maltese are *dan* (this.SGM), *din* (this.SGF), *dawn* (these), *dak* (that.SGM), *dik* (that.SGF), and *dawk* (those). In spoken Maltese, they are commonly shortened (e.g., *da'*, *di'*, *daw'*) and often fuse with the definite article (*dal-*, *dil-* and corresponding assimilated forms), as attested in the data collected. Agreement in b) and c) may occur in both pre- and post-nominal position, showing that agreement is not tied to syntactic position.[39]

Both nouns and adjectives in Maltese may take the definite article, which is cliticised and not marked for gender or number. Although definiteness is expressed on the noun and the adjective, Fabri (2001) shows that, unlike in Standard Arabic, it does not constitute an agreement feature. Indefinite adjectives can co-occur with definite nouns, indicating that agreement in Maltese is limited to gender and number. Demonstratives and adjectives must agree with the head noun in both gender and number, as illustrated in (4) below, where the demonstrative, noun, and adjective are all marked for feminine.

(4)	*Dik*	*papr-a*	*ħelw-a.*
	That.SGF	duck-SGF	sweet-SGF
	'That's a sweet duck.'		

[38] Refer to Fabri (2023) for details on grammatical agreement in Maltese.
[39] In some contexts, pre-nominal agreement with adjectives can be suspended.

2.5.1.3. Complexity and relevance to acquisition

While transparent and opaque morphological patterns occur cross-linguistically, the Maltese gender and number systems offer a particularly rich testing ground shaped by the interaction of Arabic, Romance, and, to a lesser extent, English influences on the morphosyntactic system being acquired in a bilingual context. Concatenative plural formation involves the addition of affixes to a stem and may represent a less cognitively demanding construction type than templatic plural formation. Similarly, animate gender alternations are mostly transparent, whereas the gender assignment of inanimate nouns requires both the acquisition of rules and memorisation. Agreement adds another layer of complexity, as mastery requires not only the acquisition of inflectional morphology but also its application across multiple targets within the NP.

This stands in sharp contrast to English, the other language acquired by the children in this study. English lacks morphological gender marking on nouns (with the exception of third-person singular pronouns) and adjectives. Also, in English agreement is minimal: subject–verb agreement in the present tense is limited to the third person singular (e.g., *she runs* vs *they run*), and adjectives are invariant. Number marking is affixal although non-concatenative forms also occur (e.g., ablaut: *man – men*).

Bilingual children acquiring Maltese and English are thus faced with typologically divergent systems: Maltese, with a morphologically rich inflectional and agreement system, and English, with sparse inflectional morphology and agreement. This contrast highlights the developmental significance of tracing how and when Maltese gender and number features emerge and stabilise in acquisition of MD bilingual children.

2.5.2. The language situation in Malta

Maltese is spoken by over 500,000 individuals, primarily in Malta, with additional speaker communities in countries such as Australia, the United Kingdom, and the United States. These diaspora communities have contributed to the development of distinct varieties of the language, such as *Maltraljan* – the Maltese spoken in Australia (see Bovingdon, 2004). Within Malta, several regional dialects of Maltese are actively spoken (e.g., Aquilina & Isserlin, 1981; Camilleri Grima, 2020), and Maltese English, that is, the variety of English spoken in Malta, is widespread (e.g., Caruana, 2025; Grech, 2024). While attitudes towards Maltese English vary, most generally "public perception, at least, is more likely to express concern at deteriorating standards rather than to view it with indulgence as an enriched illustration of local society" (Grech, 2015, p. 11).

Data on language use from the most recent national population census, 2021, indicate that the most widely understood languages among residents are English (96%), followed by Maltese (90.4%) and Italian (62%) (National Statistics Office, 2023a). While these numbers are partly attributable to the substantial presence of non-Maltese residents in Malta, they are also the result of the omnipresence of English in various sectors of Maltese life, but particularly in the spheres of the media and entertainment and in education. The 2021 census (National Statistics Office, 2023a) also revealed that almost a quarter (23.4%) of Maltese children aged between 5;0–9 consider English to be their main language, dropping to 14.7% for those between 10–19 years of age. While these proportions do not come as a surprise to those residing in Malta and interacting with young Maltese families, they depict a situation in which English is subtly infiltrating the linguistic spaces occupied by Maltese up until this point. As discussed in Section 2.5.2.1, below, these figures reflect a growing trend among Maltese parents to bring up their children with English as the L1 rather than Maltese.

2.5.2.1. The home setting

In her analysis of language use within Maltese families in 1995, Camilleri classified Maltese families into four ‘types’ according to the languages/varieties being acquired in these families: a) Type A: dialect as L1, Standard Maltese and English as L2, b) Type B: Standard Maltese as L1, English as L2, c) Type C: Standard Maltese and English acquired simultaneously and, d) Type D: English as L1, Standard Maltese as L2.

The influx of non-Maltese residents over the past 10 years or so, soaring at one in every five Maltese residents being non-Maltese as of 2019 (Eurostat, 2019), results in an added layer of complexity which calls for the need to supplement this list of types with a fifth type catering for those children having a foreign parent(s), acquiring the parent(s) language(s) as the L1 and then English and, potentially, Maltese later through socialisation and formal schooling (Vella, 2020). This reality is palpable in many Maltese schools that “face plurilingualism and linguistic diversity resulting from immigration” (Caruana et al., 2019, p. 331) and are not well-equipped to cater for multilingual classroom environments. It is also worth noting that a growing number of children with two Maltese parents are being brought up with English as their primary home language (e.g., Vella, 2020).

As a result of the complex linguistic situation, the question surrounding the choice of Maltese or English as the L1 in Maltese homes continues to be "an immensely loaded one for many families" (Vella, 2012, p. 542). This observation was also made in Borg (1988) who had already pointed out that some families were making the choice to bring up their children with English as the L1: the reasons for such a choice were and continue to be affected by the events and sociocultural context of the time. In this context of the presence of English alongside Maltese in everyday life, CS between the two languages, both in the speech directed to children and in the children's speech itself, is widespread and differs from family to family (e.g., Azzopardi-Alexander, 2016; Gatt, 2017).

Lastly, in recent years, increasing attention has been directed toward promoting Maltese as the L1. Language policies and initiatives (e.g., L-Aġenzija Nazzjonali tal-Litteriżmu, 2021; Vella et al., 2022) reflect a broader societal effort to strengthen the role of Maltese in education, media, and the home environment. As noted by Grech (2024, p. 280), "there is sometimes an underlying concern that the dominance of English in different spheres might threaten the ability of Maltese—and indeed other languages increasingly evident throughout the islands—to thrive and remain visible". Efforts are also being made by academics and local personalities, particularly through social media, to encourage Maltese families to use their native language at home and to increase the visibility of Maltese more generally. A number of families participating in this research fall within the category of parents who have consciously chosen to raise their children with Maltese as their L1. Other participating families have found this to be the most natural course of action, while some have adopted bilingual strategies, such as the one-parent-one-language approach, still placing primary emphasis on Maltese as the L1. Thus, all the participants in this study are MD to varying degrees.

2.5.2.2. The childcare setting

In Malta, schooling is compulsory for children aged five to sixteen and consists of six years of primary education and five years of secondary education. Childcare centres and kindergarten are optional for children up to five years old. As of 2023, 51% of children under the age of three in Malta were enrolled in formal childcare, dropping to 44.1% in 2024 (Eurostat, 2024). This marks a significant increase from 20.2% in 2013 and surpasses the EU average of 37.4%. This increase is largely attributed to the introduction of the national Free Childcare Scheme for working/studying parents of infants as young as three months, in which 187 out of 192 Maltese childcare centres are enrolled. In addition, statistics indicate that 26.8% of children under the age of three attending childcare in Malta spend 30 hours or more per week at a childcare centre (Eurostat, 2024).

A language policy for the early years (0–7: childcare, kindergarten, first two years of primary school) is in place, stating that “Maltese and English are to be introduced and developed in tandem in the Early Years” (Ministry for Education and Employment, 2016, p. 5). However, for a diverse host of reasons, this ideal is not being met. One contributing factor is the significant presence of non-Maltese workers and children in childcare centres, many of whom do not speak Maltese, leading caregivers and children to use English as the primary language of communication. While specific figures for the childcare sector alone are not provided, the National Statistics Office (2023b) data indicate that of the 15.5% of non-Maltese workforce in the care sector, one-third are employed in education. This suggests that a significant portion of non-Maltese care workers are engaged in roles related to childcare.

From my own experience with young MD families, including the participants in this study, I have observed that language dominance shifts when children enter childcare. Parents express concerns about their children's dip in the use of Maltese when attending childcare; however, they sometimes feel compelled to shift to English even in the home setting, especially during playtime. This concern was also voiced by academics, policy makers and curriculum curators during the yearly conference *Il-Malti f'Politika Lingwistika Nazzjonali* ('Maltese in a national linguistic policy') held on the 3rd of May 2025. The significance and urgency of promoting the exposure to and use of Maltese in childcare settings were highlighted at the conference, alongside the need for additional initiatives and engaging Maltese-language resources aimed at children across all age groups (Newsbook, 2025), which initiatives are currently on the rise.

Since this study is longitudinal and a number of participants were visited both prior to commencing childcare and during childcare attendance, a transition from more to less MD from one session to the next was often in evidence. Out of the eight participants in this study, only one child did not attend childcare at any point, and although many other factors must be weighed in, this child is the one who, throughout the data collection stage can be demonstrated to be by far the most MD. Interestingly, a reverse effect is sometimes observed when transitioning from childcare to kindergarten/primary school, that is, parents note a rekindled interest in Maltese, a positive attitude towards the language and an increase in Maltese dominance. This effect, however, appears to be strongly influenced by the school context, or, on a more micro level, by the individual classroom environment.

As articulated by Pace and Borg (2017, p. 83), “much still remains to be done to give Maltese its rightful place in the life of the island nation”. Their observation aligns with current efforts to promote balanced bilingualism in the early years and with initiatives increasingly emphasising the importance of providing children with rich input in both Maltese and English. Within this context, exposure to Maltese in early educational and home settings can be seen as a key factor in supporting successful bilingual acquisition and the continued vitality of Maltese as an L1.

2.5.3. Existing research in Malta

MFLA is an understudied field, and it is also underrepresented in the broader landscape of international language acquisition research (see Section 2.5.4, below). Evidence of this is Kidd and Garcia’s (2020) study, which lists a mere three articles in total available for Maltese in the four main language acquisition journals (Journal of Child Language, First Language, Language Acquisition, and Language Learning and Development) from their inception in 1974 to 2020. While published research on MFLA is available in sources beyond the four main journals accounted for by this study, research in this field is clearly still minimal. The existing published research on MFLA focuses mainly on three aspects: a) lexical development and vocabulary production (e.g., Gatt, 2017; Gatt & Dodd, 2020), b) phonological acquisition (e.g., Grech & Dodd, 2008; McLeod & Crowe, 2018), and c) language mixing (e.g., Azzopardi-Alexander, 2016; Camilleri Grima, 2016). In addition, Maltese has received limited attention in cross-linguistic CLA research (examples include Katsos et al., 2016; McLeod & Crowe, 2018; Sauerland et al., 2016).

The large proportion of MFLA studies involve unpublished dissertations submitted to the Institute of Linguistics & Language Technology, the Department of Maltese, the Faculty of Education and the Department of Human Communication Sciences & Disorders, all at the University of Malta. Of these studies, few place exclusive emphasis on the linguistic trajectory of TD children acquiring Maltese as their L1. This body of research predominantly addresses the following thematic areas: a) lexical development (e.g., Debattista, 2024; Farrugia, 2024; Frendo Wirth; 2020; Gatt, 2010), b) the effect of input and other contextual factors (e.g., Baldacchino, 2020; Nappa, 2024; Pulis, 2021), and c) bilingual exposure and interaction (e.g., Caruana Lia, 2016; Scerri, 2015; Sultana, 2014). Past local unpublished research also focuses on case studies and diary entries of the researchers' own children/child relatives centring on CLA (e.g., Bartolo, 2014; Vassallo, 2006) or simply touching upon CLA as part of broader developmental research (e.g., Camilleri, 1987; Cassar, 1986).

Table 2.3 below lists studies that focus on aspects of morphosyntactic L1 acquisition of Maltese between ages 0–5;0 from 1980 onwards in chronological order. While the list aims for completeness, certain contributions may have been inadvertently overlooked.

Table 2.3 *List of Studies Focusing on Aspects of Maltese L1 Morphosyntactic Acquisition*

Year	Title	Author	Contribution Type
1982	*Aspects of the language of Maltese five year olds*	Cachia, P.	Bachelor's dissertation
1997	*Mean length of utterance in three to four-year old Maltese speaking children*	Portelli, J.	Bachelor's dissertation
1997	*The emergence of the plural in three to five-year old Maltese speaking children*	Murray Curtis, J.	Bachelor's dissertation
2002	*The acquisition of number and gender agreement between nouns and adjectives in three to five year old Maltese speaking children*	Coppini, A.	Bachelor's dissertation
2002	*Inflection and agreement in child language: a case study*	Desira, G.	Bachelor's dissertation
2005	*The mean length of utterance in Maltese-speaking Children*	Portelli, J.	Master's dissertation
2006	*The proportion of nouns and verbs in two and three year old Maltese-speaking children*	Debono, G.	Bachelor's dissertation
2006	*The language development of a Maltese toddler: a case study*	Vassallo, M.	Bachelor's dissertation
2008	*The use of number and gender markers in the Maltese; 3 to 6 year old children*	Zammit, R.	Bachelor's dissertation

2008	*Investigating the relationship between expressive vocabulary and MLU*	Bajada, M. A.	Bachelor's dissertation
2010	*Early expressive lexical development: evidence from children brought up in Maltese-speaking families*	Gatt, D.	Doctoral thesis
2014	*Il-lingwaġġ tat-tfal: studju tat-taħdit ta' tifla ta' sitt snin [The language of children: a study of the speech of a six-year-old girl]*	Bartolo, D.	Bachelor's dissertation
2016	*Comprehension and production of nouns and verbs in Maltese five-year olds*	Attard, D.	Bachelor's dissertation

Portelli's (2005) work, listed above, is pertinent to the present study as it documents morphological development and details the adaptations made to Brown's (1973) calculations of MLUm, enabling the application of MLUm to Maltese. Also included is the doctoral thesis by Gatt (2010), which focused on lexical rather than morphosyntactic development. Although its primary emphasis differs from the studies listed in Table 2.3, it represents a significant contribution to research on MFLA. The study provides an analysis of lexical growth, closely interrelated with morphosyntactic development, and examines POS distributions as well as the acquisition of Maltese and English lexical items. While both Portelli and Gatt focus on MD children, the former adopts a cross-sectional design encompassing children aged 1;6–5;0, while the latter integrates both longitudinal and cross-sectional methodologies to explore the development of expressive vocabulary from 1;0–2;6.

2.5.4. The research gap

Although both theoretical and applied linguistic research on Maltese is gaining momentum, with steady advances across various domains, on the international scholarly stage Maltese continues to occupy a relatively marginal position. This limited visibility may be attributable to its status as a small language competing with more widely spoken ones, as well as the scarcity of open-access computational resources, such as spoken language corpora and comprehensive online dictionaries, both of which are currently in development. Despite a longstanding tradition of linguistic work on Maltese dating back to the nineteenth century, the lack of large-scale digital resources continues to hinder progress in several areas of linguistic research and restricts the inclusion of Maltese in broad cross-linguistic studies and language-model frameworks. As a result, the development of first-rate, open-access tools for Maltese, characterised by high accuracy and applicability, continues to lag behind that of many other languages, even as initiatives at the University of Malta and beyond are increasingly prioritising such resources.

One of the key factors that has likely deterred researchers from working on MFLA is the absence of an online, publicly accessible repository of Maltese child language data. This lack of accessible resources has significantly stymied the development of in-depth studies, contributing to the current research vacuum. In this respect, the most time-intensive aspect of the present study was the collection, manual transcription, and annotation of a substantial dataset. To the best of my knowledge, none of the studies listed in Section 2.5.2 span a two-year longitudinal period or encompass more than 30 hours of recorded child language, nor do most of them focus exclusively on linguistic development. For instance, although Cassar

(1986) and Camilleri (1987) each conducted longitudinal studies, only a limited portion of their work addresses language acquisition directly.

Moreover, the few existing studies on Maltese morphosyntactic development in child language have generally taken the form of small-scale dissertations, restricting their scope to one or two features and preventing broader analysis of concurrent morphosyntactic development. Given the typologically hybrid nature of Maltese, the contribution of MFLA research to the wider field of CLA is particularly salient. Work on the acquisition of morphosyntactic elements is thus not only warranted but essential for advancing our understanding of language acquisition within lesser-studied linguistic systems and for contributing to cross-linguistic theories of morphosyntactic development. The present study seeks to address this gap by investigating the early development of morphosyntactic features in MD children aged between 1;6–5;0.

2.5.5. The research questions

This study therefore set out to investigate the developmental trajectory that characterises MFLA from a morphosyntactic perspective, focusing specifically on nominal inflection and agreement, while still considering other morphosyntactic features that develop concurrently. In view of the overarching aim and the primary focal points, the study attempts to answer the following research questions (RQs) within the present sample:

RQ1. What does the combined progression of MLUm and POS category distribution indicate about the developmental trajectory of early morphosyntactic acquisition in MD children?

RQ2. When and how do gender and number features surface and stabilise across nominal inflection and agreement in the present sample of MD children in the present sample?

RQ3. How are patterns of nominal inflection and agreement related to changes in MLUm and POS category distribution in the present sample?

The methodological considerations described in the upcoming section stem from the drive to investigate the above questions.

3. Methodology

This chapter presents the methodology employed in this investigation of MFLA with a focus on nominal inflection and agreement. It begins by detailing ethical approval (Section 3.1) and participant recruitment (Section 3.2). The method section (Section 3.3) and its subsections elaborate on this study's integration of different data collection approaches and designs, and the development of a test battery tailored to address the RQs articulated in the previous chapter (Section 2.5.5). The procedures and conventions for data collection (Section 3.4) and transcription (Section 3.5) are then described. Finally, the analytical methods employed to examine MLUm (Section 3.6), POS category distribution (Section 3.7) and nominal inflection and agreement (Section 3.8) are outlined.

3.1. Ethical Clearance

Ethical clearance to collect and store all the necessary data was obtained from the University of Malta Research Ethics Committee (UREC) in January 2022 and amended to reflect minor changes to the study in March 2024 (REDP Application ID: MAKS-2021-00038), namely, the addition of a child assent form (see Section 3.2.2; Appendix F) and a short interview with a parent[40] as (one of) the child's primary language input model(s) conducted during the final visit to each family (see Sections 3.2.2 and 3.4; Appendix G).

[40] In all cases, the primary caregivers of the participants were their parents, who are referred to as such throughout this study.

3.2. Participants

3.2.1. Study design and participant overview

The study was designed to examine developmental patterns in MFLA through a combined longitudinal and cross-sectional design. Eight children were followed longitudinally over 20–25 months, with 2 children in each age cohort. At the initial stage of data collection, participants were aged 18, 24, 30, or 36 months. Children were revisited at regular intervals, every four months until they reached 30 months and every six months thereafter, resulting in recordings covering the period from 18–60 months of age. This design allowed the study to mirror a longer longitudinal investigation of three and a half years within the timeframe of a full-time doctoral project.

These age points correspond to the critical period during which L1 acquisition develops most rapidly. As noted in the literature (e.g., Clark, 2024; Ravid, 2019; see Section 1.1.1), core linguistic systems are typically established between one and five years of age, although the process continues into early adolescence. Sampling across this range therefore provided an opportunity to identify key stages in early morphosyntactic development (see Section 2.3 for more on CLA milestones).

3.2.2. Participant selection and recruitment

Participants were recruited according to predefined inclusion criteria designed to ensure comparability across developmental stages and to control for relevant linguistic variables. The objective was to obtain a sample of TD, MD children balanced across age bands and sex.

- Age: Children were selected from four age cohorts, 18, 24, 30, and 36 months, at the initial data collection stage. This arrangement allowed for both longitudinal tracking and cross-sectional comparison.
- Sex: An equal number of boys and girls was included to maintain demographic balance and to enable comparisons not influenced by sex-related variation in early language development.
- Typical development: This study focused on TD children. Parental questionnaires addressed developmental milestones, physical or neurocognitive conditions and family histories of hearing or speech–language difficulties to confirm eligibility.
- Language dominance and exposure environment: All children were receiving MD exposure at the time of recruitment. Both parents were native speakers of Maltese, and all children were born and raised in Malta. Language dominance was established through parental language background questionnaires detailing the relative amount of Maltese and English input the child received. Given Malta's bilingual context, in which both languages are used to varying degrees (e.g., Fabri, 2015; Vella, 2012), establishing language dominance at recruitment was necessary to contextualise participants' linguistic exposure (see Section 2.5.2).
- Commitment to longitudinal participation: Since the project spanned two years, only families willing to commit to long-term participation were recruited to ensure consistency across data points and reduce attrition.

Recruitment followed a short pilot phase involving two girls (aged 2;3.24 and 4;11.07) conducted in February–March 2022. Following the pilot, minor adjustments to the procedure were implemented. These included ensuring that the audio recorder remained in close proximity to the child throughout each session, repositioning it as necessary when the child moved between play areas to optimise recording quality. The selection of play materials was also expanded to include a wider range of tactile objects and toys aimed at eliciting more varied linguistic output. The main recruitment began in late March 2022 through personal networks and social-media outreach.

Each family's first home visit served as the recruitment meeting. Parents received an information sheet, recruitment letter, and consent form (see Appendices A–C) and completed both a language background and a child development questionnaire (see Section 3.2.3; Appendices D and E). All documents were available in Maltese and English. In accordance with best practice in research ethics, assent was also obtained from the children at the start of each session, prior to recording. For children aged 48 months and older, assent was also recorded on a simple form where children shaded a happy or sad face to indicate willingness to participate. The procedure of obtaining child assent encouraged child autonomy and decision-making and helped participants better understand my intentions as a researcher. All children assented, and no exclusions were necessary on these grounds.

Final Sample

In total, eight children were initially recruited (four girls, four boys). Subsequently, an additional boy and a set of twin boys were added, all within the 30–32-month group. These additional participants provided contingency against attrition and allowed for expansion of the limited corpus of Maltese child language data.

The twins' data were ultimately excluded from analysis given the well-documented developmental differences that exist between twins and singletons (e.g., Oliver & Plomin, 2007; Ozturk et al., 2021; Rice et al., 2014).[41] Data from one girl were also excluded after she developed bilateral otitis media (OM) at 3;10, also known as middle ear infection and commonly referred to as glue ear. OM can temporarily impair auditory perception and influence language acquisition. Previous research has associated OM with reduced phonological sensitivity (e.g., Nittrouer & Lowenstein, 2024), literacy delays (e.g., Winskel, 2006), smaller receptive vocabularies (e.g., Brennan-Jones et al., 2020), and delayed morphosyntactic development (e.g., Tomblin et al., 2015). Therefore, although recording sessions with this child were not discontinued, the data were excluded from the analysis.

Consequently, for the 30–32-month cohort, data were retained for two boys rather than one boy and one girl, producing a sex imbalance. This was taken into account during analysis, given that significant sex disparities have been reported in

[41] The data collected from the twin participants remain of considerable interest and may serve as a valuable contribution to the field. To the best of my knowledge, no previous research on twin language development has been conducted within the context of MFLA. These data therefore represent a potential avenue for future investigation.

CLA,[42] with boys tending to be delayed in achieving language milestones when compared to girls (e.g., Eriksson et al., 2012; Ponvannan et al, 2024), and with vocabulary composition varying as a result of sex-based differences[43] (e.g., Wallentin & Trecca, 2023).

Although regional and socio-economic factors were not formal selection criteria, they were noted to contextualise the sample. Most families were drawn from central and southern areas of Malta, with some locality overlap. SES data derived from parental occupation, educational level, and family size (see Section 3.2.3) indicated that all families fell within a middle-class profile. This homogeneity reduces SES-related variability within the present sample and limits generalisation across socioeconomic strata, which falls beyond the scope of the present study.

Table 3.1 presents the participants' exact ages (in months and days) at each of the five recording sessions. All pseudonyms used in Table 3.1 are maintained throughout the thesis to preserve anonymity.

[42] Some studies, however, report minimal or no sex-based differences, instead identifying other variables as more influential in shaping language acquisition (e.g., Ibbotson & Browne, 2024; Rinaldi et al., 2021).

[43] For instance, research shows that girls often outperform boys in lexical domains related to clothing and body parts, whereas boys tend to excel in words associated with vehicles and outdoor scenes. See Wallentin and Trecca (2023) for a detailed discussion.

Table 3.1 *Age of Participants, in Months and Days, at Each of the Five Recording Sessions, Coupled with Information on Sex*

Age Group (months)	Name (pseudonym)	Sex	Age at visit 1	Age at visit 2	Age at visit 3	Age at visit 4	Age at visit 5
18-20			**18-20**	**22-24**	**26-28**	**30-32**	**36-38**
	Anna	Female	1;6.05	1;10.27	2;2.17	2;7.05	3;0.20
	Stiefnu	Male	1;7.08	1;11.05	2;3.15	2;7.07	3;0.30
24-26			**24-26**	**28-30**	**32-34**	**38-40**	**44-46**
	Karmena	Female	2;0.30	2;4.21	2;8.02	3;2.09	3;8.15
	Lippu[44]	Male		2;3.21	2;9.05	3;2.07	3;8.25 4;2.06
30-32			**30-32**	**36-38**	**42-44**	**48-50**	**54-56**
	Duminku	Male	2;6.12	3;0.09	3;6.17	4;1.03	4;6.23
	Toninu	Male	2;6.26	3;0.25	3;7.04	4;2.07	4;8.01
36-38			**36-38**	**42-44**	**48-50**	**54-56**	**60-62**
	Kelina	Female	2;11.24	3;5.20	4;0.01	4;6.02	4;11.27
	Reno	Male	3;0.13	3;6.24	4;0.18	4;6.10	5;0.10

[44] Due to travel-related absence, the boy in this age group was 27 months old at the time of the initial data collection session. Nevertheless, five recording sessions were completed.

3.2.3. Questionnaires

In addition to the recruitment and consent materials, two questionnaires were administered during the initial visit: a child development questionnaire and a language background questionnaire (see Appendices D and E). Both were purposely designed for this study, provided in Maltese and English, and completed by one of the child's parents during the first home visit.

Child Development Questionnaire

The child development questionnaire gathered background information and gave an indication of each participant's reported developmental status at the outset of data collection. Particular attention was given to achieving a balance between comprehensiveness and practicality, that is, the questionnaire included a sufficient number of items to provide a detailed and vivid picture of each child's background, behaviour, and overall development, while remaining concise enough to avoid respondent fatigue and ensure reliability of reporting. In designing the questionnaire, a range of established developmental assessment instruments was consulted, and several items were modelled on Squires and Bricker's questionnaire (2009, 3rd ed.). Four age-specific versions were produced for the 18-, 24-, 36-, and 48-month cohorts.

The questionnaire comprised six sections:

- Child background information: sex, birth order, and SES (derived from parental education, occupation and family size), since all three factors are established correlates of early language outcomes (e.g., Berglund et al., 2005; Ensminger & Fothergill, 2014; Fernald et al., 2013; Kuvač-Kraljević et al., 2021; Pine, 1995).
- Current communicative functioning: two multiple-choice items mapping onto recognised stages of language acquisition (e.g., Clark, 2024; Lust, 2006).
- Developmental milestones: language and motor development, reflecting the synchrony between motor and linguistic growth (e.g., Diepeveen et al., 2018; Lenneberg, 1967; Sanjeevan et al., 2015).
- Prenatal and perinatal history: identifying potential risk factors associated with later language delay (e.g., Allotey et al., 2018; Sansavini et al., 2011).
- Medical history: including ear infections and OM, due to their documented long-term effects on phonological and literacy development (e.g., Brennan-Jones et al., 2020; Winskel, 2006).
- Habits and behaviour: questions on play, interaction, and screen-time exposure included to identify possible late-talker profiles and environmental risk factors, since, for example, reduced initiating social interaction behaviours have been associated with delayed language development (e.g., Hutton et al., 2019; Vuksanović, 2015).

Language Background Questionnaire

The language background questionnaire assessed language dominance and exposure patterns within each household. Parents provided self-ratings of proficiency in Maltese and English across listening, speaking, reading, and writing on a four-point scale (1 = limited, 4 = advanced). Mean scores across the four domains were calculated for each language, and these were considered alongside reported language use in the home to determine relative language dominance (MD, English-dominant, or balanced bilingual).

The questionnaire also collected information about language use by parents, caregivers, and children across contexts (home, extended family, childcare). Given that Maltese dominance was an inclusion criterion established during the initial recruitment visit, the questionnaire served to corroborate the initial information obtained and to describe the linguistic input environment in greater detail, with the understanding that self-report measures provide indicative rather than objective data. The self-assessment measures of Maltese and English proficiency used in this language background questionnaire were adapted from Dandria (2002), who in turn based them on Romaine's (1995) self-rating test.

Parental Speech Sample

Several questionnaire items were revisited in a short interview with one parent, conducted immediately after the final data collection session with each child (see Appendix G). These brief recordings were designed to provide representative samples of parental speech, forming the primary linguistic input model for the child. Parental speech is widely recognised as playing a central role in language acquisition, reflecting the dynamic interaction between parent and child (e.g., Laalo & Argus, 2020). It also influences children's linguistic development through both genetic factors (e.g., hereditary speech sound disorders) and exposure to language use within the home. Each interview comprised a short spot-the-difference task and a semi-structured conversation between the parent and the researcher. Further details of the parental speech recordings are provided in Section 3.3.

3.3. Method

3.3.1. Combined longitudinal and cross-sectional design

Both longitudinal and cross-sectional methods have contributed substantially to the field of CLA. While longitudinal designs provide insights into developmental trajectories and capture the dynamic nature of acquisition, cross-sectional designs offer a snapshot of language development at a specific moment and facilitate the identification of age-related differences (Eisenbeiß, 2006).

The present study employed a combined longitudinal and cross-sectional design. The same children were followed at regular intervals, and comparisons were also made across children at the same point in time. Additionally, this design allowed both cross-sectional (between age-group) comparisons (e.g., 18-, 36-, and 60-month-olds) and cross-participant (within age-group) comparisons. At each data collection point, at least two children were of comparable age, and at certain points, such as around 35–36 months, up to six children could be compared directly.

This enabled analysis not only of developmental change within individuals but also of contemporaneous variation between children.

As Snyder (2021, p. 873) observes, “a longitudinal corpus provides by far the best available record of an individual child’s acquisitional time course” as long as data collection occurs at regular intervals. In this study, data collection visits were conducted every four to six months. Although individual linguistic development was tracked across sessions, more frequent visits (e.g., weekly or monthly) would have allowed for a more fine-grained account of each child’s linguistic progression. However, given the number of participants and the 20–25-month data collection period, more frequent visits were not practical.

CLA research benefits from designs that combine longitudinal and cross-sectional approaches, as these permit detailed analyses of individual development while simultaneously capturing cohort-level variation at specific time points (e.g., Conboy & Thal, 2006). In the present study, the relatively small number of participants (from a cross-sectional perspective) and the regular but relatively spaced out visit schedule (from a longitudinal perspective) offset one another, as the combined design generated a substantial dataset comprising over 40 recordings of child language. These were collected at approximately two-month intervals (owing to participants’ staggered recruitment ages) over a maximum period of 42 months (see Figures 3.1 and 3.2 in Section 3.5).

3.3.2. Naturalistic and experimental techniques

In addition to combining longitudinal, cross-sectional and cross-participant designs, this study collected both naturalistic speech data and semi-structured experimental data. To mitigate the limitations and capitalise on the strengths of naturalistic, experimental, and semi-structured research methods (See Section 2.4.3), a mixed-methods approach was adopted. The primary dataset consists of naturalistic language samples, complemented by a series of semi-structured elicitation tasks, including a comprehension activity and a novel-word plural task.

This integrative design enabled a comprehensive investigation of morphosyntactic development in Maltese, a language still with relatively limited research on L1 acquisition. The combination of methods not only enhanced the ecological validity of the findings but also permitted controlled exploration of specific linguistic phenomena.

3.3.2.1. Naturalistic data

The study was designed to collect naturalistic data through play and spontaneous speech interactions. A set of age-appropriate toys with minimal noise emission (e.g., no rattling toys or Velcro fasteners) was selected for presentation at each session to invite children to play and encourage interaction with the researcher. These toys were carefully chosen to elicit a range of morphosyntactic constructions. For instance, an assortment of plastic fruit and vegetables (farmer's market sorting set) was used to elicit noun–adjective agreement, while a set of toy vehicles was intended to prompt the production of SP and BP forms.

The curated selection also ensured a variety of sensory experiences (e.g., wood, fabric, plastic, and felt) and excluded all potentially hazardous items. Several alternative activities, such as craft-making and baking, were considered but excluded to avoid a) unnecessary accidents, b) excessive noise that could have compromised recording quality, c) immersive tasks that might have hindered elicitation or spontaneous conversation (e.g., jigsaw puzzles requiring concentration), and d) activities that risked creating messy or intrusive settings. Table 3.2 lists the toys used and the rationale for their inclusion in the study design.

Table 3.2 *List of Toys Used and the Reason for their Inclusion in the Study Design*

Item	Reason for inclusion
Dog hand-puppet	Facilitate elicitation
Soft foam balls	Elicit colour terms, broken plural, prepositions
Farmer’s market sorting set	Elicit colour terms, agreement
Animal figurines	Conversation starter – animals (favourite, pets…)
Noah’s Ark wooden set	Elicit sound and broken plural, prepositions
Toy vehicles	Elicit agreement, degree adjectives
Two-storey parking lot	Prompt pretend play, elicit verbs
City roads mat	Elicit directions, vocabulary related to locations
Felt pizza	Conversation starter – food (favourite, cooking…)
Hand-painted wooden furniture set	Prompt pretend play, elicit verbs
Playmobil playground set	Conversation starter – fun activities with family
Playmobil adventure treehouse	Prompt pretend play, elicit verbs
Playmobil ski lodge	Conversation starter – holidays, travelling, winter
Barbie vintage fabric game	Elicit verbs, adjectives, vocabulary related to fashion

All the items listed in Table 3.2 were either the researcher's own childhood toys, borrowed from friends or family, or purchased (new or second-hand in good condition) for the purpose of this study (see Appendix H for images). Not all toys listed above were available at each session, as suitability varied according to the target age group. Toys were presented one at a time, and each set of items was cleared away before introducing a new set to maintain order and consistency. In addition, conversation starters with and without the use of toys were planned in advance and adapted during each session. For instance, while settling down, the researcher asked participants about previous activities carried out that day, and, while playing with pretend food items (e.g., farmer's marking sorting set, felt pizza), the researcher asked about the child's favourite food.

The study design entailed that the same set of toys be presented to children of approximately the same age to facilitate the collection of comparable samples (Eisenbeiß, 2006). However, it was anticipated that full consistency would not always be achievable, as a) some participants occasionally preferred to play with their own toys during sessions, and b) play preferences varied, with some children engaging with a single toy for an extended period while others rotated through several. These variations did not compromise the study, as the researcher adapted elicitation strategies and conversation prompts to ensure that relevant morphosyntactic constructions were still targeted. A degree of flexibility is essential in studies involving infants, particularly in longitudinal designs. The safety, wellbeing, and comfort of each child were therefore prioritised from the study's conception.

3.3.2.2. Experimental data

Since spontaneous speech recordings are not always likely to produce data which are "more representative of the full range of usage" (Brown, 1973, p. 326), semi-structured experimental techniques in the form of receptive and production tasks were also administered. Receptive tasks were conducted with participants aged 18–36 months, while production tasks were administered to participants aged 30 months and older, depending on the task requirements. Consistent with previous language acquisition research (Eisenbeiß, 2010), most production tasks were introduced when children were at least 36 months old. Table 3.3 presents a comprehensive list of all the experimental tasks designed for this study, together with the materials required for each task and target age group. All tasks generated data that were included in the analysis.

In a number of the experimental tasks, the dog hand-puppet, Bobo, was employed (see Section 3.3.2.1). The aim was to introduce Bobo at the start of each recording as an icebreaker and to incorporate him into the session as necessary to ease elicitation during free play and experimental tasks. Additionally, since all participants in this study are bilingual to different degrees, the aim was to use Bobo as an engaging tool with which to interact in Maltese, as he was consistently introduced as a Maltese monolingual dog. He was also described as having poor eyesight to encourage participants to explain what they were seeing/doing from time to time, particularly during narration tasks. Following Eisenbeiß's (2010) advice, the puppet was not blindfolded nor fitted with broken glasses (common practices in early production studies with young children), as both have been found to upset children at times. See Section 2.4.3 for more on the use of puppets in CLA studies.

Table 3.3 *List of Experimental Tasks with Materials Required and Target Age Group*

Task	Materials Required	Target Age (months)
Colour-sorting activity	Farmer's market sorting set	18–28
Quantifier comprehension activity	Noah's Ark, toy animals in sets of 4, dog hand-puppet	18–36
Real-word plural task	Screen, slides with carrier phrase and images (Microsoft PowerPoint)	
-Abridged version	10 items – 20 slides	36–48
-Full version	20 items – 40 slides	36–54
Real-word plural task bingo	2 laminated bingo sheets in colour, 2 corresponding packs of flashcards in colour, 2 whiteboard markers, wipes	44–60
Novel-word plural task	Screen, slides with carrier phrase and images (Microsoft PowerPoint – 20 items – 40 slides)	44–60
Opposites task	Screen, slides with questions and images (Microsoft PowerPoint – 12 items – 12 slides)	38–60
Spot-the-difference	8 sets of spot-the-difference laminated sheets in colour (16 sheets in total)	42–60
Narrative Elicitation -Picture books	3 small animal board books in colour – domestics, farm and jungle animals	18–28
-Short animated videos	4 short clips (duration: 3–5 minutes), dog hand-puppet	36–60
-Picture stories	5 laminated picture stories in colour with 4 tiles per story	42–60
-Illustrated books	3 short, illustrated children's books in colour, in Maltese	42–60
-Mayer's (1969) *Frog Story*	Screen, slides containing Mercer Mayer's wordless story 'Frog, where are you?' (1 illustration per slide), microphone	54–60

Most of the experimental tasks also made use of a screen. In all cases, the screen used was a 13-inch laptop monitor, placed directly in front of the child. Each experimental task lasted no longer than 5 minutes. Procedures were explained in simple terms, and tasks were conducted in Maltese with the help of the dog hand-puppet. The visual stimuli used for most tasks were simple, easily recognisable clip-art-style images sourced from open-access repositories that permit free reuse for research and educational purposes.

The experimental tasks listed in Table 3.3 are described below. Each task description includes a brief account of its purpose and methodological relevance. Protocol instructions, that is, short researcher scripts used to introduce and guide participants through each task, are provided in Appendix I, with only one representative example (from the colour-sorting activity) retained here for illustration.

Colour-sorting activity

In this hands-on conceptual task, designed to investigate children's colour-sorting abilities, participants were asked to place all the orange fruit and vegetables (carrot, tangerine, apricot, pumpkin) in one bowl and all the rest in another. Colour-sorting activities play a significant role in CLA, particularly in the development of colour term capabilities (Saji et al., 2020; Wagner et al., 2013). Research indicates that children's understanding of colour is closely linked to their linguistic abilities and environmental exposure (Roberson et al., 2004; Scott et al., 2023). The fruit and vegetables used in this task were all previously named by the child and the researcher, thus familiarising the child with all items prior to testing. This task was conducted with children aged 18–28 months.

<u>Protocol instruction example:</u>
Tista' tpoġġi l-affarijiet oranġjo kollha hawn ġew?
'Can you put all the orange things in here?'
(Researcher pointing to a plastic bowl)

L-oranġjo biss. L-oħrajn agħmilhom hawnhekk.
'Just the orange ones. Put the rest here.'
(Researcher pointing to another plastic bowl, both in full view and reach of the child).

Quantifier comprehension activity (one vs all)

Modelled on the Give-a-Number task (Wynn, 1992) and similar numerical development comprehension tasks (e.g., Barner et al., 2007; Meyer, 2019; Socarrás, 2011), this activity required children to follow the dog puppet's requests to place 'one' (*wieħed*) or 'all' (*kollha*) of a certain animal (four of each were available: elephants, giraffes, tigers, lions, zebras) in a wooden vessel (Noah's Ark). Prior to the task, all animals were named, grouped by type, and kept in full view of the child. None of the animals were in or on the vessel at the start. This task was conducted with children aged 18–36 months.

Real-word plural task: Full and abridged version

The full and abridged versions of the real-word plural elicited-production tasks were designed to investigate the acquisition of the SP and BP in Maltese (see Section 2.5.1 for an account of the Maltese plural system). Nouns were presented by the researcher in the singular, supported by images, and the child was asked to produce the plural by completing a carrier phrase. The abridged version contained 10 items; the full version expanded this to 20 items to capture a wider range of plural suffixes and patterns. The target items were selected carefully so that, for example, words featuring the phoneme /tʃ/ were excluded from the set since, according to Grech and Dodd (2008), this phoneme is absent from the repertoire of children aged 36–47 months both those using Maltese only at home and those using Maltese and English at home. See Appendix J for the full list of items and the rationale behind their selection.

A set of slides was created using Microsoft PowerPoint. Each item corresponded to two slides: one slide with a single image of the item and the following slide with the same image replicated four times to depict multiple instances and, hence, elicit the plural. The task was piloted with two MD adults prior to child testing. The items were presented in the same order to all participants. The abridged task was conducted with children aged 36–48 months, and the full task with 36–54-month-olds. See Section 2.4.3 for more on production tasks employed in CLA studies.

Real-word plural task: Bingo

To reduce restlessness during sessions containing multiple elicitation tasks, the full plural task was later adapted into a bingo game and presented as a fun activity. See Section 2.4.3 for more on the adaptation of classic games for the purpose of language acquisition research. Two laminated bingo sheets (6 × 6 grid, 36 images) were created, incorporating 20 target items and 16 fillers.[45] The researcher and child alternated drawing and naming flashcards, marking corresponding images on their boards. When participants were in doubt or produced a different word from that targeted, the researcher provided the singular and then asked for the plural (as in the real-word plural task). This task was conducted with children aged 44–60 months.

Novel-word plural task

A novel-word plural task, modelled on Berko's (1958) *Wug Test* and Nieder et al.'s (2020) novel-word experiment with native Maltese adults, was designed to examine children's ability to pluralise novel nouns. The novel images[46] employed here were those used by Nieder et al. (2020), created for van de Vijver and Baer-Henney's (2014) study. In addition, all novel words were phonotactically legal[47] singular nouns in Maltese, constructed by Nieder et al. (2020) by systematically changing

[45] The filler items were presented in the singular form to avoid potential morphological priming effects.

[46] A degree of semantic bias cannot be fully excluded when interpreting results from this task, as the novel images vary in colour and shape (e.g., some resemble land animals, whereas others evoke marine creatures).

[47] Although all novel words were phonotactically legal in Maltese, some sounded less natural than others, particularly when both consonantal and vocalic segments were altered.

either consonants (C), vowels (V), or both (CV) of existing Maltese nouns. Phonotactic viability was considered crucial since research has demonstrated that children as young as 19 months distinguish between phonotactically legal and illegal words in their L1, treating only the former as potential lexical candidates (e.g., Friedrich & Friederici, 2005; Graf et al., 2011).

Twenty items were drawn from Nieder et al.'s (2020) corpus of over 700 novel words. Two items were used for the trial phase. The remaining eighteen test-phase items were equally divided into three sets of six (C change, V change, CV change) and were further subdivided into three pairs, depending on whether the existing base word took SP, BP, or both. One example of each set is provided in (1) – (3) below. For each pair, one noun is an unmarked feminine noun, and the other is an unmarked masculine noun. Words containing the phoneme /tʃ/ were excluded (see above). See Appendix K for the full list of items.

(1) Consonant change: *pannat* (*tallab* – *tallaba/tlaleb* 'beggars'; SP, BP)

(2) Vowel change: *tekkina* (*takkuna* – *tkaken* 'heels'; BP)

(3) Consonant and vowel change: *lepa* (*nota* – *noti* 'notes'; SP)

The puppet Bobo was used to introduce the scenario and maintain engagement. The researcher stressed that there was no right or wrong answer. When both this task and the real-word plural task were conducted within the same session, the novel-word plural task was conducted first so as not to risk having real-word plural items influencing responses. The task was conducted with children aged 54 months and older.

Opposites task

The opposites task explored gender agreement between nouns and adjectives. Various elicitation tasks have been used in CLA to reveal insights into gender assignment and agreement for several languages, such as Icelandic (e.g., Björnsdóttir, 2021), Spanish (e.g., Socarrás, 2011), and Irish (e.g., Nic Fhlannchadha & Hickey, 2017). Such tasks have also been employed to investigate factors influencing the acquisition of gender agreement, namely bilingual language exposure (e.g., Nicoladis & Marchak, 2011; Vorobyeva & Voeikova, 2024), specific language impairment (SLI; e.g., Leonard et al., 2001), and developmental language disorder (DLD; e.g., Tribushinina & Mak, 2022).

Twelve slides were used to elicit noun–adjective combinations (6 masculine, 6 feminine), each depicting 2 contrasting items (e.g., a big and a small ball; see Appendix L for the full list of items). The slides were presented to all participants in the same order and were ordered in such a way to avoid consecutive testing of masculine and feminine (or vice versa) adjectives and to include a mix of masculine and feminine nouns, rather than testing all masculine nouns followed by all feminine nouns. Children were prompted with *X'qed tara hawnhekk?* 'What are you seeing here?', followed by *X'hemm differenti bejniethom?* 'How are they different?' if further prompting was required. The nouns and adjectives selected were integral child-language vocabulary items[48] that were easy to depict. This task was conducted with children aged 36–42 months.

Spot-the-difference

This familiar activity was used to elicit spontaneous language involving colour, number, and negation. Spot-the-difference is widely used in elicited-production tasks, notably in L2 studies (e.g., Bagheri et al., 2012; Perpiñán, 2015). Eight pairs of laminated spot-the-difference sheets featuring cartoon-style scenes were used. Children were asked to describe as many differences as possible. The task was conducted with children aged 42 months and older and was presented as a playful challenge to sustain engagement.

Narrative Elicitation

As stated by Eisenbeiß (2006, p. 113), "a good documentation of child language must provide a basis for studies on narrative development and the linguistic socialization process that leads to conversational competence in a particular community". In this study, semi-structured narrative tasks targeting conversational competence and story construction were interspersed with play. Five age-appropriate tasks incorporating visual stimuli were developed. All materials were in colour except

[48] The selected items appear in one or both early-years Maltese vocabulary lists (Gatt 2010; Muscat et al., 2017). Although to date no corpus-based frequency estimates of word class or lexical frequency in Maltese CDS are available, see work by Gatt et al. (2023) for word-usage data relating to children aged 12–30 months.

Mayer's (1969) *Frog Story*; picture books and stories were physical, whereas animated videos and *Frog Story* were digital.

All narrative tasks were introduced with open prompts to encourage spontaneous responses. Typical examples included *X'qed jiġri hawnhekk?* ('What's happening here?') and *Ara issa, x'ġara?* ('Look now, what happened?'). Depending on the brevity and specificity of the replies, additional open questions and more specific questions were employed. Following recommendations in child language research, direct questions were avoided with children aged around 36 months and younger since such questions "tend to have an overwhelming and demotivating effect" with this age group (Kauschke et al., 2024, p. 5).

Younger participants (18–28 months). Wordless board books depicting domestic, farm, and jungle animals were used primarily to elicit vocabulary related to animals, actions, and attributes, as well as to prompt basic descriptive utterances. Board books are ideal for very young children as they are difficult to bend and tear, wipeable and easy for toddlers to turn over the pages. According to Applebee (1978), at 24 months or so children are at the first stage of story schema development, in which they are able to label and describe character actions, without however linking them to the story's theme. In addition, narrative ability typically emerges around 36 months (e.g., Gardner-Neblett, 2024; Khan et al., 2016; Piasta et al., 2018). Hence, the board books were used as early narrative precursors.

Older participants (42–60 months). For older children, four one-page picture stories and three short, illustrated Maltese storybooks – *Barnuża Hamra* (*Little Red Riding Hood*), *It-Tliet Qżieqeż* (*The Three Little Pigs*), and *Il-Prinċep Żrinġ* (*The Frog Prince*) – were used to elicit short narratives. The picture stories, retrieved from an open-access online source, each depicted a simple everyday event featuring the same young girl across four illustrated panels.[49] The three storybooks, forming part of a nine-book Maltese series,[50] presented classic tales adapted for early readers. The materials were comparable in length, visual complexity, and thematic appeal, and no rote retelling was observed.

Mayer's *Frog Story*. *Frog, Where Are You?*, a wordless picture book designed by Mercer Mayer in 1969, is "one of the best-known broad-spectrum elicitation tools in child language research" (Eisenbeiß, 2009, p. 6). The story forms part of a six-frog-book series and follows a boy and his dog on an adventure to find their escaped pet frog. In this study, it was presented digitally to participants aged 54 months and older. Children were provided with a toy microphone and asked to imagine narrating the story to a Maltese-speaking audience without visual access to the images, thereby encouraging detailed narration in Maltese. As stated by Gardner-Neblett (2024, p. 1223), "the frog storybooks are ideal for eliciting children's oral narratives as they portray a main character's attempts to achieve a goal and thus provide children with rich material from which to craft goal-oriented stories". Being

[49] These picture stories are freely available for personal use at http://clairedelvaux.blogspot.com/2019/06/des-nouvelles-de-nina.html. They were included in the study design because the depicted situations are culturally relevant to Maltese children, the illustrations are age-appropriate and engaging, and the material provides effective prompts for spontaneous conversation.

[50] The three books used form part of a nine-volume series of classic tales entitled *Il-Hrejjef Maġiċi* (*Magical Fairy Tales*), originally written by Paolo Valentino and illustrated by Valeria Abatzoglu, with Maltese translations by Noel Tanti.

one of the most widely used cross-linguistic narrative elicitation tools, its inclusion allows for potential cross-language comparison and eventual incorporation of Maltese into the *Frog Story* corpora (CHILDES).[51]

Animated videos. Four sound-free animated clips, each approximately 3–5 minutes long (e.g., from the Swiss–German series *Pingu*), were used with children aged 36 months and older. Silent visual stimuli have long been employed to elicit narrative production in language acquisition research, providing a standardised yet engaging context for assessing linguistic performance. Classic examples include the *Fish Film* (Tomlin, 1997), the *Pear Stories* (Chafe, 1980), *Charlie Chaplin* clips (e.g., Kim, 2014), and other silent videos (e.g., von Stutterheim et al., 2012). Recent findings highlight methodological advantages of animated over static stimuli. Diehm et al. (2020) reported that TD young children produced longer retells, demonstrated greater lexical diversity, and used more complex syntax when recounting animated stories compared with picture-book versions of the same narratives. They also found that adults provided fewer prompts and that children displayed greater engagement during video-based tasks. Accordingly, the inclusion of animated clips in the present study was expected to elicit richer, more structurally complex narratives and to reveal potential differences in children's responses to static versus dynamic visual stimuli.

[51] The *Frog Story* corpora are publicly accessible via https://childes.talkbank.org/access/Frogs/.

Having outlined the study's integrated design, combining longitudinal, cross-sectional and cross-participant components, and the complementary use of both naturalistic and experimental data collection methods, I now turn to the specific procedures employed to obtain the data. The following section details how the samples were collected from TD, MD children between 18–60 months.

3.4. Data collection procedure

Data collection spanned from March 2022 up until June 2024. During this period, 50 recording sessions were conducted, of which 40 were transcribed and analysed.[52] Each session lasted approximately 45 minutes to one hour and was recorded using one video camera and two audio devices: a ZOOM H2n Handy Recorder and a laptop (backup). As outlined in Section 3.3.2, both naturalistic and semi-experimental data were collected during these sessions. All recordings were conducted in the children's homes, typically in a play area familiar to them.

I arrived at each recording session with a Mary-Poppins-like bag in which all the toys and materials were stored. The recording equipment and documents to be completed by a parent were stored separately. Once parental consent and child assent had been obtained, the devices were set up. When semi-experimental tasks were scheduled, these were as much as possible conducted towards the start of the session. Overall, participants were highly cooperative and enthusiastic and looked

[52] Five recordings from one participant who developed OM and five from a pair of twins were excluded from transcription and analysis (see Section 3.2.2 for details).

forward to the sessions. Playmobil toys were all-time favourites among participants.

All recording sessions were carried out by the researcher in the presence of a parent, ensuring that participants felt secure and relaxed. No pressure was placed on the children at any point, and their wellbeing, safety, and comfort were prioritised throughout. A small non-consequential reward (e.g., a packet of stickers) was given to each child after every session in agreement with the parents.

Session timing varied according to family availability. Wherever possible, sessions were arranged when children were well-rested and alert. Parents could choose whether to participate or simply observe; if they participated, they were asked not to correct grammatical errors, answer on behalf of the child, or encourage rote-learned phrases. In one case, a sibling who was five years older than the target child joined recordings to facilitate interaction with a particularly timid participant.

Necessary measures were taken to minimise background noise, including turning off any noise-emitting devices and gadgets and avoiding toys likely to produce creaks or rattles. A number of play areas in the families' households were located in open-plan or street-level rooms where external noise (e.g., traffic, construction) could occasionally be heard. However, the presence of soft furnishings (e.g., carpets, curtains, sofas) typically absorbed echoes and dampened background interference.

At the end of the fifth and final recording session with each child, a brief parental speech sample was collected (see Section 3.2.3). These audio-only recordings, lasting no longer than ten minutes, were conducted with one parent who had been notified in advance and could opt out. The recording consisted of two parts: a short spot-the-difference task, followed by a semi-structured conversation exploring a) the child's primary language input models, b) dialect use and attitudes, c) language dominance and patterns of CS, and d) developmental concerns, both general and linguistic (see Appendix G).

The parental recordings served to a) follow up on questionnaire data, b) address topics not covered in the questionnaires (e.g., dialect attitudes), and c) provide a representative speech sample from (one of) the child's main language model(s). Note that these data do not constitute CDS and were not intended for full transcription or analysis unless features directly relevant to the child's linguistic output emerged (e.g., restricted phonetic repertoire or dialectal variation). After each recording, all toys were sanitised and reorganised for the next visit.

All data collected during this study were managed and stored in accordance with the University of Malta's Research Ethics and Data Protection Guidelines and the General Data Protection Regulation (GDPR). Each recording was assigned a unique identification code, and all personal identifiers were replaced with pseudonyms. Audio and video files, together with transcriptions and metadata, were stored in password-protected folders. Backups were maintained on secure institutional servers to prevent data loss. Hard copies of consent forms and questionnaires were stored separately in a locked cabinet.

3.5. Transcribing the data

All recordings, including spontaneous conversation, elicited narratives, and experimental data, were transcribed orthographically[53] using a format developed for this study and aligned with recognised standards in child-language research, particularly the CHAT (MacWhinney, 2000). CHAT, used throughout the TalkBank repositories, is compatible with CLAN software. CLAN programs "provide systems for the automatic checking of transcription accuracy, methods for the automatic analysis of morphology and syntax, and tools for the semi-automatic entry of codes" (MacWhinney, 2000, p. 20). To date, CLAN has not been applied to Maltese data. Only a subset of CHAT conventions (e.g., role labels such as *CHI* for *Target Child*) was implemented, but the format was designed to be readily convertible to full CHAT in future.

Each transcript corresponds to one recording session and begins with metadata on the participant, including pseudonym, chronological age, recording date, session duration, and total number of child utterances.[54] A brief contextual paragraph follows, describing the setting, the child's engagement, and any other relevant information. This is succeeded by a list of all speakers involved and a summary of the activities undertaken. The transcription layout comprises four columns:

[53] The data were not transcribed phonetically for the purposes of the current study. However, the recordings are of sufficient quality to permit future phonetic transcription and analyses focusing, for instance, on phonological development.
[54] See Section 3.6.1 for the working definition of *utterance* adopted in this study.

- Time stamp (five-minute intervals, facilitating reference to the recording);
- Speaker label (e.g., *CHI*, *INV*);
- Utterance (verbatim);
- Comments, where applicable (additional notes in English).

All transcripts were anonymised to protect participant identity. In the metadata, only pseudonyms appear; within transcripts, personal names (of the child, relatives, or pets) are represented by the initial letter followed by the relevant consonant-vowel pattern to preserve phonological information without revealing identity (see Appendix M for transcription guidelines). The guidelines also specify conventions for fillers, non-verbal vocalisations, baby words, and onomatopoeia to ensure consistency (e.g., *mm*, *wijj*). Criteria for segmenting speech into utterances were established and applied consistently, following recommendations by Eisenbeiß (2006) and others (see Section 3.6.1). [55]

Typically, the main portion of each session was transcribed, excluding brief settling-in or closing segments and any periods dominated by background noise. Audio data were processed in PRAAT (Boersma & Weenink, 2009), which allows playback at variable speeds and facilitates precise segmentation. Video recordings were also consulted regularly to clarify contextual cues lost in the audio (e.g., MacWhinney, 2000).

[55] Inter-pausal units (IPUs) were not used to segment utterances (see, for example, Bigi & Priego-Valverde, 2022; Nguyen et al., 2023), as this measure may potentially skew MLU calculations.

Figure 3.1 presents the 40 recordings transcribed and analysed, covering ages 18–60 months with a minimum of one recording every three months across this 42-month developmental span. Figure 3.2 illustrates the distribution of transcriptions within six-month age bands, showing the highest concentration between 36–41 months, corresponding to the overlap of early and late data-collection stages. Altogether, the dataset includes over 30 hours of Maltese child language recordings, averaging 46 minutes per session. On average, 7 minutes of each recording contain semi-experimental data, with the remainder representing spontaneous speech collected in naturalistic settings.

Figure 3.1 *Timeline Heatmap of Available Transcriptions Per Month from 18–60 months*

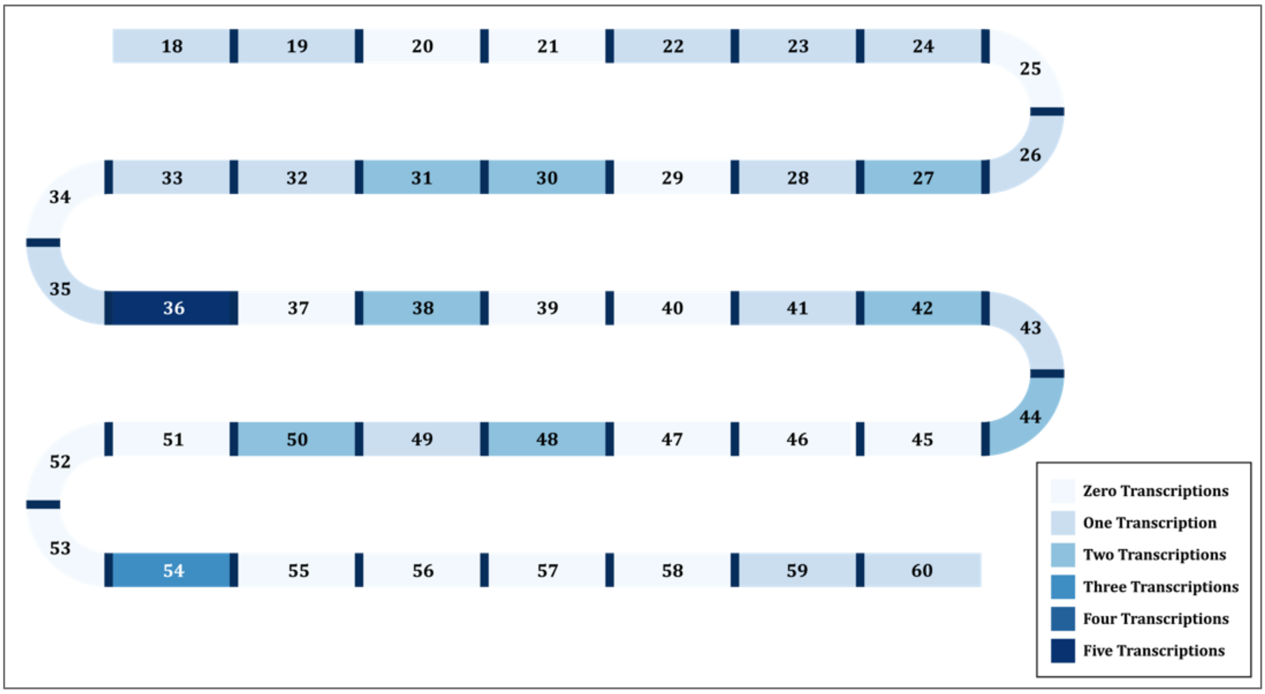
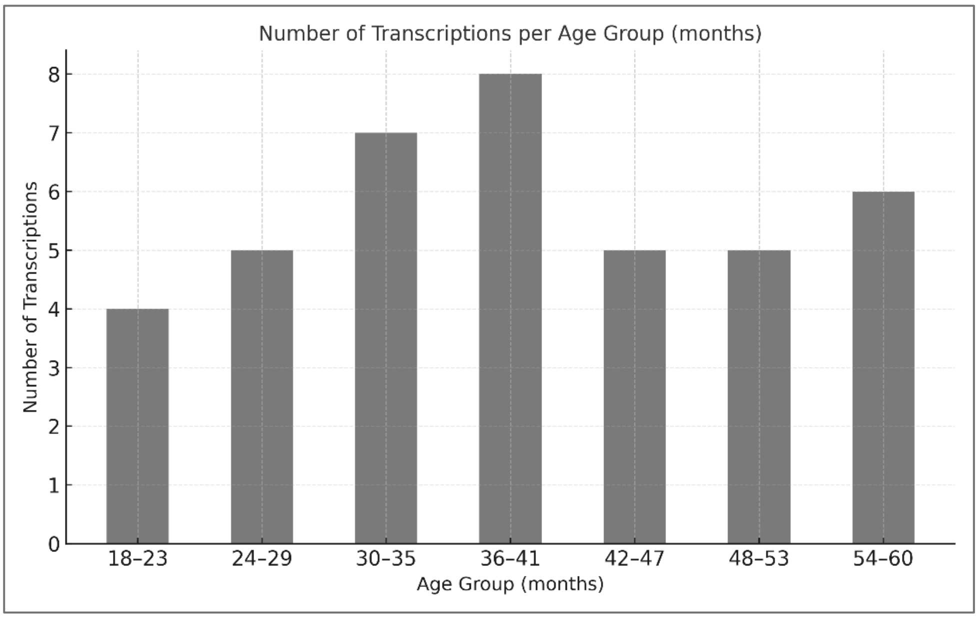


Figure 3.2 *Distribution of Available Transcriptions by 6-month Age Groups over the Entire Data Collection Period*



As MacWhinney (2000, p. 18) observes, “whatever form a transcript may take, it will never contain a fully accurate record of what went on in an interaction”. Nevertheless, supplementary annotation can enhance the fidelity of transcription. In this study, fourteen annotation conventions were adopted (see Appendix M for transcription guidelines), several drawn from Jefferson’s Conversation Analysis system and others from Vella et al. (2011). Given the time-intensive nature of child-language transcription, annotation was deliberately limited to features relevant to the study’s morphosyntactic focus to ensure manageability and consistency. While full translations are not provided for the Maltese child language transcripts (which at times include English items), all metadata and researcher comments are in English (see Appendix R for transcription excerpts). Translations and interlinear glosses are supplied in later chapters for any examples cited directly from the data.

3.6. Calculating MLUm

To address RQ1, which concerns early morphosyntactic development, MLUm was employed. MLUm is widely recognised in CLA research as a more sensitive indicator of linguistic development than chronological age and serves both diagnostic and analytical purposes (see Section 2.3.2 on MLU/MLUm in the context of Brown's (1973) work). Given the morphologically complex and hybrid nature of Maltese, MLUm offers a more accurate reflection of developing morphosyntactic competence than measures based on word counts (MLUw; see Section 4.3.1). MLUm is calculated by collecting a representative sample of utterances, counting the number of morphemes in each, summing the total, and dividing by the number of utterances.

$$\text{MLUm} = \frac{\text{Total number of morphemes}}{\text{Total number of utterances}}$$

However, using MLUm as an index of morphosyntactic development in Maltese raises two theoretically and methodologically loaded questions:

- What constitutes an utterance in child language?
- What constitutes a morpheme in Maltese?

A reliable and consistent method for segmenting continuous child speech into discrete utterances and identifying and counting morphemes is essential, particularly given the complex interaction of concatenative and non-concatenative morphological processes in Maltese. The methodological choices made in these areas carry significant implications since inconsistencies in utterance segmentation and counting morphemes can result in divergent outcomes, undermining both the comparability and interpretability of findings. Accordingly, rigorous standards are essential in ensuring intra- and inter-study reliability in both clinical and research contexts.

This section begins by outlining the principles guiding utterance segmentation and morpheme identification in child language, reviewing key criteria proposed in the literature and evaluating their implications for the reliability and validity of developmental analyses. The MLUm calculation protocol developed for the current study is presented in Appendix N.

3.6.1. What constitutes an utterance in child language?

In CLA, the utterance is the primary unit of transcription and coding of most child language corpora as well as the unit of measurement most widely used in metrics such as MLU. Yet, despite its central role and widespread use, the definition of what constitutes an utterance in child language is complex and controversial, particularly due to the fragmentary and evolving nature of children's speech. Hence, the utterance in its broadest sense, as a unit of speech bound by silence or change in speaker, requires re-evaluation in the light of child language. An utterance in the context of CLA is not defined by its grammatical structure, rather it is an operational construct based on a set of practical, observable criteria.

The most influential framework for defining and segmenting utterances in CLA was established by Brown (1973) in his seminal longitudinal study of the language of English-speaking children. Brown's criteria, developed for the purpose of calculating MLU, have become the *de facto* standard, having influenced corpus annotation practices and coding manuals such as those developed for the CHILDES database (MacWhinney, 2000). The CHAT transcription format used in CHILDES provides a systematic framework for utterance segmentation. According to the CHAT manual, an utterance is typically defined by a) a change in speaker, b) a pause exceeding a given threshold (often 2 seconds), or c) a clear prosodic boundary, that is, a complete intonational contour, such as a terminal fall for a declarative statement (MacWhinney, 2000). These criteria are a means of balancing pragmatic, syntactic, and phonological cues to determine utterance boundaries, though variation across languages and developmental stages often necessitates the judgement of the analyst. Thus, a consistent methodology for handling characteristic disfluencies and the inherent messiness of spontaneous child language is critical, especially when the utterance is a core unit of analysis.

In spontaneous speech, particularly among young children, utterances may be incomplete, ungrammatical, or consist of single words. This has led researchers to consider both structural and functional approaches to defining utterance boundaries. Bloom et al. (1975), for instance, argue for a function-based approach, emphasising the communicative intent of the child over grammatical completeness. In their study, utterances were defined as any stretch of speech that forms a distinct communicative act, regardless of syntactic integrity. This view aligns with the broader shift in developmental psycholinguistics from structuralist to functionalist perspectives (e.g., Tomasello, 2003). In the case of preverbal or early verbal children, further complications arise. Vocalisations such as babbling or protowords are often counted as utterances if they are used consistently with communicative intent, though researchers highlight the role of situational context and analyst interpretation in determining whether such productions qualify as utterances (e.g., Nelson, 1973).

The question of repetition also features prominently in the literature. Imitations of adult language are generally included as utterances if they are complete and immediately follow the adult model, though some researchers exclude such imitations when calculating MLU or lexical diversity, arguing that they do not reflect independent linguistic competence (e.g., Peters, 1983). Similarly, formulaic or rote-learned chunks may be treated as single units in early developmental stages, as such structures are not evidence of internal linguistic complexity and raise theoretical questions about productivity and compositionality. Additionally, partial or complete repetitions within a single speech turn (e.g., she walks *like... like a duck*) are considered part of a single utterance, though the repeated elements are typically excluded from the word/morpheme count for MLU calculation. In the same way, fillers such as *erm* and *hmm* are transcribed as part of the utterance; however, they are not included in the morpheme count (MacWhinney, 2000).

Several structural and functional approaches to utterance segmentation have been proposed; however, the present study adopts the three operational criteria defined in the CHAT manual (MacWhinney, 2000) and outlined earlier in this section: change in speaker, pause length, and prosodic boundary. Furthermore, a protocol for calculating MLUm[56] was developed. This protocol accounts for the complexities inherent in spontaneous child language data in relation to utterance segmentation and calculation, including repetitions, fillers, and false starts, as

[56] See above for the rationale behind the focus on MLUm instead of MLUw in this study.

discussed earlier. Appendix O presents a representative extract of this protocol, illustrating how specific language features were identified and scored.

3.6.2. What constitutes a morpheme in Maltese?

One of the central challenges in morphology is formulating a precise definition of *word* and *morpheme*. The term *morpheme* is commonly attributed to Jan Baudouin de Courtenay (1895), who defined it as "the part of a word which is endowed with psychological autonomy and is for the very same reason not further divisible" (Stankiewicz's (1972) translation, p. 153). Another canonical definition comes from Bloomfield (1933), who characterised the morpheme as a minimal form with a constant meaning. This classical view posits the morpheme as the smallest meaningful constituent of a word, an indivisible pairing of sound and meaning. Though influential, this formalist definition has been refined by generative and post-generative frameworks to account for morphologically complex and non-concatenative structures found across languages. Nonetheless, the concept remains controversial.[57] Anderson (2015, p. 32), for instance, questions "whether or not there is any such thing as a morpheme", underscoring ongoing debates about its theoretical validity. More broadly, Haspelmath (2020, p. 124) calls for "clarity and consistency" in morphological theory.

The notion of the morpheme is central to what Hockett (1958) termed *duality of patterning*, whereby a finite set of meaningless phonemes combines to form a vast inventory of meaningful morphemes, which in turn combine to create

[57] See, for example, Bauer (2016, 2019), Haspelmath (2020) and Anderson (2015).

words, phrases, and sentences. Hockett (1958) further advanced morpheme theory by distinguishing between segmental and suprasegmental morphemes and highlighting *allomorphy*, where a single morpheme has multiple phonological realisations. A morpheme is therefore an abstract grammatical entity, whose physical form is referred to as a *morph*. A morpheme can be realised by different morphs, depending on the phonological and morphological context, which are known as *allomorphs*. A classic example is the English regular plural morpheme, which has at least three phonologically conditioned allomorphs: /s/ as in *morph**s***, /z/ as in *morphem**es***, and /ɪz/ as in *languag**es***. An example from Maltese is the dual form morpheme, which has at least two allomorphs: /ɛjn/ as in *minkbejn* '(two) elbows' and /ɐjn/ as in *dirgħajn* '(two) arms'.[58]

Morphemes are generally classified as *free* or *bound*.[59] Free morphemes function independently as words, whereas bound morphemes must attach to others and cannot occur in isolation, as with the regular plural morpheme discussed above. Free morphemes divide further into *lexical* and *functional* types. Lexical (or content) morphemes are open-class items carrying substantial semantic content – typically nouns, verbs, adjectives, and adverbs. Functional (or grammatical) morphemes serve mainly structural or relational roles, including determiners and conjunctions. In child language research, functional morphemes are especially central, as their acquisition tends to be slower and more variable than that of lexical morphemes (e.g., Brown, 1973).

[58] See, for example, Haspelmath and Sims (2010) for a detailed account of allomorphy, and Bezzina (2020) for allomorphy in Maltese.
[59] See, for example, Fromkin and Rodman (2011).

Bound morphemes are likewise subdivided into *inflectional* and *derivational* types.[60] Inflectional morphemes do not alter the lexical category of the base word but express grammatical relations such as number, gender, or case. Derivational morphemes, by contrast, participate in word formation, changing a word's category or meaning, for example, the English suffix *-er* in *painter* and the Maltese prefix *arċ-* in *arċisqof* 'archbishop', attached to *isqof* 'bishop'.

In languages with root-and-pattern morphology, the linear assumptions of morpheme-based models pose challenges for morphological segmentation. Here, consonantal roots interdigitate with vocalic templates to produce meanings and grammatical categories, rendering morphemes less easily isolable. Theoretical models such as Prosodic Morphology (McCarthy & Prince, 1990) and Morphological Realisation theories (Stump, 2001) have been developed to address these limitations.

While the definition and analysis of the morpheme have evolved over time to accommodate typological diversity and theoretical innovations, the morpheme endures as a central unit of form-meaning mapping and an empirically fruitful concept in linguistic theory. In present day research, the morpheme continues to be an essential analytical unit in linguistic description and acquisition studies. Within the realm of language acquisition, the morpheme serves both as a unit of measurement (e.g., Brown's (1973) MLUm) and as a focus of morphological development. The study of how children acquire bound morphemes, namely functional ones, has provided key insights into cognitive and linguistic maturation,

[60] See, for example, Lieber (2020).

the role of input, and cross-linguistic variability (e.g., Engelmann et al., 2019; Slobin, 1985).

Researchers investigating language acquisition in Semitic languages often begin with the rules established for calculating MLUm in English (Brown, 1973), subsequently developing language-specific adaptations to better capture developmental patterns in morphologically rich contexts. For example, Dromi and Berman (1982), in their study of child language in Modern Hebrew, introduced the metric *Morpheme Per Utterance* (MPU), highlighting morpheme count, rather than utterance length, as a more appropriate indicator of morphosyntactic complexity in such languages. Similarly, Portelli (1997, 2005) adapted Brown's (1973) MLUm framework to Maltese.

As stated by Tallas-Mahajna and Dromi (2023, p. 131), "the bound character of Hebrew and Arabic morphology, compared to English morphology, is such that linguistic units which fall under only syntax in English, correspond in these Semitic languages to both morphology and syntax". This observation also applies to Maltese, as illustrated in examples such as *idejna* /ɪˈdɛjnɐ/ 'our hands' and *m'għamilnihomlhiex* /mɐmɪlnɪjɔmˈlɪːʃ/ 'we did not do them for her'. Although a spoken corpus of Maltese is not yet available, with one currently in development, these examples reflect naturally occurring usage in spontaneous speech.

The measure used for calculating MLUm in the present study (see Appendix N) is adapted from the Arabic-MPU framework developed and tested by Tallas-Mahajna and Dromi (2023) for Palestinian Arabic, with adjustments to accommodate features specific to Maltese morphology. This model, itself an adaptation of the MPU metric originally proposed by Dromi and Berman (1982) for Hebrew, employs a feature-score system in which morphemes are assigned values ranging from one to three points. The Arabic-MPU offers a well-defined framework, particularly suited to Maltese which shares morphosyntactic features with Palestinian Arabic. From the standpoint of usability and replicability, the Arabic-MPU presents a systematic and practical approach, with logically structured handling of exceptions. A few minor adjustments were necessary to tailor the framework more precisely to the morphological properties of Maltese. These adaptations aim to preserve consistency and clarity in the application of scoring rules. Importantly, adopting the feature-score model used by Tallas-Mahajna and Dromi (2023) facilitates comparability between the present findings and their results on Palestinian Arabic language acquisition. This comparison, however, must be interpreted with caution due to key differences in participant profiles – most notably, their study involved monolingual children, whereas the present study examines a bilingual cohort.

The current scoring method also bears similarities to that proposed by Portelli (1997, 2005) for calculating MLUm in Maltese, whose framework represents a valuable and language-specific contribution that considers important morphological features of Maltese. However, when applied to the data in the present study, some aspects of the scoring system used by Portelli, such as ambiguities and relatively inflated scores, posed challenges for consistency and comparability. Therefore, an alternative scoring model was adopted. By providing four examples, Table 3.4 illustrates differences between Portelli's framework and the present scoring system, highlighting how certain elements, such as plural pronouns and 'unmarked' verb forms, receive different values.[61]

Table 3.4 *Differences between Portelli's (1997, 2005) Scoring System and that of This Study*

Linguistic element	Example	Portelli (2005)	Present study
Personal pronoun	*aħna* 'we'	2	1
Demonstrative pronoun	*dak* that.SGM 'that'	2	1
'Unmarked' verb	*rebaħ* won.SGM.PFV 'he won'	4	1
Noun with possessive suffix	*sieqek* foot-2SG.POSS 'your foot'	2	2
	sieqkom foot-2PL.POSS 'our hand'	3	2

To further examine the practical impact of these distinctions, a small subset of child utterances (n = 20) was rescored using both Portelli's and the current system. The comparison (Appendix O) shows that Portelli's method consistently produced higher MLUm values. Nevertheless, MLUm values from the present study are discussed alongside Portelli's for contextual reference, given that it remains the only prior MLUm study on Maltese, with the acknowledgment that the underlying metrics differ substantially.

[61] Examples are illustrative and not exhaustive of either system.

That said, the chosen scoring model is not without its limitations. Two potential areas for critique include:

- The simplification of complex categories, such as assigning both the SP and BP a score of 2 (e.g., *tapit* - *tapiti* (SP), *twapet* (BP) 'carpets'). See Section 2.5.5.1 on nominal inflection in Maltese, including the plural system.
- The potential for scorer subjectivity as certain rules allow for interpretation and full objectivity is not guaranteed. For instance, although the 3rd person masculine perfective form is treated here as the unmarked verb form with a score of 1 (see Table 3.4), one could argue that its overt marking merits a higher value.

These limitations do not undermine the validity of the method but rather exemplify a classic trade-off in linguistic analysis: the balance between descriptive granularity and practical applicability.

MLUm, as a metric, inherently involves a degree of subjectivity, particularly in languages like Maltese that combine concatenative and non-concatenative morphological processes. Fundamental concepts such as *morpheme* and *utterance* are themselves often fluid and context-sensitive, as discussed above. Therefore, once a scoring system is selected, it is incumbent upon the researcher to maintain consistency throughout, especially in instances where the scoring involves arbitrariness. Even Brown's (1973) original MLUm framework – despite being foundational and widely used for English has not been immune to critique. As stated by Crystal (1974), Brown's method presents ambiguities and lacks clarity in certain areas, often requiring researchers to make judgement calls during analysis. In the present study, consistency was as much as possible ensured by having all MLUm scores assigned by the same researcher (i.e., myself).[62] Broader implementation of this system would benefit from the development of a comprehensive scoring manual, including detailed guidelines and a wide range of examples, to support replicability and minimise scorer-related variability. Such a resource would ideally be established and maintained collaboratively within the Maltese linguistic research community, ensuring a shared and consistent standard across studies.

3.6.3. MLUm calculation strategies

To analyse the progression of MLUm, as specified in RQ1, a total of 14 transcriptions (3,898 child utterances after cleaning) were selected from the larger dataset comprising 40 recordings, each between 30–45 minutes in length. Manually calculating MLUm of child utterances in all 40 sessions would have been too time-consuming for the purposes of this research project. The selected subset reflects the study's hybrid data collection design, enabling both developmental tracking and cross-participant comparisons.

Eight transcriptions (2,122 child utterances after cleaning) were selected to capture development across the full age range of 18–60 months, based on data from 6 children. This subset includes pairs of recordings from children recorded at 18, 32, 44 and 60 months, and at 32 and 44 months, respectively. These samples collectively provide a developmental trajectory while incorporating some true

[62] Refer to Section 3.7 for a note on inter-rater reliability.

longitudinal data, since the set selected at 32 and 44 months are the same children, as illustrated in Table 3.5 below.

Table 3.5 *Participants Selected for Cross-sectional Tracking*

	18 months	32 months	44 months	60 months
Cross-sectional Tracking	Anna	Karmena	Karmena	Kelina
	Stiefnu	Lippu	Lippu	Reno

In addition, 6 transcriptions (1,776 child utterances after cleaning) were selected from all children for whom data were available at 36 months of age, offering a cross-participant snapshot at a key developmental stage, as illustrated in Table 3.6 below. Notably, the children recorded at 18 and 60 months, and selected as part of the developmental tracking sample, also form part of this 36-month group, making it possible to track within-child development between their earliest and latest recordings (i.e., 18–36 months and 36–60 months, respectively). This dual approach allows for both between-child comparisons across participants at the same age, and within-child comparisons over time, thereby enriching this analysis of morphosyntactic development.

Table 3.6 *Participants Selected for Cross-participant Snapshot*

	36 months					
Cross-participant snapshot	Anna	Stiefnu	Duminku	Toninu	Kelina	Reno

The data from all 14 orthographic transcriptions, manually transcribed and annotated as described in Section 3.5, were cleaned prior to MLUm calculation, following the extraction of child utterances. Transcription, annotation, data cleaning, and calculation of MLUm were performed manually to ensure maximal accuracy and consistency. The cleaning process involved the removal of yes/no responses and full adult repetitions (adjacent imitations), among other non-lexical or non-spontaneous items. A detailed list of exclusion criteria and examples is provided in Appendix M.

Cleaning was applied at the utterance level: only entire utterances that met exclusion criteria were removed from the dataset. The internal structure of remaining utterances was preserved; elements such as fillers or repetitions occurring within an utterance were retained in the transcription but excluded from the morpheme count during MLUm calculation. Hence, in calculating MLUm, the following were assigned a score of 0: a) fillers (e.g., mm), b) non-verbal elements (e.g., <sigh>), c) repeated segments within a child's own utterance (e.g., in *she walks like… like a duck*, 'like…'=0), d) unintelligible speech, e) passages deemed by the researcher to be rote-learned (e.g., nursery rhymes), and f) partial repetitions of the adult's preceding utterance (e.g., adult: *he climbed up the tree*; child: *he climbed up there and fell down* – the repeated segment 'he climbed up' receives a morpheme score of 0).

All exclusion decisions made during the cleaning and analysis of the data are consistent with established protocols in MLUm research. Elements such as fillers, false starts, noises without semantic content, and repeated words are commonly excluded from MLUm calculations, as they do not contribute meaningfully to an understanding of the child's morphosyntactic development.[63] Oosthuizen and Southwood (2009), in their study on Afrikaans-speaking children and MLU calculation issues, emphasise the methodological implications of including or excluding certain elements, demonstrating that even minor inconsistencies can result in significant variations in MLU outcomes. They advocate for clearly defined and consistently applied criteria, particularly when working with spontaneous child language data. The approach adopted in the present study therefore reflects widely accepted practices within the field and ensures the reliability and comparability of the resulting MLUm measures.

Data from the real-word (slides/bingo), novel-word plural, and opposites tasks were excluded from the MLUm calculation. These tasks tend to elicit highly specific, patterned constructions that substantially influence MLUm values and do not reflect the child’s spontaneous language production. All other tasks, such as narrative elicitation and receptive tasks, were retained as they yielded spontaneous speech. In addition, utterances containing instances of language mixing were included in MLUm computations.

Once the above exclusion criteria were applied and a clean dataset of child utterances was produced, MLUm was calculated, in accordance with the scoring system largely adapted from Tallas-Mahajna and Dromi (2023) and presented in Appendix N, as discussed earlier in Section 3.6.2. While it is widely accepted that a

[63] See Brown (1973) and Miller and Chapman (1981) for more traditional and widely used cleaning criteria, and Johnston (2001) for an alternate method.

sample length of approximately 100 utterances (Brown, 1973), and in some cases fewer (Miller & Chapman, 1981), is sufficient to obtain a reliable MLUm count, all available utterances from the selected recordings were included in the calculation. This decision was motivated by the relatively small number of recordings (n = 14) and by the fact that including the full dataset increases representativeness and reduces sampling variability.

The mean number of utterances per recording was 278, substantially exceeding the commonly cited minimum threshold of approximately 100 utterances for reliable MLUm estimation (see above). Longer samples were methodologically preferable to increasing the number of recordings, as the most time-intensive stage of the process was not the MLUm computation itself but the data-cleaning phase, particularly the identification and removal of partial and full repetitions. The results corresponding to the MLUm analytical method described above are presented in Section 4.2.1.

In addition to MLUm, MLUw was also calculated for a subset of the data. For MLUw, eight transcriptions (2,156 clean child utterances) were selected; these transcriptions are also included in the analysis of MLUm. Together, they provide true longitudinal data spanning the full 18–60-month period (see Table 3.7). The youngest set of children, Anna and Stiefnu, were first recorded at 18 months and last at 36 months, while the oldest set, Kelina and Reno, were first recorded at 36 months and last at 60 months.

Table 3.7 *Participants Selected for Longitudinal Tracking*

	18 months	**36 months**	**60 months**
Longitudinal tracking	Anna	Anna	Kelina
	Stiefnu	Stiefnu	Reno
		Kelina	
		Reno	

MLUw was calculated alongside MLUm in order to enable a direct comparison between the two measures, particularly given that Maltese is characterised by both concatenative and non-concatenative morphology. In addition, the inclusion of MLUw provides a complementary metric that does not rely on prior assumptions about morphological segmentation**.** At the earliest stages, MLUm and MLUw are expected to yield comparable values (MLUm ≈ MLUw), as most lexical items consist of single morphemes. As children's morphosyntactic skills expand, however, MLUm would be expected to rise more steeply than MLUw (MLUm > MLUw), reflecting increased use of morphological processes in marking inflection and participating in agreement, among others.

While MLUw is often described as being automatically extractable from orthographically transcribed data, in practice additional preprocessing was required to ensure comparability and accuracy. The transcriptions were therefore subjected to further cleaning prior to calculation. This involved the removal of non-lexical annotation conventions, including pause markers (.) and non-speech/paralinguistic material, which would otherwise inflate word counts.

Following this cleaning stage, MLUw was calculated as the mean number of orthographic words per utterance.

A key methodological decision when calculating MLUw concerns the definition of a *word* in morphologically complex contexts. In line with standard practice in the field and following established approaches in both English and Semitic CLA research (e.g., Brown, 1973; Dromi & Berman, 1982; Ntelitheos & Idrissi, 2017), an orthographic criterion was adopted: units separated by whitespace were treated as individual words. This entails that morphologically bound elements written as a single orthographic unit were counted as one word. This decision has the following consequences:

- The Maltese definite article was counted as part of the noun (e.g., *it-tfal* 'the children'= one word), consistent with treatments of bound definiteness markers in Semitic languages.
- Fused preposition–article–noun forms (e.g., *għat-tfal* 'for the children') were counted as one word when realised orthographically as a single unit.
- Sequences written as separate orthographic words, such as preposition + NP where no fusion occurs (e.g., għal tfal 'for children'), were counted as multiple words.
- In code-switched English material, contractions (e.g., *I'm*, *doesn't*) were counted as single words, in line with standard MLUw practice in English, where orthographic words rather than underlying grammatical components are counted.

This orthography-based approach was adopted for reasons of transparency, replicability, and cross-linguistic comparability. Once transcripts had been standardised, MLUw was consistently across the dataset without requiring language-specific morphological segmentation.

3.7. Corpus compilation and POS tagging procedure

To fully address RQ1, POS tagging was conducted in addition to MLUm calculation to examine POS category distribution in TD, MD children aged 18–60 months. For this purpose, I compiled a corpus of 18 clean transcriptions, following the protocol described in Section 3.6. Fourteen of these transcriptions were the same as those used for the MLUm calculations, with an additional 4 included to balance the dataset by age. This adjustment was necessary because 6 of the original 14 transcriptions represented 36-month-old children, selected primarily to capture a cross-participant snapshot. In total, the corpus comprises over 5,200 clean utterances of Maltese child language. The corpus served two primary purposes: a) to identify the most frequent tokens of nouns, verbs, and adjectives in both Maltese and English, acknowledging the bilingual nature of the data while recognising that all participating children were reported to be MD, and b) to examine the range and frequency of semantic categories.

In addition to determining frequency distributions for the three target POS and selected semantic categories, I sought to trace the developmental trajectory of these POS over time. To achieve this, I applied manual POS tagging to longitudinal data. Given the time-intensive nature of manual tagging, such tagging was limited to eight transcriptions (2,122 child utterances after cleaning). The eight transcriptions selected for tagging and analysis were ones previously selected for developmental MLUm analysis: the study's youngest set, Anna and Stiefnu, first recorded at 18 months and last recorded at 36 months, and oldest set, Kelina and Reno, first recorded at 36 months and last recorded at 60 months.

Recognising the methodological challenges associated with manual POS tagging, particularly in a child language corpus, a set of tagging guidelines was developed to ensure consistency, accuracy, and transparent documentation. Adherence to best practices in corpus annotation was considered essential for both reliability and usability. A streamlined five-tag scheme was devised to facilitate consistent and replicable POS annotation while accommodating the specific characteristics of Maltese child language data. For the purposes of this study, the system focused on nouns, verbs, and adjectives, tagging only these three parts of speech explicitly, alongside two additional labels, *others* and *child-language specific*, to capture remaining tokens. The tagging of nouns and adjectives proved particularly valuable for subsequent analytical stages. The guidelines were designed to balance analytical precision with practical feasibility, given the constraints of manual tagging. Tag descriptors, inclusion and exclusion criteria, and illustrative examples for each category are summarised in Table 3.8. A sample tagged transcription following this system is provided in Appendix P.

Table 3.8 *POS Tagging Categories and Criteria, with Examples from the Data*

TAG LABEL	DESCRIPTION	INCLUSION CRITERIA	EXCLUSION CRITERIA	EXAMPLE
NOUNS	Lexical/content nouns	Common nouns referring to entities, objects, or concepts	Proper nouns	*mejda* 'table' *frog*
VERBS	Lexical/content verbs	Main verbs expressing actions, states or processes	Auxiliaries, pseudoverbs, participles	*irrid* 'I want' *jumped*
ADJECTIVES	Descriptive adjectives, passive participles (functioning adjectivally)	Words modifying nouns to denote qualities or attributes	Proper adjectives derived from names	*kbir* 'big' *yellow*
OTHERS	All other word classes	Function words (e.g., determiners, pronouns), adverbs, auxiliaries, pseudoverbs, active participles		*ma'* 'with' *quickly*
CHILD-LANGUAGE SPECIFIC	Fillers, onomatopoeia, ideophones/prewords conveying meaning	Non-standard forms and sound-symbolic words produced by children		*aħħ* (danger) *wuff* (bark)

CDIs were not employed in the present study, as the analysis centres on the distribution of POS rather than lexical development. Similarly, type–token ratio measures were not considered appropriate, given that the focus lies on relative category proportions rather than lexical diversity or productivity. Instead, proportional use of POS categories was examined, and within the three principal categories – nouns, verbs, and adjectives – the most frequent tokens were extracted to characterise category composition. The results for POS category distribution are presented in Section 4.2.2.

A Note on the Use of Advanced Technologies and Inter-Rater Reliability

Preliminary trials with several advanced speech and language processing tools were conducted to assess their suitability for Maltese child language data.[64] However, across all stages of data preparation and analysis – transcription, MLUm calculation, and POS tagging – automated methods proved unreliable or inefficient.

For transcription, trials with commercial speech-to-text models revealed high rates of recognition error and extensive post-editing requirements, making manual transcription more accurate and time-efficient. For MLUm calculation, recent computational approaches to automatic utterance segmentation have shown promising results for well-resourced languages (e.g., Cabiddu et al., 2023; French et al., 2011). However, comparable tools for Maltese are lacking, as large, annotated corpora are required to train accurate models. For POS tagging, existing Maltese taggers perform well on formal texts but underperform on child data. Their training

[64] See Gatt and Čéplö (2013) on electronic resources for Maltese.

on adult, standardised registers leads to frequent misclassifications when applied to data containing non-target-like forms, morphological simplifications, and CS. Trials with automatic POS tagging tools confirmed these limitations, as post-processing demands outweighed potential benefits.

Regarding inter-rater reliability, engaging multiple annotators and calculating inter-rater or inter-tagger agreement is standard practice for ensuring consistency and rigour in the annotation process. However, time and funding constraints made this unfeasible. Therefore, I performed all MLUm calculations and POS tagging myself, following the protocols and guidelines described earlier to ensure the uniform application of criteria throughout the dataset.

3.8. Analysing nominal morphosyntax and developmental timing

Analysis of nominal inflection and agreement constituted the next stage of investigation, corresponding to RQ2. Insights from the POS tagging conducted for RQ1 were instrumental in facilitating a systematic assessment of these morphosyntactic features. In relation to RQ3, the analysis did not require new coding procedures; instead, it entailed synthesising the principal findings – those concerning MLUm, POS category distribution, and nominal inflection and agreement – to explore their interrelation within the developmental progression of TD, MD children.

With respect to nominal inflection, in the present study I focus on number marking only, given that the hybrid morphology of Maltese plays a central role in shaping the plural system (see Section 5.1 – Hypothesis 2A). Nominal gender inflection is therefore excluded from the present analysis, both because it is only sparsely represented in the dataset and in order to maintain a focused investigation of number inflection rather than an overly broad treatment of multiple morphological features. Results pertaining to nominal inflection are presented in Section 5.2.1.

With respect to nominal agreement (see Section 5.1 – Hypothesis 2B), I restrict the analysis to agreement between nouns (or free-standing pronouns) and demonstratives, adjectives, and quantificational elements (including numerals) that occur in the nominal domain and can trigger or reflect agreement features (e.g., *wieħed* 'one.SGM' – *waħda* 'one-SGF' – *ħafna* 'all'). Both attributive and predicative adjectives are included, given the relative scarcity of adjectives in the naturalistic data collected. Moreover, adjective-like forms, such as passive participles (e.g., *mimli* (SGM) – *mimlija* (SGF) – *mimlijin* (PL) 'full'), are also treated under this category.[65] Noun–verb agreement (both subject and object agreement) is not included as it involves additional morphosyntactic and discourse-pragmatic mechanisms that extend beyond the nominal domain and require a separate, in-depth treatment. The corresponding results on nominal inflection are discussed in Section 5.2.2.

To address RQ2 and RQ3 and test the corresponding hypotheses (see Section 5.1), both naturalistic and experimental data were analysed. For the naturalistic data, the eight transcriptions (2,156 clean child utterances) used for developmental tracking in the analysis of RQ1 were selected for in-depth analysis (see Table 3.7).

[65] See Bartolo (2021) for an account of the participle in Maltese, including its adjectival function.

The discussion section draws examples both from this core set of analysed transcriptions and from additional transcriptions not used for results extraction (see Section 5.3).

The experimental data were obtained from 3 elicitation tasks administered to children aged 36 months or older: an existing-word plural task, a novel-word plural task, and an opposites task. The first 2 tasks target nominal number inflection, while the opposites task focuses on gender agreement. The tasks were conducted informally in the home setting, embedded within free-play sessions, and were designed to supplement the naturalistic data. Further details of the task designs and administration are provided in Section 3.2.2.2.

A note about non-target-like productions

Here a brief note on terminology is warranted. Throughout this thesis, the terms *target-like* and *non-target-like* are used as the preferred descriptors of children's productions. The labels *correct/incorrect* or *errors* also appear; these terms are used interchangeably and carry no theoretical implications.

As part of the analysis of nominal inflection and agreement, I conducted an analysis of non-target-like productions, focusing on the systematicity of such productions as evidence of the learning process. Non-target productions are widely recognised as an important window into the developing grammatical system, as they often reveal the rules or patterns that children are in the process of acquiring. Research in CLA, across both naturalistic and elicited data, has shown that such productions are often systematic rather than random, reflecting children's attempts to identify patterns in the input and generalise them.

Among the most extensively studied non-target patterns are overgeneralisations, in which children successfully extract a grammatical rule but extend it to forms that do not conform to the rule (e.g., Marcus et al., 1992). Overgeneralisation has been documented across a range of languages (e.g., Hržica, 2012), linguistic structures (e.g., Ambridge et al., 2012), and methodological approaches (e.g., Bidgood et al., 2021). The emergence of such forms reflects children's active engagement with the grammatical system and their increasing willingness to apply emerging rules productively.

To assess the systematicity of non-target-like productions, all instances of agreement and inflection in obligatory contexts were extracted from both the naturalistic and experimental data, and non-target-like productions were coded by type. Here, the term *obligatory contexts* refers to instances where the grammar requires overt marking, with a minimum of 5–10 such contexts serving as the accepted threshold in the field. This procedure also enabled the calculation of accuracy and stability, two key measures relevant to Hypotheses 2A–2C (see Section 5.1). Accuracy is calculated as the proportion of correctly produced tokens relative to all tokens (i.e., correct + errors) within the context being analysed (e.g., total plural contexts in the case of accuracy of the plural).

$$\text{Accuracy} = \frac{\text{Target-like tokens}}{\text{Total tokens in context}}$$

In the literature, a feature system is typically considered stable once accuracy in all obligatory contexts reaches and remains at ≥90% for three consecutive sessions, with sensitivity checks conducted at the 85–95% range. This operational criterion, first proposed in Brown (1973) and widely adopted since (e.g., Davison & Hammer, 2012; Padilla, 1978), provides a clear benchmark for developmental attainment (see Section 2.3.2).

While many of the non-target forms in the data can be understood as overgeneralisations or developmental misanalyses, it is important to acknowledge that not all deviations from the adult norm are necessarily 'errors' in a strict sense. Research has shown that children frequently engage in creative and playful manipulation of language, inventing novel forms or deliberately producing unconventional utterances, particularly in contexts of pretend play or humour (e.g., Crystal, 1998; Nwokah & Graves, 2009). This suggests that some of the forms documented here may reflect not only limitations in linguistic competence, but also children's experimentation and enjoyment of language as a symbolic resource. Such instances support the view of acquisition as an active process in which children are not merely passive recipients of input but creative agents who explore the expressive potential of their language system (e.g., Berko Gleason, 1958; Clark, 2009).

In essence, creativity and errors are two sides of the same coin: the child's active, constructive journey to mastering a complex linguistic system. It is challenging for researchers to definitively determine a young child's intent. However, the predictable patterns of these non-target-like productions across different children and even different languages strongly suggest a developmental origin rather than a purely creative or humorous one.

3.9. Concluding Remarks

This chapter has outlined the methodological framework adopted for investigating MFLA, detailing ethical considerations, the study design, participant recruitment, data collection and transcription procedure and analytical methods employed. In particular, it has described the calculation of MLUm as the primary measure of early morphosyntactic development and the POS tagging framework adapted for analysing MD children's productions. The subsequent chapters present the results and assess the hypotheses (Sections 4.1 and 5.1) associated with the three RQs (Section 2.5.5).

4. MLUm and POS category distribution

This chapter addresses RQ1 by investigating the progression of MLUm and the distribution of POS categories, considered both independently and jointly, to characterise early morphosyntactic growth in TD, MD children. The chapter begins by outlining Hypothesis 1 and its subhypotheses, which together predict systematic associations between MLUm, age, and POS category distribution (Section 4.1). The ensuing sections present results for each subhypothesis (Section 4.2), followed by a discussion that interprets these findings within the broader framework of early language development in Maltese (Section 4.3).

4.1. Hypothesis 1

Guided by RQ1, this section outlines the hypotheses and subhypotheses underpinning the analysis of MLUm and POS category distribution.

Overarching hypothesis

The developmental trajectory of early morphosyntactic acquisition in MD children is characterised by a significant positive correlation between MLUm and shifts in POS category distribution. Specifically, higher MLUm is predicted by an increase in the proportion of verbs and grammatical function words relative to nouns in the structural composition of child language.

Subhypotheses

The overarching hypothesis gives rise to the following more specific subhypotheses:

1A: MLUm will correlate positively with chronological age, with developmental trajectories comparable to those reported in Brown (1973), Portelli (2005), and Tallas-Mahajna and Dromi (2023), which underpin the MLUm calculation strategy adopted in the present study, particularly within the developmental window of 30–48 months.

1B: A noun bias[66] is expected in the youngest age group, with its prevalence gradually decreasing over time as the proportion of verbs and other non-nominal elements increases.

1C: An increase in MLUm is predicted to correlate negatively with the proportion of noun tokens, such that higher MLUm values will be associated with fewer nouns relative to other POS.

By tracking MLUm and POS category distribution in tandem, it will be possible to capture both quantitative increases in utterance length and qualitative shifts in structural composition, thereby offering a more comprehensive account of early morphosyntactic development than were either measure to be considered alone.

[66] In this study, the term noun bias is used descriptively to denote a higher proportion of nouns in early productions, rather than as a claim about universality or discrete developmental stage.

4.2. Results

4.2.1. MLUm

This section presents results based on the dataset outlined in Section 3.6.3, encompassing both the longitudinal component, eight recordings charting development from 18–60 months, and the cross-participant component, consisting of six recordings at 36 months, which serve as a point of comparison for the developmental analysis. The subsection opens with an outline of the findings addressing Hypothesis 1A (Section 4.2.1.1), proceeds to discussing sex-based differences (Section 4.2.1.2), and concludes with a note on CS (Section 4.2.1.3).

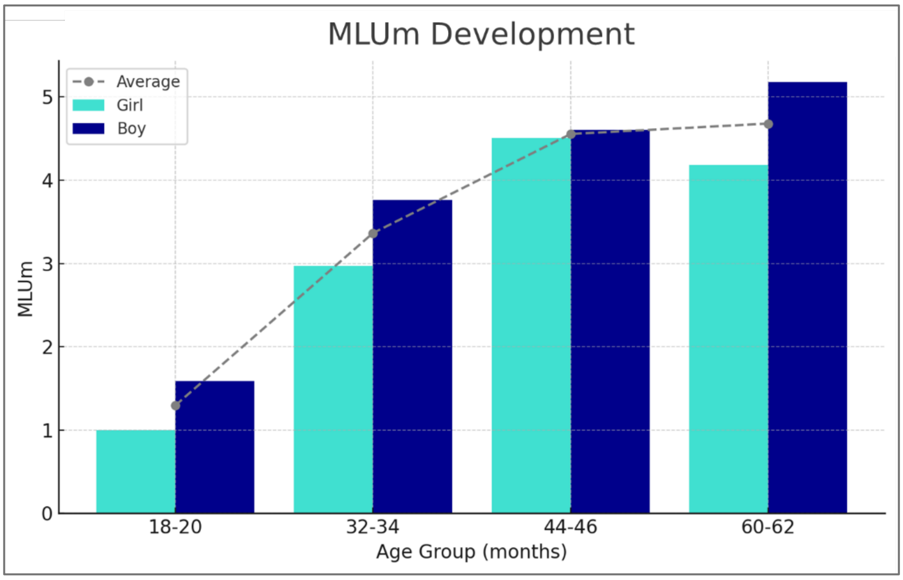
Although sex-based differences and CS are not examined as formal hypotheses, descriptive patterns are reported for both. Given the small sample size and the fact that these variables are not the primary focus of the study, no formal statistical testing is undertaken. Nevertheless, as both areas are understudied in the Maltese context, these observations are presented as exploratory, offering directions for future research.

4.2.1.1. Hypothesis 1A: MLUm and chronological age

Hypothesis 1A predicts that MLUm will correlate positively with chronological age, with developmental trajectories comparable to those reported in Brown (1973), Portelli (2005), and Tallas-Mahajna and Dromi (2023), particularly within the developmental window of 30–48 months.

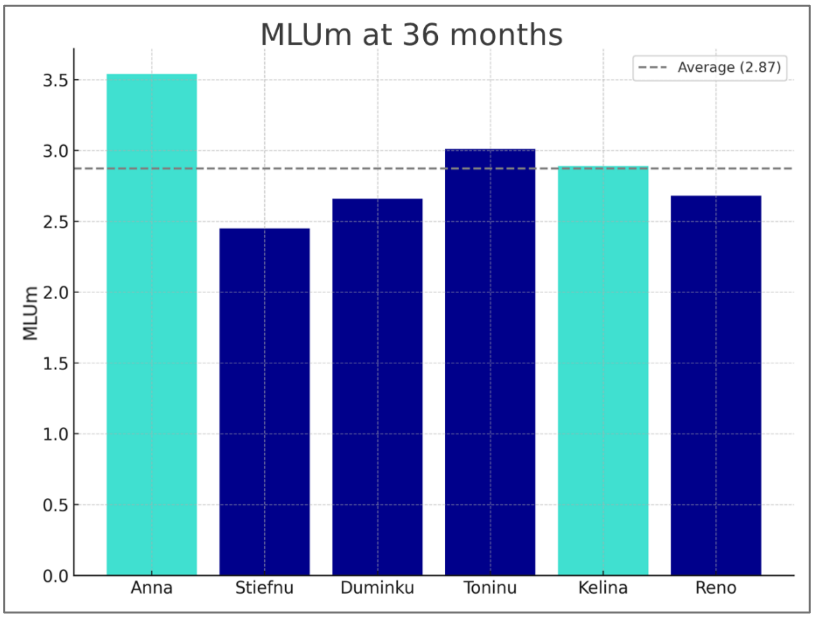
MLUm development was examined using both longitudinal and cross-participant datasets. In the developmental cohort (Figure 4.1), as shown by the average trend line, MLUm generally increases with age, with growth rates slowing down as children grow older.

Figure 4.1 *MLUm for Girl–Boy Cohorts at Four Age Points with the Average Line Plotted Across the Full Age Span*



The cross-participant dataset provides a single-age snapshot of six children at 36 months (Figure 4.2). Five of the children produced MLUm scores between 2.45 and 3.01, with one girl scoring 3.54. This higher score increases the group mean to 2.87; however, overall variation within the group is limited.

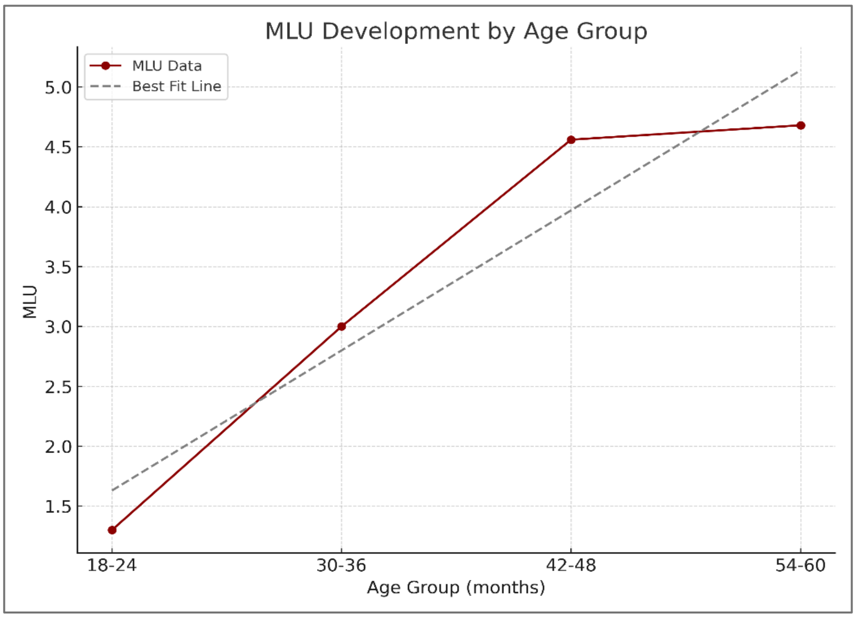
Figure 4.2 *MLUm Scores for Six Children in the Cross-participant Dataset at 36 Months, with the Group Average Indicated*



When all MLUm scores from both datasets are plotted together (Figure 4.3), the best fit regression line reveals a strong positive association between chronological age and MLUm. Pearson's correlation coefficient, calculated on the combined data, confirms a strong positive correlation, $r = .91$, $p < .001$.[67] This finding aligns with Tallas-Mahajna and Dromi's (2023) study of Spoken Palestinian Arabic, which also reports a significant positive correlation between Arabic-MPU and chronological age ($r = .727$, $p < .001$).

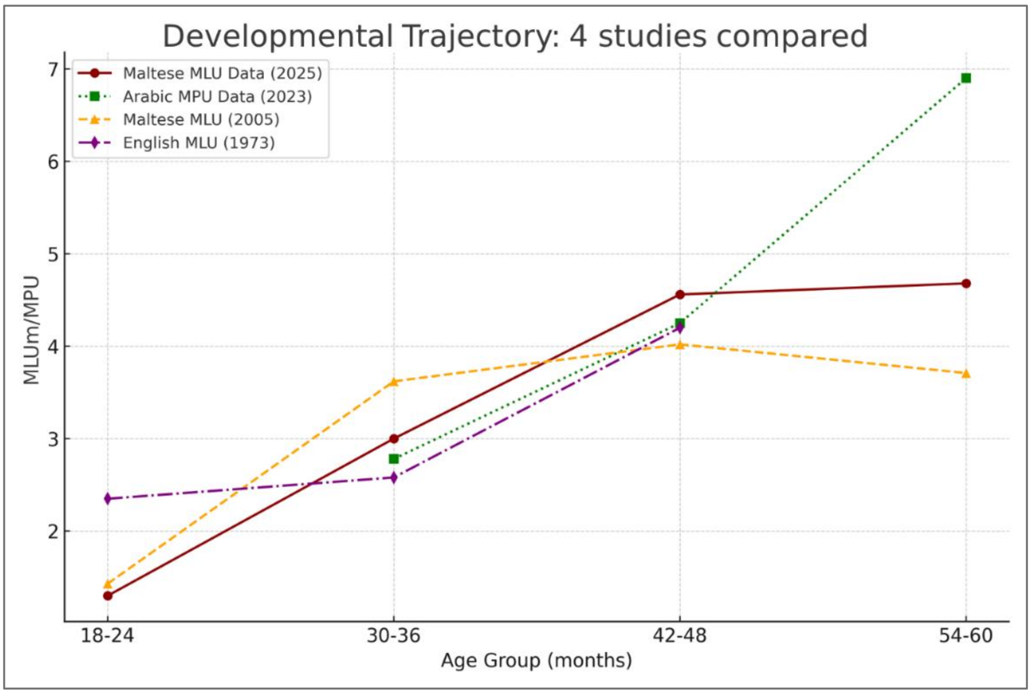
[67] This result is statistically significant since the p-value is well below the .05 significance threshold.

Figure 4.3 *MLUm Scores from All 14 Sessions Across Both Datasets, with the Best-Fit Regression Line Illustrating the Association Between Chronological Age and MLUm*



To further contextualise the developmental pattern observed here for TD, MD children, Figure 4.4 illustrates the MLUm trajectory observed in the present study alongside those reported in Brown (1973), Portelli (2005), and Tallas-Mahajna and Dromi (2023). These studies were selected as points of comparison as they directly inform the MLUm calculation strategy adopted here: Brown (1973) provides the foundational framework for MLU, Portelli (2005) offers the only prior MLUm-based analysis of MFLA, and Tallas-Mahajna and Dromi (2023) present a recent study of MPU in Spoken Palestinian Arabic, which is typologically comparable to Maltese and methodologically aligned with the present study (see Section 3.6.2). The comparison is centred on the 30–48-month developmental window, as this represents the period for which the present dataset is most robust and for which all three studies provide comparable data since Brown's (1973) dataset does not extend beyond 48 months, while Tallas-Mahajna and Dromi's (2023) cohort begins at 30 months.

Figure 4.4 *MLUm/MPU Scores of Four Language Acquisition Studies Compared*



The comparison reveals that, within the 30–48-month window, the trajectory observed in the present study aligns closely with those reported in all three studies. Between 18–24 and 30–36 months, the findings in this study are most similar to those of Portelli (2005), reflecting comparable developmental patterns in Maltese.[68] Between 30–36 and 42–48 months, the trajectory parallels those reported in both Brown (1973) and Tallas-Mahajna and Dromi (2023), indicating cross-linguistic consistency in the rate of morphosyntactic growth. Beyond 42–48 months, divergence is observed. Tallas-Mahajna and Dromi (2023) report a sharp increase in MPU, whereas both the present study and Portelli (2005) show relative stability, with the former increasing only marginally and the latter showing a slight decline.

[68] As noted previously (see Section 3.6. and Appendix O), Portelli's (2005) MLUm scores are higher due to her different scoring system, which tends to inflate values.

In the present study, utterance length ranged from 1 morpheme at 18 months to 20 morphemes at 60 months.[69] Portelli (2005) reports similar results for Maltese, with values ranging from 1 to 24[70] within the same age range. Examples illustrating variation in utterance length and structural complexity from the present dataset are provided below:

(6) 19m:	*baħħ*	*it-/teddy/*
	gone.BT	DEF-teddy
	'The teddy's gone.'	morphemes = 3

(7) 33m:	*u*	*kien*	*hemm*	*/story/*	*biex*	*ma*
	and	be.3SGM.PFV	there	story	so	NEG
	j-ixxerj-a-x		*u*	*ġie*		*/beast/*
	3PFV-share-SGM-NEG		and	came.3SGM.PFV		beast

'And there was a story so that he doesn't share and he turned into a beast.' morphemes = 12

(8) 44m: *when we're going outside to play and it's raining and there is the slides wet*
morphemes = 20

[69] For MLUm results, see Section 4.2.1.
[70] The higher maxima in her data are likely attributable to the inflation inherent in her scoring system.

(9) 60m:	*ija*	*u*	*l-ħmar-a*		*tiegħi*	*ġie*
	yes	and	DEF-donkey-SGF		mine	come.3SGM.PFV
	ħmar		*jismu*	*Ċvccv*	*u qed*	*j-igdim-ha*
	donkey.SGM		called.3SGM	Ċvccv	and PROG	3SG.M.IPFV-bite-3SG.F.OBJ
	minn	*hawnhekk*				
	from	here				

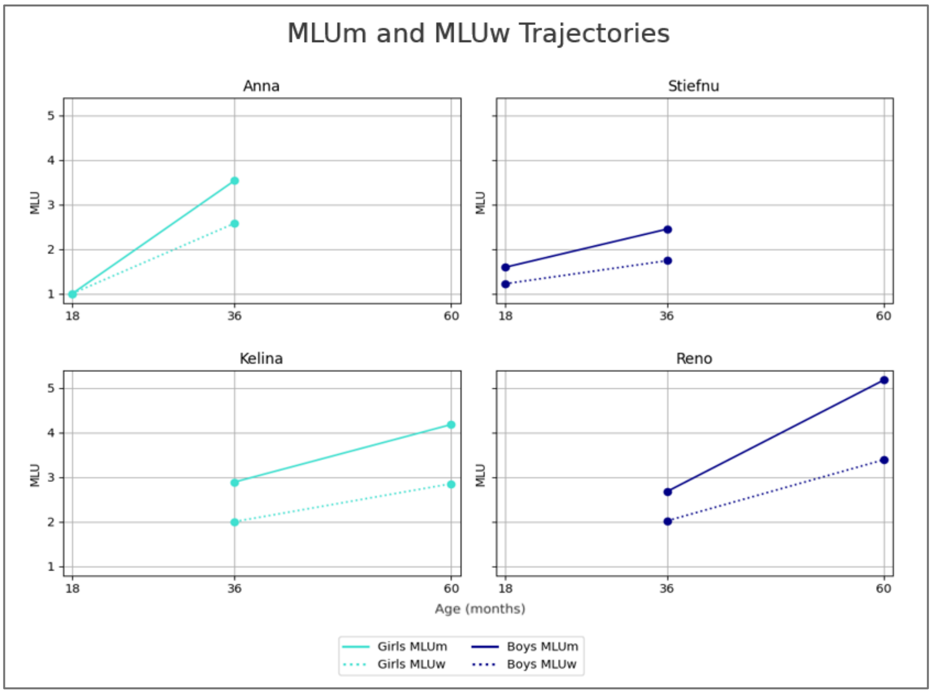
‘Yes, and my donkey, a donkey named Ċvccv came, and he’s biting her from here.’ morphemes = 19

The examples above illustrate how the quantitative trends identified in Figures 4.3 and 4.12 are reflected in the children’s actual productions. At the earliest stages, utterances are short, typically consisting of single morphemes or isolated lexical items, often nouns, onomatopoeic forms, or babytalk vocabulary (see Section 2.2.2). As development progresses, utterances become longer and more structurally complex, incorporating verb inflections, function words, and multi-morphemic constructions. This progression accounts for the steep rise in MLUm observed between 18–36 months, followed by a more gradual increase thereafter. The contrast between the earliest and later utterances also illustrates the transition from lexical to grammatical growth, as children’s expanding morphosyntactic repertoires enable the production of embedded and thematically elaborate structures. These qualitative observations therefore corroborate the statistical findings, demonstrating how increases in MLUm index both shifts in POS category expansion and the gradual mastery of Maltese morphosyntax.

MLUw was also calculated and analysed alongside MLUm for a subset of the data containing longitudinal observations across two developmental windows. The youngest and oldest girl–boy cohorts in the study were selected, and MLUm and

MLUw scores were compared between their first and last visits. The youngest set, Anna and Stiefnu, were first recorded at 18 months and last at 36 months; the oldest set, Kelina and Reno, were first recorded at 36 months and last at 60 months (see Table 3.7). The results of the MLUw calculation are presented in Figure 4.5 alongside the corresponding MLUm values.

Figure 4.5 *Developmental Trajectories of MLUm and MLUw Across Four Participants (18–60 months)*



As expected, the results presented in Figure 4.5 indicate that at the earliest stages, MLUm and MLUw yield comparable values (MLUm ≈ MLUw). However, as children's morphosyntactic skills develop, a divergence emerges, with MLUm rising more steeply than MLUw (MLUm > MLUw). This pattern reflects the increasing contribution of morphological marking to utterance length, which is captured by MLUm but not by MLUw.

No formal hypothesis was proposed regarding the relationship between MLUm and MLUw, as MLUm remains the primary measure of morphological development in the present study and is better suited to capturing the structural properties of Maltese. Accordingly, MLUw was calculated on a subset of eight transcriptions only. Nevertheless, the results presented here provide preliminary evidence that MLUm and MLUw diverge over developmental time in Maltese, supporting the view that MLUm is more sensitive to morphosyntactic growth in Maltese. Future work may examine this relationship in greater depth using a larger dataset.

In sum, the analysis confirms Hypothesis 1A since it demonstrates that MLUm increases with chronological age and, crucially, that the developmental trajectory observed in the present study is broadly comparable to those reported in Brown (1973), Portelli (2005), and Tallas-Mahajna and Dromi (2023), particularly within the 30–48-month window. These findings support the use of MLUm as a robust measure of morphosyntactic development in Maltese and situate the present study within both Maltese-specific and cross-linguistic CLA research.

While Hypothesis 1A addresses the developmental trajectory of MLUm across the cohort as a whole, closer inspection of the longitudinal data suggests potential variation between boys and girls. This is explored in Section 4.2.1.2 below.

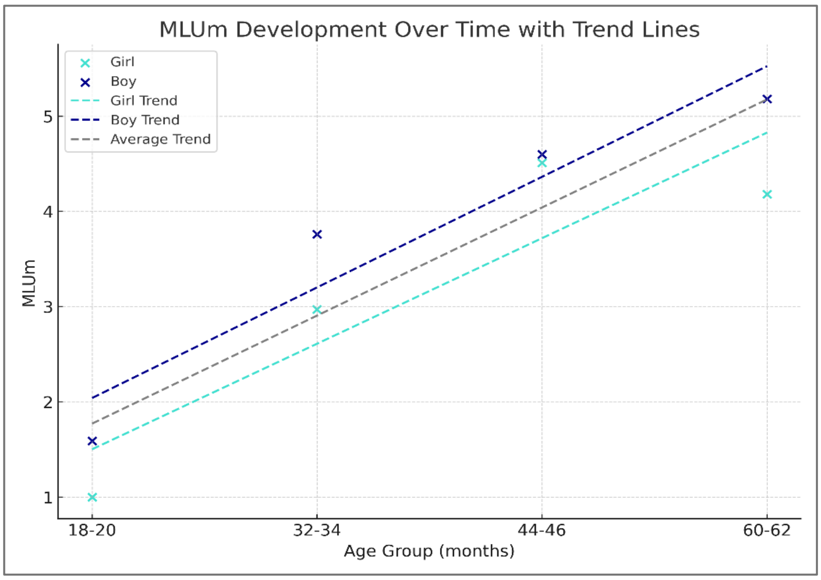
4.2.1.2. A note on sex-based differences

Given that sex-based differences may contribute to individual and group variation in MLUm, a brief consideration of these patterns is included here. This allows for a more nuanced interpretation of the results presented in the preceding section, where variation in MLUm may partly reflect sex-related differences alongside other factors.

Although sex-based differences are not treated as a formal hypothesis in the present study, MLUm values were examined across boys and girls in the developmental cohort. Given the small sample size and the structure of the dataset, no formal statistical testing is undertaken. Instead, the analysis is intended to document patterns observed in the data, without making inferential claims.

Across the four age points in the developmental cohort, boys consistently exhibit higher MLUm values than girls (see Figure 4.1 above). Figure 4.6 displays individual MLUm scores for each transcription, along with separate trend lines for boys, girls, and the overall average. The boys' trend line remains above that of the girls throughout the observed age range. This pattern diverges from accounts in the literature, which generally report higher MLUm values for girls in early childhood. However, caution is warranted when interpreting this trend, as the children at each age point are not the same individuals, and each group comprises only one boy and one girl.

Figure 4.6 *MLUm Scores for the Four Age Points in the Developmental Dataset, with Separate Trend Lines for Boys, Girls, and the Overall Average*



To address the limitation of non-overlapping participants across age points, the eight transcriptions comprising the subset of true longitudinal data was analysed (see Table 3.7). As shown in Figure 4.7 and corresponding Table 4.3, Anna demonstrates the steepest increase in MLUm, followed closely by Reno, whereas Kelina and Stiefnu exhibit more moderate growth.

Figure 4.7 *MLUm Scores for Four Children at their First and Last Recording Sessions*

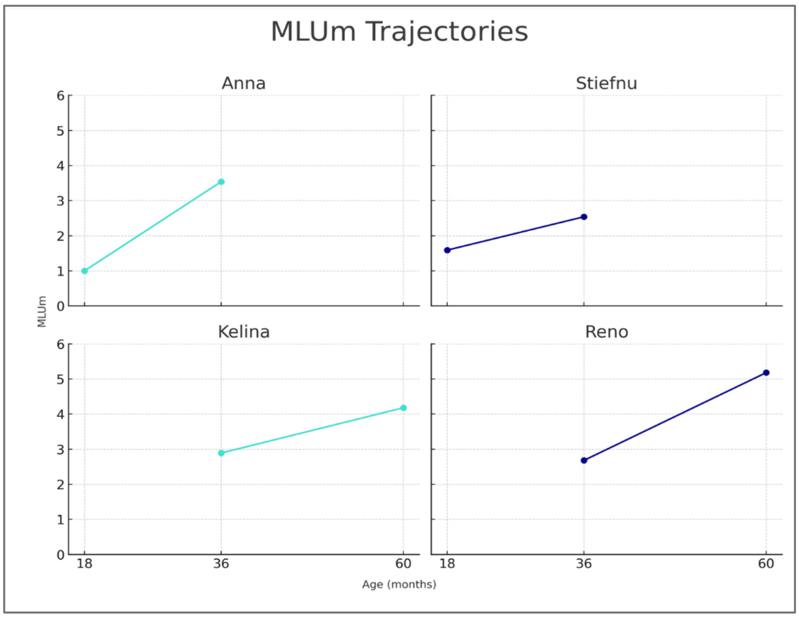


Table 4.1 *MLUm Increase Between First and Last Recording Session for Four Children*

	MLUm increase
Anna	+2.54
Stiefnu	+0.95
Kelina	+1.29
Reno	+2.50

Taken together, these two analyses yield conflicting patterns. The first developmental dataset suggests that boys consistently produce higher MLUm values than girls across the sampled age range. In contrast, the true longitudinal subset reveals substantial individual variation in growth trajectories, with no consistent differences in MLUm increase between girls and boys. These findings therefore preclude drawing firm conclusions regarding sex-based differences in MLUm. Instead, they highlight the role of IDs and methodological constraints, and underscore the need for larger, more densely sampled longitudinal datasets to determine whether such differences are robust.

4.2.1.3. A note on code-switching

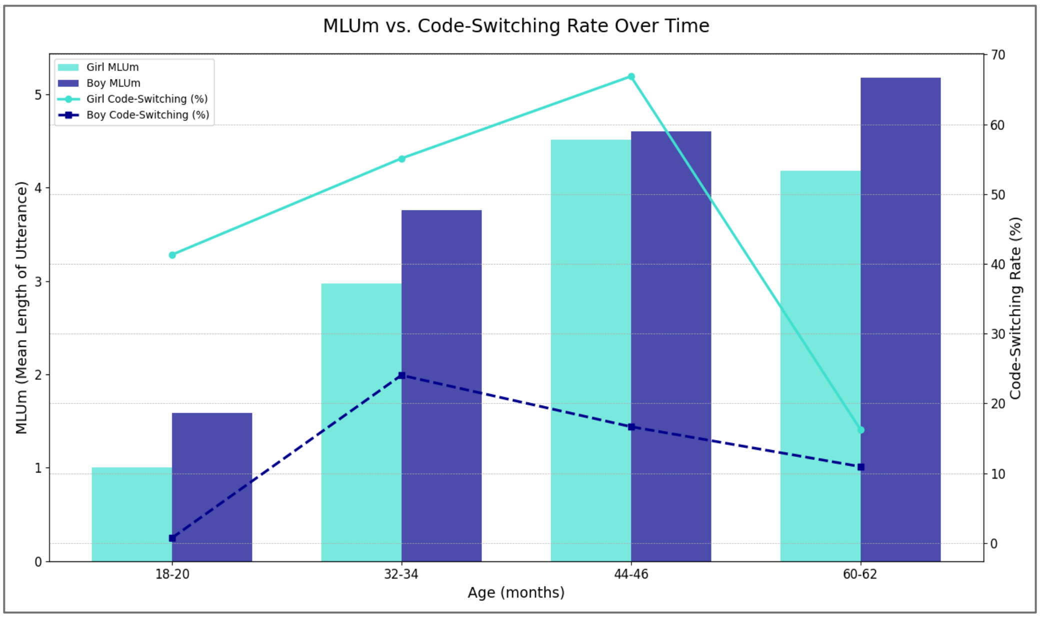
Given that patterns of language mixing may contribute to individual and group variation in MLUm, particularly in bilingual contexts, a brief consideration of CS is included here. This allows for a more nuanced interpretation of the results presented in the preceding section, where exposure to and use of English may partly account for observed differences between participants (e.g., by sex or individual background).

Although CS was not a primary focus of this research and no formal hypothesis was formulated, rates of CS per transcription were nonetheless calculated, given the distinctive sociolinguistic context of MFLA, in which Maltese is rarely acquired in isolation from English. For the purposes of this study, a code-switch is defined as the insertion of one or more lexical or functional items from English within an otherwise Maltese utterance, or vice versa.

In the cross-participant group (six children at 36 months), CS rates range from 15% to 40%. This wide variation is likely attributable to a combination of factors, including differences in individual children's degree of Maltese dominance, the linguistic choices and practices of their family members (e.g., one-parent-one-language approach), childcare attendance and differences in the conversational contexts captured during the recordings.

For the developmental group (eight recordings from six children, spanning ages 18–60 months), CS rates are consistently higher for girls than for boys at each sampled age point. Figure 4.8 presents these CS rates alongside the MLUm values reported earlier (Figure 4.1), allowing a visual comparison between the two measures. As with the analysis of sex-based differences in MLUm, caution is warranted: the children at each age point are not the same individuals, and only one girl–boy cohort is represented at each interval. Consequently, the apparent decrease in CS between 44–46 months and 60–62 months should not be interpreted as a systematic developmental decrease or as directly linked to changes in MLUm. Nonetheless, the observed patterns raise interesting questions about the relationship between CS, morphosyntactic development, age, and sex. Future research focusing on CS in Maltese child language could distinguish between inter-sentential and intra-sentential switching, an approach that has been shown to yield further insights into bilingual language processing and developmental trajectories in young children (e.g., Kapantzoglou et al., 2021; Meisel, 1994).

Figure 4.8 *MLUm Scores per Girl–Boy Cohort Alongside the CS Rate (%) for the Four Age Points Selected*



Across the analyses in this section, MLUm is found to be strongly and positively correlated with chronological age (Hypothesis 1A), with developmental trajectories comparable to those reported in Brown (1973), Portelli (2005), and Tallas-Mahajna and Dromi (2023). In examining potential sex-based differences, however, the results are less straightforward: developmental data suggest higher MLUm values for boys, whereas the true longitudinal subset reveals substantial individual variation with no consistent sex advantage. These conflicting patterns preclude firm conclusions and point to the need for larger datasets. Finally, the CS analysis reveals notable variability in rates between children, as well as a tendency for higher CS rates among girls in the developmental group. Together, these findings underscore the complexity of morphosyntactic growth in a bilingual environment and highlight promising avenues for future research on the interplay between age, sex, and bilingual language use.

4.2.2. POS category distribution

Having addressed the first component of RQ1, attention now turns to the second: the distribution of POS in the language of TD, MD children. This section begins with descriptive results on POS distribution and semantic categories in the corpus (Section 4.2.2.1). Focusing on nouns, verbs, and adjectives, I then examine Hypothesis 1B, which predicts that the youngest cohort will exhibit a noun bias (Section 4.2.2.2), and Hypothesis 1C, which posits a negative association between MLUm and the proportion of nouns (Section 4.2.2.3). In view of their relevance to later analyses (Hypothesis 2), particular attention is devoted to nouns and adjectives.

4.2.2.1. Descriptive results on POS distribution and semantic categories

In what follows, I present descriptive results on POS distribution in the child language corpus together with an examination of semantic categories. Token frequencies for Maltese and English nouns, verbs, and adjectives were calculated across more than 5,200 utterances. Semantic categories were derived from a subset of the most frequent nouns and adjectives (see Section 3.7).

Frequency lists were examined for each POS and cut-off points were chosen where token frequency dropped sharply. Token cut-off points per POS category were selected to capture the most representative lexical items in the corpus while excluding low-frequency forms since the distribution of lexical items in the corpus follows a typical Zipfian pattern. The distribution of these tokens by language (Maltese vs English) varies across POS, as illustrated in Table 4.2. A summary of the overall frequency of all five categories in relation to the total number of tokens in the corpus is provided in Appendix Q.

Table 4.2 *Distribution of the Most Frequent Noun, Verb, and Adjective Tokens by Language*

	Nouns	**Verbs**	**Adjectives**
Maltese tokens	57.5%	90.91%	53.33%
English tokens	42.5%	9.09%	46.67%

Given their theoretical and analytical relevance to the present study, nouns and adjectives were examined in greater detail. Tables 4.3 and 4.4 list the most frequent tokens within each category, ranked by frequency of occurrence in the corpus. These lists highlight both the lexical preferences of MD children and the extent of English influence on their expressive vocabulary.

Table 4.3 *Fifteen Most Frequent Noun Tokens in the Corpus, Ranked by Frequency*

Token no.	Rank	Noun	Language	Translation
1	1	mamà	Maltese	mum
2	2	dar	Maltese	house
3	3	kelb	Maltese	dog
4	4	nanna	Maltese	grandmother
5	5	/slide/	English	—
6	6	bandli	Maltese	swings/playground
7	7	/toys/	English	—
8	8	/mummy/	English	—
9	9	papà	Maltese	dad
10	9	pizza	Maltese	pizza
11	10	tifel	Maltese	boy
12	10	tifla	Maltese	girl
13	11	ħuta	Maltese	fish
14	11	/wolf/	English	—
15	11	żrinġ	Maltese	frog

Table 4.4 *Fifteen Most Frequent Adjective Tokens in the Corpus, Ranked by Frequency*

Token no.	Rank	Adjective	Language	Translation
1	1	/big/	English	—
2	1	żgħir-a	Maltese	small-SGF
3	2	/red/	English	—
4	3	kbir	Maltese	big.SGM
5	4	/green/	English	—
6	4	/yellow/	English	—
7	5	blu	Maltese	blue.SG

8	6	ħamr-a	Maltese	red-SGF
9	7	żgħir	Maltese	small.SGM
10	8	/angry/	English	—
11	8	/black/	English	—
12	9	kbir-a	Maltese	big-SGF
13	9	tajjeb	Maltese	good.SGM
14	10	/silly/	English	—
15	11	blu-na	Maltese	blue-SGF

Nouns and adjectives were further classified by semantic category, with proportions of Maltese and English tokens calculated for each category (Figures 4.9 and 4.10). This analysis enabled examination of whether particular categories were dominated by Maltese or English items. The following analysis is exploratory and is not tied to a specific hypothesis, but is included to provide further insight into POS category distribution across Maltese and English within the bilingual dataset.

For nouns, the categories comprise a) *persons*, b) *furniture and household objects*, c) *animals and nature*, d) *food and drink*, e) *places*, f) *vehicles*, g) *bodily functions*, and h) *play*. For adjectives, the categories include a) *size*, b) *colour*, c) *quality*, and d) *emotion*. These categories were informed by broad semantic fields commonly used in studies of early lexical development (e.g., Clark, 1993) and were designed to capture conceptually salient domains that frequently emerge in early child language while remaining appropriate for the bilingual Maltese–English context. Such domains are generally considered to reflect underlying conceptual structures that are shared across languages, even though their lexical realisation may vary depending on language-specific patterns.

Figure 4.9 *Distribution of the 40 Most Frequent Nouns by Semantic Category and Language*

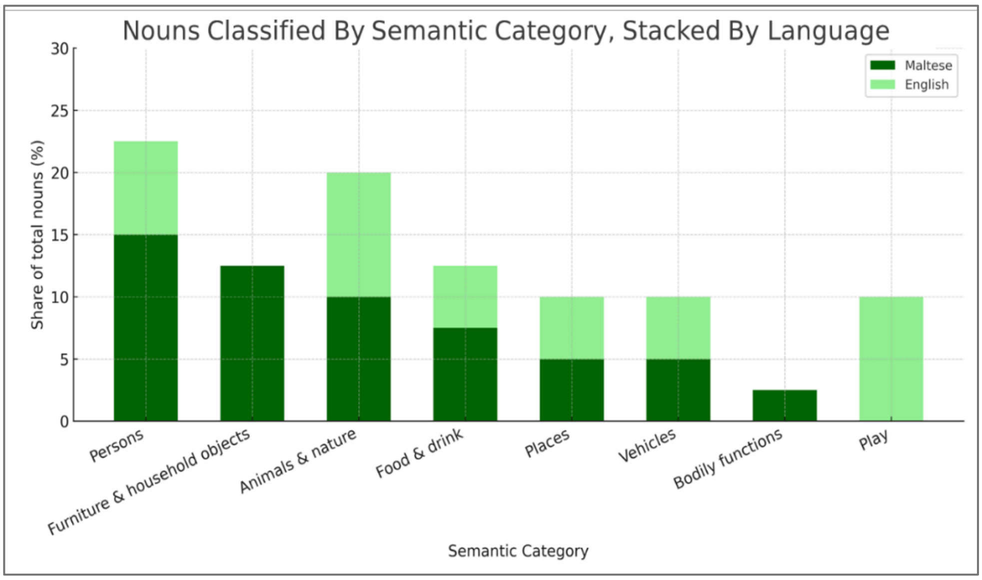
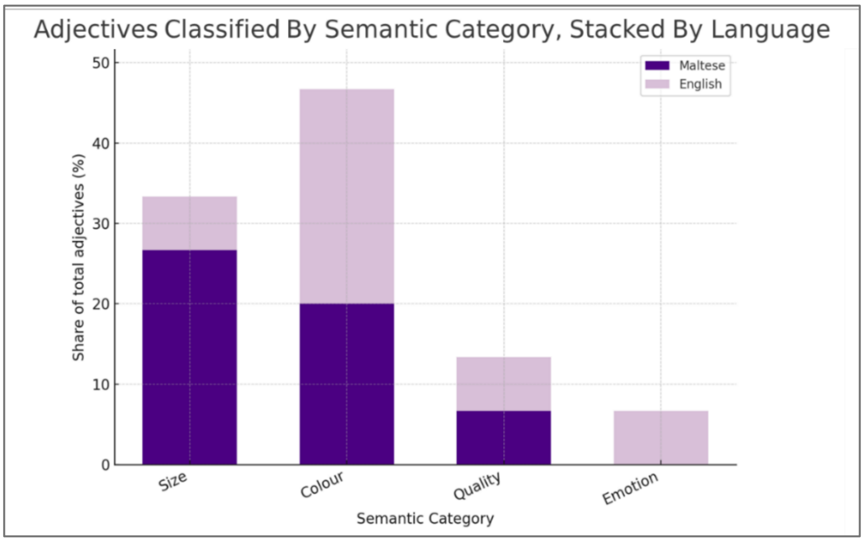


Figure 4.10 *Distribution of the 15 Most Frequent Adjectives by Semantic Category and Language*



In summary, the corpus analysis suggests that the relative proportions of Maltese and English tokens differed across POS. Nouns and adjectives draw from both languages, but verbs are mostly Maltese. The analysis of nouns and adjectives by semantic category further demonstrates that some domains are dominated by Maltese and others more influenced by English.

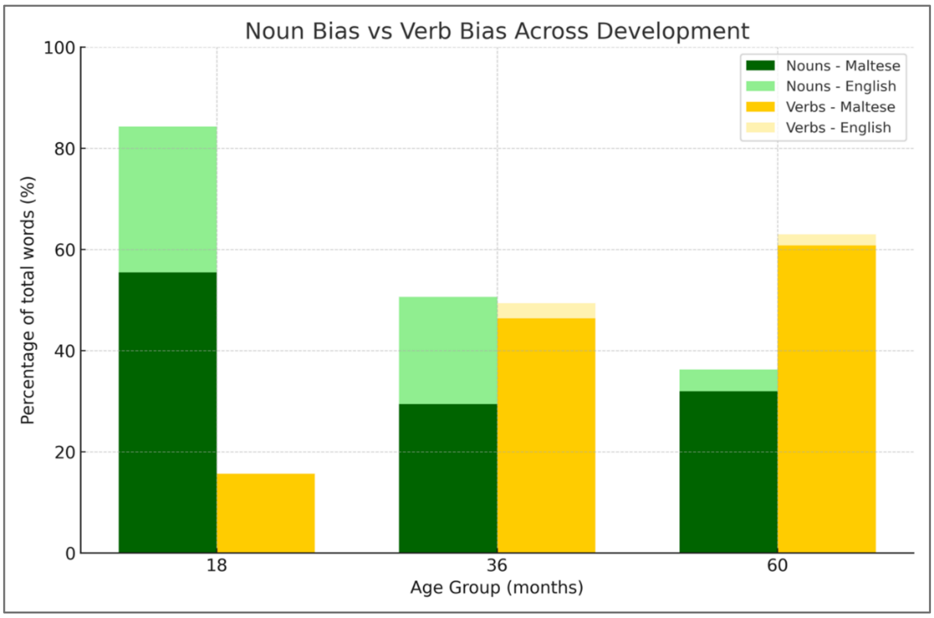
4.2.2.2. Hypothesis 1B: Noun bias over time

Hypothesis 1B predicts that a noun bias will characterise the youngest age group, but that its prevalence will gradually decrease over time as the proportion of verbs and other non-nominal elements increases. To test this, longitudinal data were analysed using the five-tag system described above, focusing on nouns, verbs, and adjectives.

A summary of results for each of the eight transcriptions, including the proportions of Maltese and English tokens within each category, is provided in Appendix Q. In brief, the distribution of Maltese and English varies both by POS category and by each child's language dominance at a given point in time. Verbs are consistently produced in Maltese, with a minimum of 90% Maltese tokens across all children and ages. As expected, the proportion of tokens tagged as *child-language specific* is highest among the 18-month-olds (24–30%), declining to below 5% in the older groups. In contrast, tokens classified as *others* (function words, adverbs, auxiliaries, pseudoverbs, and active participles) account for 30% or less at 18 months but rise to 47% or more in the older groups. Similar to verbs, this category is also dominated by Maltese tokens.

Since the debate in the literature is framed around a noun bias versus a verb bias, nouns and verbs were aggregated across age groups, with tokens tagged as *adjectives*, *others*, and *child-language specific* excluded. The results are presented in Figure 4.11, which also shows the relative proportions of Maltese versus English tokens. The results reveal a consistent preference for Maltese verbs, while the proportion of English nouns relative to Maltese nouns decreases with age.

Figure 4.11 *Noun Versus Verb Bias Across Development at 18, 36, and 60 Months, by Language*



In the data from the four children analysed, a noun bias is indeed observed in the youngest age group (84%). However, by 36 months verbs nearly equal nouns (51%) in frequency, and by 60 months verbs predominate as noun proportions decrease further (36%). These results support Hypothesis 1B, demonstrating that the proportion of nouns decreases over time, giving way to verbs.

To further analyse whether the proportion of noun productions decreases with age, a mixed-effects logistic regression model was fitted, with noun production (noun = 1, verb = 0) as the dependent variable and centred age as the fixed predictor. Child was included as a random intercept to account for repeated observations.

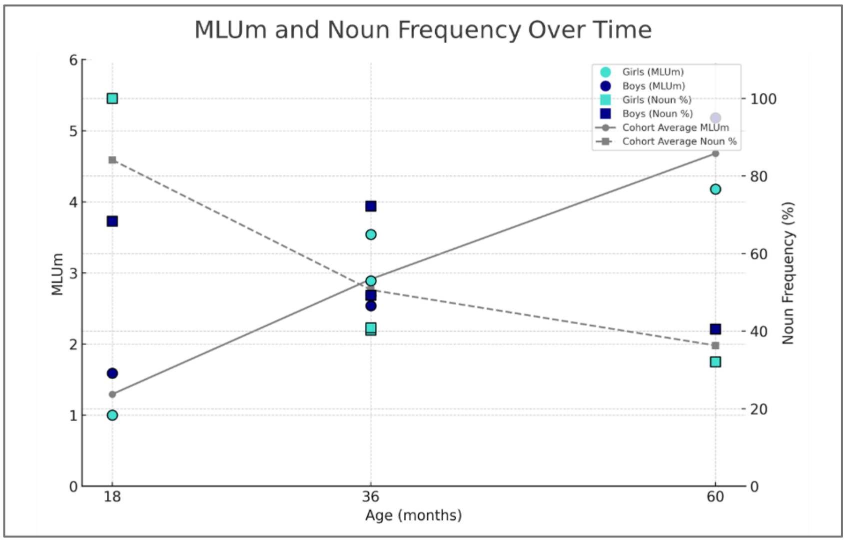
The dataset comprised 715 observations across four children (Anna: n = 126; Stiefnu: n = 118; Kelina: n = 199; Reno: n = 272), with 341 noun tokens and 374 verb tokens. Age was mean-centred (mean = 44.26 months) prior to modelling. The model (*noun ~ age_c + (1 | child_id)*) revealed a significant negative effect of age (β

= −0.043, SE = 0.008, $z = -5.64$, $p < .001$), indicating that the likelihood of producing a noun decreases as age increases. The intercept was not significant (β = −0.177, SE = 0.176, $p = .313$), reflecting the log-odds of noun production at the mean age. The variance associated with the random intercept for child was relatively small ($\sigma^2 = 0.093$), suggesting limited between-child variability in baseline noun production. Therefore, these findings also support Hypothesis 1B, indicating a developmental shift away from an early noun bias towards increased use of verbs and other non-nominal elements over time.

4.2.2.3. Hypothesis 1C: MLUm and noun proportions

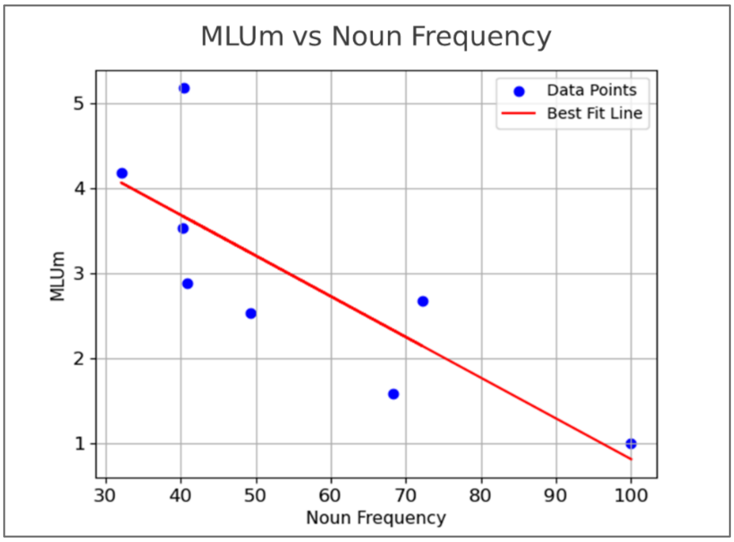
Hypothesis 1C predicts a negative correlation between MLUm and the proportion of noun tokens. To provide an initial descriptive overview, Pearson's correlation coefficient was calculated for the aggregated data. As illustrated in Figure 4.12, noun proportions (dotted line) decrease as MLUm (solid line) increases across the four children analysed.

Figure 4.12 *Longitudinal Development of MLUm and Proportion of Noun Tokens for Four Children*



For the aggregated dataset, the correlation between MLUm and noun proportions was strong and statistically significant ($r = -.81$, $p < .001$), indicating that higher MLUm values are associated with lower proportions of noun tokens (Figure 4.13).

Figure 4.13 *Scatter Plot Showing the Negative Correlation Between MLUm and Noun Proportions*



To examine this relationship more formally while accounting for repeated observations within children, a mixed-effects logistic regression model was fitted with noun production (noun = 1, verb = 0) as the dependent variable and centred MLUm as the predictor, with child included as a random intercept.

MLUm was mean-centred (mean = 3.58 morphemes) prior to modelling. The model (*noun ~ mlu_m_c + (1 | child_id)*) revealed a significant negative effect of MLUm ($\beta = -0.607$, SE = 0.095, $z = -6.40$, $p < .001$), indicating that higher MLUm values are associated with a lower probability of noun production relative to verbs.

The random intercept variance ($\sigma^2 = 0.202$) suggests meaningful individual variability once utterance length is taken into account.

Model comparison further showed that MLUm provided a better fit to the data than age (AIC = 936.61 vs. 948.46), indicating that utterance length is a stronger predictor of noun–verb distribution than chronological age. A parallel model using MLUw (*noun ~ mlu_w_c + (1 | child_id)*) yielded comparable results ($\beta = -1.040$, SE = 0.162, $z = -6.42$, $p < .001$; AIC = 936.64), reinforcing the robustness of this relationship across measures of utterance length.

Taken together, these findings support Hypothesis 1C: increases in utterance length (MLUm and MLUw) are associated with a decrease in the proportion of noun tokens, marking a shift in POS category distribution.

Overall, the analyses addressing Hypothesis 1 (and its subcomponents, 1A–1C) reveal a clear developmental shift in TD, MD children's language use: an early noun bias evident at 18 months gradually gives way to a growing reliance on verbs and other grammatical elements as morphosyntactic complexity increases. Together, these findings highlight the interplay between MLUm and POS category distribution in early bilingual acquisition. The following section (Section 4.3) offers an interpretation of the findings, situating them within the broader framework.

4.3. Discussion

4.3.1. Hypothesis 1A: MLUm and chronological age

The findings of Hypothesis 1A demonstrate a clear increase in MLUm with chronological age and, importantly, show that the developmental trajectory observed in the present study is broadly comparable to those reported in Brown (1973), Portelli (2005), and Tallas-Mahajna and Dromi (2023), particularly within the 30–48-month window. This developmental window is especially informative, as it represents both the most densely sampled period in the present dataset and the only interval for which all three comparison studies provide directly comparable data.

The close alignment with Portelli (2005) in the earlier stages supports the validity of the present findings within a Maltese-specific context, while the convergence with Brown (1973) and Tallas-Mahajna and Dromi (2023) in the mid-developmental window points to cross-linguistic regularities in morphosyntactic growth. The divergence observed beyond 42–48 months, particularly in comparison with Tallas-Mahajna and Dromi (2023), may reflect differences in morphological systems, dataset composition, or methodological choices, including the treatment of non-linear morphology and code-switched utterances.

This progression aligns with Portelli's (2005) three-stage model of Maltese MLUm development, adapted from Brown's (1973) five foundation stages (MLUm 1.75–4.0). According to Portelli, children in the initial stage (MLUm 1.0–2.0) rely heavily on gestures and onomatopoeic forms, gradually moving towards single words (predominantly nouns) and simple noun–verb combinations. The second stage (MLUm 2.0–4.0) is marked by rapid lexical growth, particularly in nouns and verbs, and the emergence of longer, more complex sentences. At this stage, children begin to master grammatical structures such as negation and the use of direct and indirect objects. In the final stage (MLUm 4.0–5.5), the rate of MLUm growth slows, and the focus shifts from structural acquisition to the functional use of language for reasoning and interaction. The findings of the present study (Hypotheses 1A–C) are broadly consistent with these developmental stages.

MLUw was also calculated in order to be compared with MLUm, although no formal hypothesis was advanced regarding this comparison. The results provide preliminary support for an expected divergence between the two measures in a language such as Maltese, characterised by both concatenative and non-concatenative morphology. While the utterances of younger participants are relatively similar in length when measured in morphemes and words, a clear divergence emerges with age. In older participants, MLUm consistently exceeds MLUw, reflecting the increasing contribution of morphological complexity to utterance length. Representative examples from the corpus are given below: [71]

(1) 19m:	*ara,*	*il-kokka*	
	see.IMP	DEF-owl.SGF	
	'Look, the owl.'		morphemes = 3, words = 2

[71] To enhance readability, annotation conventions have been omitted. Only those conventions directly pertinent to the analysis are retained.

(2) 27m: *n-agħml-ilhom* */punishment/*
1SG.IPFV-do-3PL.OBJ punishment
'I'll punish them.' morphemes = 4, words = 2

(3) 33m: *kon-t* *kissir-t-ilha* *ras-ha*
be-1SG.PFV break-1SG.PFV-3SGF.OBJ head.SGF-3SGF.OBJ
'I had broken her head.' morphemes = 7, words = 3

(4) 36m: *inti* *ġib-t-hie-li*
you.SG bring-2SG.PFV-3SGF.OBJ-1SG
'You brought it to me.' morphemes = 5, words = 2

(5) 44m: *ma* *n-agħml-u-lu-x* *il-/face/*
NEG 1PL.IPFV-do-2PL-3SGM.OBJ-NEG DEF-face
'We don't do the face for him.' morphemes = 7, words = 3

Ntelitheos and Idrissi (2017, p. 237) calculate both MLUm and MLUw in their study of Emirati Arabic and report an "extremely strong correlation" between the two measures. Crucially, however, their analysis considers only linear morphological processes and excludes the non-linear component of Emirati Arabic morphology. This methodological choice may account for the strength of the reported correlation, as the inclusion of non-linear morphology would be expected to reduce the correspondence between MLUm and MLUw, potentially revealing a divergence over developmental time similar to that observed here for Maltese.

In conclusion, the findings support the use of MLUm as a reliable index of early morphosyntactic development in TD, MD children, while also demonstrating that its developmental trajectory is comparable to those reported in both Maltese-specific and cross-linguistic research. While the primary aim of the analysis is theoretical and descriptive, the MLUm values reported here also provide a developmental reference that may be of relevance in clinical contexts, particularly in the absence of normative data for MFLA.

4.3.2. Hypothesis 1B: Noun bias over time (and insights from the descriptive results)

Hypothesis 1B posits that a noun bias will characterise early child language, diminishing over time as verb use increases. The findings from the longitudinal dataset align with this prediction: at 18 months, children's vocabularies are strongly dominated by nouns; by 36 months, verbs approach nouns in frequency; and by 60 months, verbs overtake nouns, marking the expected developmental shift. The results of the mixed-effects logistic regression model further support this pattern, confirming a significant decrease in noun production with increasing age.

This trajectory is consistent with the results of previous research on Maltese, which show that children's early lexicon is typically rich in nouns, particularly in the semantic domains of objects, people, and familiar routines (e.g., Gatt, 2010; Portelli, 2005), with more recent evidence also showing a strong noun bias in caregiver-reported vocabularies, alongside a later increase in verbs and function words as children develop (Gatt et al., 2023). Gatt (2010, p. 374) concludes that her findings on early Maltese vocabularies "[add] to the cross-linguistic evidence in favour of a noun bias in early vocabularies", a pattern first theorised by Gentner (1982) but challenged by studies of so-called verb-friendly languages such as Mandarin (e.g., Tardif, 1996; see Section 2.3.3). The subsequent decrease in noun bias observed here reflects increasing morphosyntactic maturity, wherein verbs assume the role of syntactic pivot, encoding nominal arguments through inflectional processes such as subject–verb agreement and pronominal affixation.

Importantly, an early noun bias may also be an outcome of properties of input. CDS in early stages may include a higher proportion of referential prompts (e.g., “what’s that?” questions), which could favour noun production. While the present study does not quantify such patterns, this possibility highlights the role of interactional context in shaping early POS category distributions. Future research incorporating parallel analyses of adult input would be necessary to disentangle input-driven and child-internal contributions to the observed noun bias.

With regard to the distribution of Maltese versus English tokens, it is well established in the literature that code-switches, particularly single-word insertions, most frequently involve nouns and adjectives, and less commonly verbs and function words (e.g., Myers-Scotton, 2002). Nouns are generally identified cross-linguistically as the most frequently code-switched category, a tendency attributed to their labelling function and their ease of borrowing to fill lexical gaps. Yet, in this dataset, both the descriptive corpus results (Section 4.2.2.1) and the hypothesis-testing (Section 4.2.2.2) indicate that adjectives are the most frequent category of English insertions, followed closely by nouns. This finding is congruent with patterns reported in Maltese CS, where lexical fields such as numbers and colour terms are often expressed in English. The high frequency of colour terms in English inflates the adjectival category in the code-switched data, possibly explaining why adjectives marginally surpass nouns in this study.

The classification of nouns and adjectives by semantic category reveals domain-specific patterns of language use. Some categories are dominated by Maltese tokens, while others show greater proportions of English, indicating that bilingual influence is uneven and shaped by semantic domain. For instance, household and routine-related items are largely expressed in Maltese, reflecting the children's immediate home environment, whereas references to toys and animals contain more English items, pointing to contexts where English input is especially salient. These findings demonstrate that bilingual development is influenced not only by overall language dominance but also by the functional and situational distribution of input. Analysing semantic categories therefore provides a more nuanced account of how lexical influence from each language emerges across different areas of experience.

By contrast, the strong predominance of verbs in Maltese is particularly notable. This may be explained by the relatively rich inflectional morphology of Maltese verbs, which increases their salience in the input and renders them more resistant to borrowing. The data therefore suggest that while English exerts a measurable influence on children's nominal and adjectival vocabularies, verbs remain firmly anchored in Maltese. This stability extends to the category labelled *others* (function words, adverbs, auxiliaries, pseudoverbs, and active participles), which also remains largely unaffected by English influence. Together, the results indicate that while lexical categories are more permeable to cross-linguistic influence, grammatical and morphosyntactic domains appear relatively stable.

Nevertheless, the small sample size used to test Hypothesis 1B (a total of eight transcriptions from four children) warrants an element of caution. The degree of noun bias and the rate of its decrease may vary substantially across individuals depending on factors such as language dominance, input conditions, and exposure to English in specific contexts such as preschool. Future investigation should therefore make use of larger and more diverse datasets to test the generalisability of these developmental patterns in MD children.

In conclusion, the present findings reinforce cross-linguistic evidence for an early noun bias. The decline of noun dominance alongside the rise of verbs emphasises the increasingly central role of morphosyntax in children's utterances. This observation provides a natural bridge to Hypothesis 1C, which examines how the proportions of nouns in children's language relate to their MLUm, thereby connecting POS category distribution with broader measures of morphosyntactic growth.

4.3.3. Hypothesis 1C: MLUm and noun proportions

Hypothesis 1C predicts a negative correlation between MLUm and the proportion of nouns in the early language of TD, MD children aged 18–60 months. The results confirm this prediction due to the stronger predictive value of MLUm relative to age. However, this relationship should not be interpreted as evidence of a unidirectional causal effect. The observed association may reflect bidirectional influences or a shared underlying developmental process linking utterance length and POS category distribution.

This relationship reinforces the view that lexical and grammatical development in bilingual children should be seen as interconnected processes. Early child language dominated by nouns may provide a foundation for subsequent grammatical growth, and expansion in one POS category is often associated with growth in others (e.g., Bornstein et al., 2004; Piccin & Waxman, 2007). As morphosyntactic structures increase in complexity, verbs and other categories become proportionally more prominent. The pattern observed here aligns with developmental theories which propose that nouns initially serve as lexical building blocks for referential communication, while verbs assume greater importance once children begin producing longer, syntactically structured utterances. Verb acquisition is often described as more complex (e.g., Gentner & Boroditsky, 2001; Swearengin et al., 2025).

This finding also highlights the utility of MLUm as an index of development in bilingual contexts. Although the MD children in this study are bilingual to varying degrees, the negative effect of MLUm remains significant. This suggests that MLUm is a reliable measure of morphosyntactic growth even in contexts where the data include elements from two languages. However, given the limited sample size and the Maltese-dominance of children tested, further research with larger and more diverse samples is required to establish whether the same relationship holds for children with greater English dominance or for those in more balanced bilingual environments.

In conclusion, the significant negative effect of MLUm observed here provides quantitative support for the developmental trajectory described in

Hypothesis 1B. The decrease in noun dominance and the concurrent increase of verbs are associated with the steady MLUm gains, reinforcing the view that lexical composition and morphosyntactic complexity develop in tandem. This interconnectedness highlights the value of MLUm as an index of grammatical growth and the importance of considering POS category distribution when examining language acquisition milestones.

4.3.4. A note on sex-based differences and interactional context

Although sex-based differences are not treated as a formal hypothesis in the present study, patterns relating to MLUm across boys and girls are briefly considered here as part of a broader interpretation of variability in the dataset. Previous research in the field has often reported a marginal advantage for girls in early morphosyntactic development, although effect sizes are typically small and diminish with age (e.g., Eriksson et al., 2012; Hadley, 2011). [72]

In the present study, the findings are mixed. Cross-sectional patterns suggest that boys exhibit higher MLUm values than girls across the sampled age points, whereas longitudinal comparisons reveal substantial individual variation, with no consistent differences in rates of development between the two groups (see Table 4.1. and Figure 4.6.). These patterns do not permit firm conclusions regarding sex-based differences and best interpreted as exploratory.

Crucially, the data highlight the role of interactional and contextual factors in shaping language production. During recording sessions, boys tended to be more

[72] See Wallentin (2009) for a detailed overview.

outgoing with me as an unfamiliar investigator and engaged more independently, whereas girls were more likely to rely on parental presence and appeared more reserved. These differences may have influenced language output and, consequently, MLUm values. It is plausible that had the recordings been restricted to child–parent/sibling interactions without the presence of an external investigator, the results may have been more in line with the literature.

Family dynamics also appear to play a role. Changes in family structure, particularly the birth of siblings, were reflected in individual developmental trajectories. Three of the eight participating families experienced such changes, which were in turn reflected in the samples collected. For instance, Anna, one of the four children tested in the longitudinal set, assumed the role of older sibling by the final recording and displayed the steepest MLUm increase among the four children in this subgroup. In contrast, Stiefnu's family also welcomed a second child, but his language development showed little change associated with this transition. Interestingly, Anna explicitly referred to her new caregiving role in her language, as illustrated in examples (10–12).

(10) 36m: *ilbieraħ hawnhekk kien jaqa'* 'yesterday here he used to fall'
(11) 36m: *imbagħad għamel naqra sad u ixrob* 'and then he got a bit sad and drink'
(12) 36m: *jien ġibtlu l-water fil-high-chair* 'I got him the water in the high-chair'

Kelina, another child tested in the longitudinal set, was anticipating the birth of a sibling at the time of the final recording, at 60 months. Her behaviour during this session contrasted with earlier recordings, as she appeared more withdrawn and dependent on her mother, perhaps reflecting heightened awareness of impending familial change. As a result, her MLUm at 60 months was 4.18, which is relatively low compared with her 60–month-old counterpart, whose MLUm score was 5.18, and also when viewed in relation to her own earlier recording sessions. Furthermore, the only child in the sample with an older sibling, Karmena, consistently sought her sister's presence during sessions and frequently modelled her behaviour on that of her sibling. Taken together, these observations, well-documented in CLA research,[73] illustrate how sibling dynamics and responses to familial transitions can shape developmental trajectories and significantly mediate language outcomes. In such cases, the presence of a third-party investigator, rather than an exclusive focus on child–parent/sibling interactions, may also have contributed to greater withdrawal and behavioural differences

In addition to influencing MLUm outcomes, the interactional context may have affected CS rates. The recordings primarily captured interactions between the investigator and the children in a play/educational context, rather than in everyday routines, such as mealtimes or bedtime, which may have contributed to the observed rates of CS (15–40%). Parents frequently reported that English was the primary language of instruction in childcare centres and the preferred language of play. Therefore, the fact that the investigator was often perceived by the children as a play partner, affectionately referred to as *Michela tat-toys*, literally 'Michela of the toys', may have encouraged greater use of English despite deliberate efforts to

[73] The role of siblings and shifts in family dynamics are briefly outlined in Section 2.2.2.

promote Maltese (e.g., through the use of a Maltese-speaking puppet; see Section 3.3.2).

Taken together, these observations underscore the complex interplay between individual variation, family dynamics, and interactional context in shaping children's language production. The present data reveal patterns that are not fully captured by broad developmental generalisations, including differences in children's engagement with the investigator, shifts in family structure, and the influence of sibling relationships on language use. Importantly, the ethnographic detail afforded by close and repeated contact with participants highlights factors that are often invisible in large-scale quantitative studies, yet which can significantly shape language samples in small cohorts.

These findings further suggest that results derived from small samples, particularly those based on spontaneous speech data, are highly sensitive to contextual and interpersonal variables. While large-scale studies remain essential for identifying average developmental trends, fine-grained, participant-level insights are crucial for interpreting variability and situating findings within the broader landscape of CLA research. As such, the patterns reported here are best understood as exploratory and hypothesis-generating, providing a basis for future research with larger and more controlled datasets.

4.3.5. Concluding remarks

Taken together, Hypotheses 1A–C, provide converging evidence for Hypothesis 1: the developmental trajectory of early morphosyntactic acquisition in TD, MD children is characterised by a significant positive correlation between MLUm and shifts in POS category distribution.

From a theoretical perspective, these findings support models of language development that posit a close interaction between POS category distribution and grammatical growth, whereby the lexicon provides an early scaffold for morphosyntactic elaboration, and grammatical structures in turn reshape patterns of lexical use. Crucially, as utterances become longer and more complex, nouns assume a progressively grammaticalised role, contributing not only as lexical labels but also as arguments subject to inflection and agreement. The discussion now turns to nominal inflection and agreement, situating these processes within the acquisition trajectory of morphosyntactic development in TD, MD children and assessing their contribution to the consolidation of grammatical structure between 18–60 months.

In sum, the findings reported here in response to RQ1 establish that MLUm and POS category distribution provide significant indices of early morphosyntactic growth in TD, MD children. RQ2 and RQ3 provide the springboard for extending the analysis by examining how specific morphosyntactic features, gender and number, emerge and stabilise across nominal inflection and agreement, and how their developmental pathways intersect with the production measures identified as being present in response to RQ1.

5. From nominal inflection and agreement to developmental timing

Building on the developmental patterns established in Chapter 4, this chapter addresses RQ2 and RQ3, focusing on the acquisition of gender and number features and their connection to overall linguistic development. Specifically, it investigates how nominal inflection and agreement develop across age and task, and how these feature-specific trajectories align with broader measures of morphosyntactic expansion, namely MLUm and POS category distribution. The chapter opens with Hypothesis 2 and its subhypotheses, which outline expected patterns in nominal inflection, agreement, and developmental timing (Section 5.1). Results from both naturalistic and elicited data are then presented and discussed in relation to the wider developmental trends observed in MFLA (Sections 5.2–5.3). For a description of the core properties of Maltese nominal inflection and NP-internal agreement analysed here, see Sections 2.5.1–2.5.1.3.

5.1. Hypothesis 2

This section sets out the hypotheses that guide the analysis of nominal inflection and agreement with respect to overall production growth.

Overarching hypothesis

In TD, MD children, the acquisition of gender and number features follows a systematic developmental trajectory in which nominal inflection and agreement become increasingly stable with overall language growth.

Subhypotheses

The overarching hypothesis is further specified in the following subhypotheses:

2A: Children will produce concatenative plural forms with higher accuracy and consistency than templatic plural forms.

2B: Accuracy in gender and number agreement (on demonstratives, adjectives, and quantifiers) will increase with age, with gender agreement stabilising earlier than number agreement.

2C: The development of nominal inflection and agreement will coincide with increases in MLUm and a lexical shift from a noun-dominant vocabulary to one with greater POS diversity.

These hypotheses focus on the development of nominal inflection and agreement and their relation to overall language growth. Gender and number represent a particularly rich testing ground, as they surface early in nominal inflection and require integration across agreement domains in order to stabilise fully. Examining their developmental pathways therefore allows us to ask both when and how these features emerge and stabilise in MFLA, and crucially, how these trajectories align with changes in MLUm and POS category distribution identified in Chapter 4.

5.2. Results

The upcoming subsections address Hypotheses 2A (Section 5.2.1) and 2B (Section 5.2.2), presenting the results in two stages: first through the analysis of naturalistic data, and subsequently through the analysis of the experimental data that serve to confirm or challenge the proposed hypotheses. The following section (Section 5.2.3) then turns to Hypothesis 2C, integrating the results of Hypothesis 1 with those of Hypotheses 2A and 2B, thus providing a comprehensive synthesis that concludes the examination of the hypotheses.

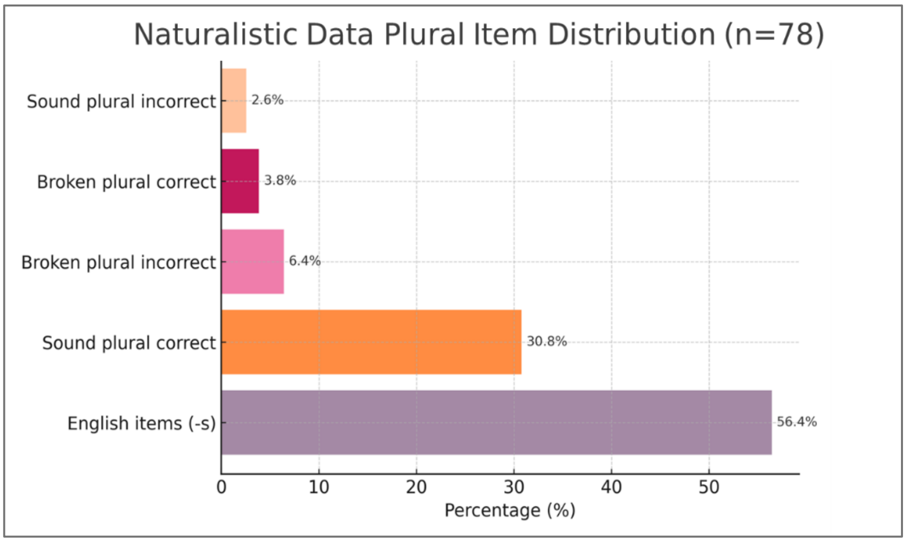
5.2.1. Hypothesis 2A: Nominal inflection

Hypothesis 2A centres around nominal inflection, focusing exclusively on number marking (see Section 3.8). The hypothesis predicts that children will produce concatenative plural forms with greater accuracy and consistency than templatic plural forms.

5.2.1.1. Naturalistic data

In the naturalistic dataset (eight transcriptions analysed, 2,156 clean child utterances), a total of 78 plural tokens were identified representing 3.6% of all utterances in this corpus. No plural forms were produced at 18 months. At 36 and 60 months, the proportion of plural use ranged between 1.44% and 7.24% of utterances per transcription. The 78 plural tokens included English plurals, target-like SP and BP forms and non-target-like plural forms, distributed as follows: 44 English plurals, 24 target-like SPs, 3 target-like BPs, 2 non-target-like SPs, and 5 non-target-like BPs (see Figure 5.1). Thus, when producing plurals spontaneously, children tended to rely heavily on either English plural morphology or concatenative Maltese plural forms. Importantly, they committed only two errors with SPs, while BPs were rarely attempted (8 tokens total) and were more often erroneous than correct (5 errors vs 3 correct). Moreover, the two errors classified as SPs were not entirely clear-cut and could have plausibly been assigned to another category. The rationale for this classification is provided in Section 5.3.1.

Figure 5.1 *Distribution of Plural Tokens in the Naturalistic Data*



In addition to the plural tokens discussed above, 42 instances of collective forms were identified in the naturalistic data, all of which were produced without error. Table 5.1 presents up to five examples, where available, for each plural category attested in the data, including the collective. All non-target-like tokens were included, which motivated the choice of five examples to ensure full representation. In the naturalistic data, four target-like SP suffixes were used: -*i* (n = 13), -*s* (n = 6), -*ijiet* (n = 4), and -*ien* (n = 1). The BP patterns attested were Type A and Type C.[74] All SP and BP non-target-like productions involved overgeneralisation of SP suffixes, specifically *-i* (n = 3), *-iet* (n = 1), and *-s* (n = 1). These examples illustrate both the systematicity and the restricted range of overgeneralisations used in naturalistic contexts. In coding the data, to avoid artificially inflating error counts due to continuous repetition, identical repetitions of the same error were capped at two instances per child.

Table 5.1 *Plural Types in the Naturalistic Data with Five Examples of Each (where available)*

Plural Type	**Examples from the naturalistic data**					
English	***bubbles***	***chocolates***	***mummies***	***piggies***	***toys***	
Target-like SP	***bibien*** *bieb* 'doors'	***iljunfanti*** iljunfant 'elephants'	***karozzi*** karozza 'cars'	***siġġijiet*** siġġu 'chairs'	***wajers*** wajer 'wires'	Item produced Singular form Translation
Target-like BP	***għasafar*** *għasfur* 'birds'	***tfal*** *tifel/tifla* 'children'				Item produced Singular form Translation
Non-target-like SP/ type-noun plural	****brokkolis*** *brokk(o)la* *brokk(o)li* 'broccoli'	****tadami*** *tadama* *tadam(iet)* 'tomatoes'				Item produced Singular form Target-like pl Translation
Non-target-like BP	****furketti*** *furketta* *frieket* 'forks'	****pizziet*** *pizza* *pizez* 'pizzas'	****sikkini*** sikkina skieken 'knives'			Item produced Singular form Target-like pl Translation
Collective	***bajd*** *bajda* 'eggs'	***dud*** *dudu* 'worms'	***frott*** *frotta* 'fruit'	***kappar*** *kappara* 'capers'	***żebbuġ*** *żebbuġa* 'olives'	Item produced Singular form Translation

[74] As different researchers employ varying typologies for BPs, all references to BP types in this thesis are standardised to Schembri's (2006, 2012) framework, which provides a comprehensive and still widely adopted account of both SP and BP formation in Maltese. See Section 2.5.1.1 for an overview of the Maltese plural system.

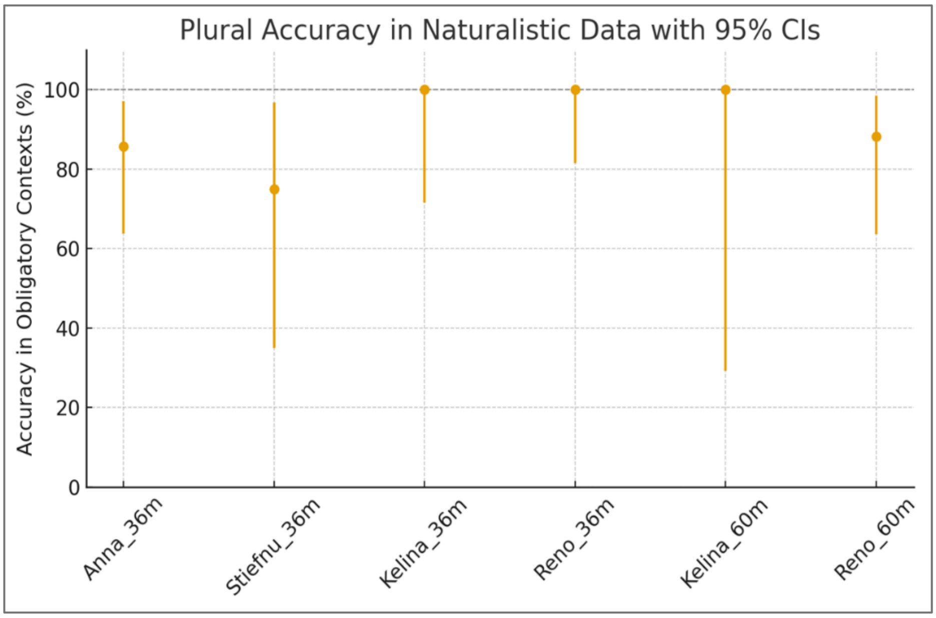
Children's spontaneous productions were examined for obligatory plural contexts. Accuracy was defined as the proportion of correctly produced plurals relative to all plural tokens (correct + errors). In keeping with the bilingual nature of the dataset, accuracy scores included both Maltese and English plural tokens; English forms were retained as they represent an important dimension of children's developing bilingual system. Collective nouns were excluded, as they typically function as unmarked base forms. To account for variation in token counts, accuracy estimates were supplemented with 95% binomial confidence intervals (CIs; Clopper–Pearson exact). CIs were compared against a chance baseline of 50%; narrow intervals (derived from many tokens) provide robust evidence of stability, whereas wide intervals (from few tokens) warrant caution in interpretation.

At 18 months, neither Anna nor Stiefnu produced plural forms. At 36 months, all four children attempted plurals, with accuracy typically above 75% and in some cases reaching 100%. At 60 months, both Reno and Kelina produced plurals: Reno continued to show high accuracy, while Kelina's very small number of tokens precluded reliable interpretation. Results are presented in Table 5.2 and visualised in Figure 5.2.

Table 5.2 *Accuracy of Plural Production in Obligatory Contexts by Child and Age*

Child	Age (m)	Correct/Total	Accuracy (%)	95% CI	Errors	Notes
Anna	18	—	—	—	—	No plural tokens
Stiefnu	18	—	—	—	—	No plural tokens
Anna	36	18/21	85.7	63.7–97.0	3 (BP)	Accuracy above chance
Stiefnu	36	6/8	75.0	34.9–96.8	2 (SP)	CI includes chance
Kelina	36	11/11	100	71.5–100	0	Stable, no errors
Reno	36	18/18	100	81.5–100	0	Narrow CI, robust evidence
Kelina	60	3/3	100	29.2–100	0	CI too wide; not interpretable
Reno	60	15/17	88.2	63.6–98.5	2 (BP)	Accuracy above chance

Figure 5.2 *Plural Accuracy in Naturalistic Data for Four Children with 95% CIs*



Because the dataset includes both Maltese and English productions, accuracy scores reflect overall plural performance across the two languages. The low number of Maltese plural tokens limits the reliable estimation of accuracy for Maltese alone:

in half the eligible cases, the minimum threshold of five obligatory contexts was not reached. Although children were reported as MD at the first timepoint, most spontaneous plurals were produced in English. This was anticipated to some extent and raises the possibility that children may have been drawing on English plural morphology as a strategy, whether consciously or unconsciously, to avoid non-target-like productions in Maltese. This possibility merits further investigation (see Section 5.3.1).

While accuracy, as defined above, could not be reliably calculated for Maltese items alone due to the low number of tokens, the association between plural type (SP vs BP) and correctness was analysed using Fisher's exact test, which is specifically designed for small contingency tables and does not rely on large-sample assumptions.[75] The test revealed a significant association between plural type and correctness ($p = .00352$), with substantially lower odds of correct production for BPs compared to SPs (odds ratio = 0.058, 95% CI [0.0038, 0.5265]). This indicates that, within the available Maltese data, children were more likely to produce SPs correctly than BPs in spontaneous speech.

Taken together, the naturalistic data provide preliminary support for Hypothesis 2A. Children show greater reliance on concatenative plural formation, which appears more frequently and with higher accuracy, whereas templatic plurals remain both infrequent and error-prone, even at later ages. Indeed, five of the seven observed non-target-like productions involve the BP, underscoring its status as the

[75] A mixed-effects model was not fitted due to the very small number of observations and sparse distribution of tokens across conditions. Fisher's exact test was therefore used as a more appropriate alternative for analysing the association.

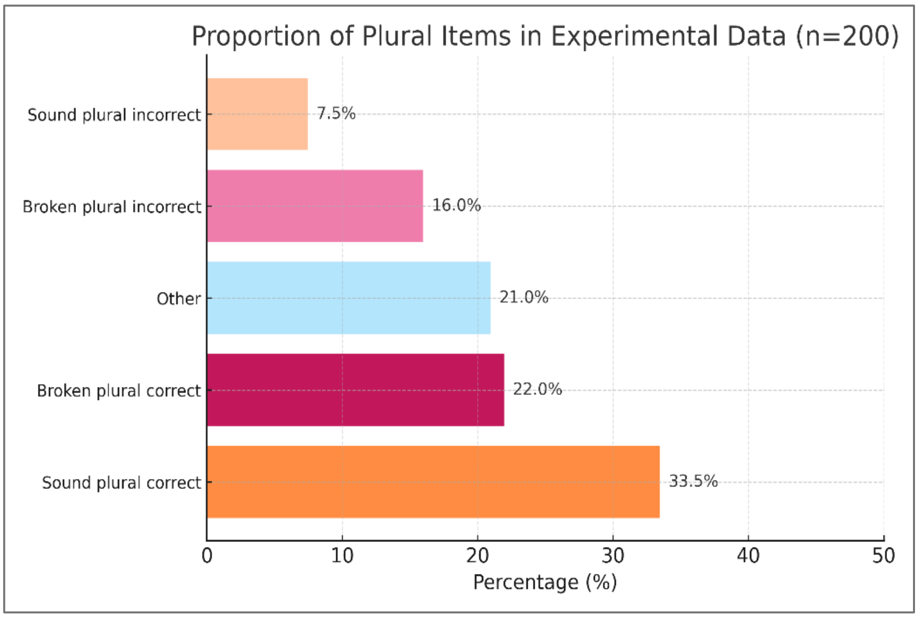
main locus of difficulty. At the same time, these findings must be interpreted with caution since the naturalistic data analysed contain a limited number of plural tokens, with even fewer tokens occurring in Maltese. To address these limitations and test Hypothesis 2A more systematically, performance was therefore examined in the existing-word plural task and the novel-word plural task.

5.2.1.2. Existing-word plural task

The existing-word plural task was administered to all eight children, with two children completing it twice, six months apart, yielding a total of ten task administrations. Each task consisted of 20 target items, resulting in 200 responses overall from children aged between 36–60 months. This task was designed to address three limitations of the naturalistic data: a) limited opportunities for plural elicitation, b) the possibility of responding in English, minimised here by presenting the singular forms in Maltese and actively involving a 'monolingual' puppet interlocutor, and c) the lack of direct comparability across responses. For younger children, the task was presented in a straightforward elicitation format, whereas for older children it was embedded in a bingo game. The two formats differed only in the order of item presentation (fixed in the elicitation version, random in the bingo version); no systematic differences in outcomes were observed. Full details of task design and administration are provided in Section 3.2.2.2.

Figure 5.3 presents children's responses in the existing-word plural task, allowing direct comparison with the proportions observed in spontaneous speech (Figure 5.2). The largest proportion of target-like responses involves SPs (33.5%), only slightly higher than in the naturalistic data (30.8%). Target-like BPs account for 22% of responses, remaining clearly lower than SPs despite the equal number of items of each type in the task. The category *other* (21%) comprises English responses, no responses, non-applicable responses, and unintelligible productions, all of which are common in elicitation tasks with young children. Compared with the naturalistic data, the proportion of English responses is substantially smaller (21% vs 56.4%). As anticipated, the elicitation format yields a higher proportion of non-target-like responses than spontaneous speech. Within these, BPs are particularly vulnerable: errors in the BP (16%) occur at more than twice the rate of errors in the SP (7.5%).

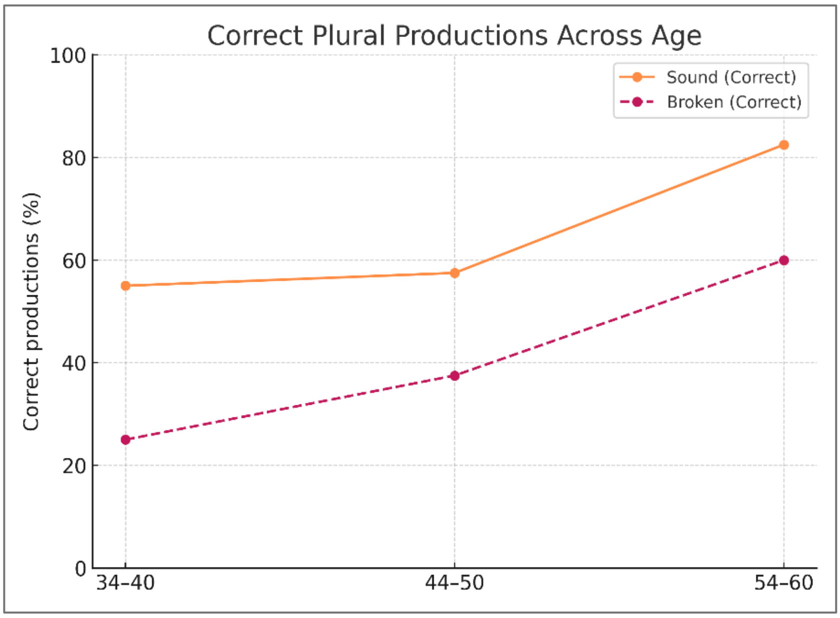
Figure 5.3 *Distribution of Response Types in the Existing-word Plural Task*



When target-like productions are examined across age groups, the developmental trajectories of the SP and BP appear broadly parallel, particularly in the steep increase observed between 44–50 and 54–60 months. However, the proportion of correct SP responses is consistently higher than that of correct BP ones at every stage. Accuracy for SPs rises from 55% at 34–40 months to 82.5% at

54–60 months, while accuracy for BPs lags behind, increasing from 25% to 60%. These results indicate that the BP persists as a challenge, even as overall performance improves with age. The developmental trajectories are illustrated in Figure 5.4.

Figure 5.4 *Developmental Trajectories of Target-like SPs and BPs Across Age Groups in the Existing-word Plural Task*



To complement the descriptive patterns shown above, plural accuracy was analysed using a mixed-effects logistic regression model (glmer) implemented in the lme4 package in R. The dependent variable was whether an individual plural attempt was correct (1) or incorrect (0). Responses involving English plurals or English singular forms were coded as NA and excluded from the analysis. Fixed effects included plural type (SP vs BP, contrast-coded as +0.5 and −0.5) and age in months (centred on the mean), as well as their interaction. Random intercepts were included for child and lexical item to account for IDs and item-level variability.

The model revealed a significant main effect of plural type, with SPs more likely to be produced correctly than BPs at the mean age ($\beta = 1.48$, SE = 0.57, z = 2.59, p = .0096). There was also a significant positive effect of age ($\beta = 0.10$, SE = 0.05, z = 2.20, p = .028), indicating that performance improved over time. The interaction between plural type and age was not significant ($\beta = 0.02$, SE = 0.05, z = 0.31, p = .76), suggesting that the advantage for SPs remained stable across the age range. Random intercepts for child and lexical item showed substantial variability, supporting their inclusion.

These results converge with the developmental patterns in Figure 5.4 and indicate that both plural type and age independently shape children's performance in the elicited task. The effects of plural type and age are visualised in Figures 5.5 and 5.6.

Figure 5.5 *Predicted Probability of Correct Responses by Plural Type in the Existing-word Plural Task*

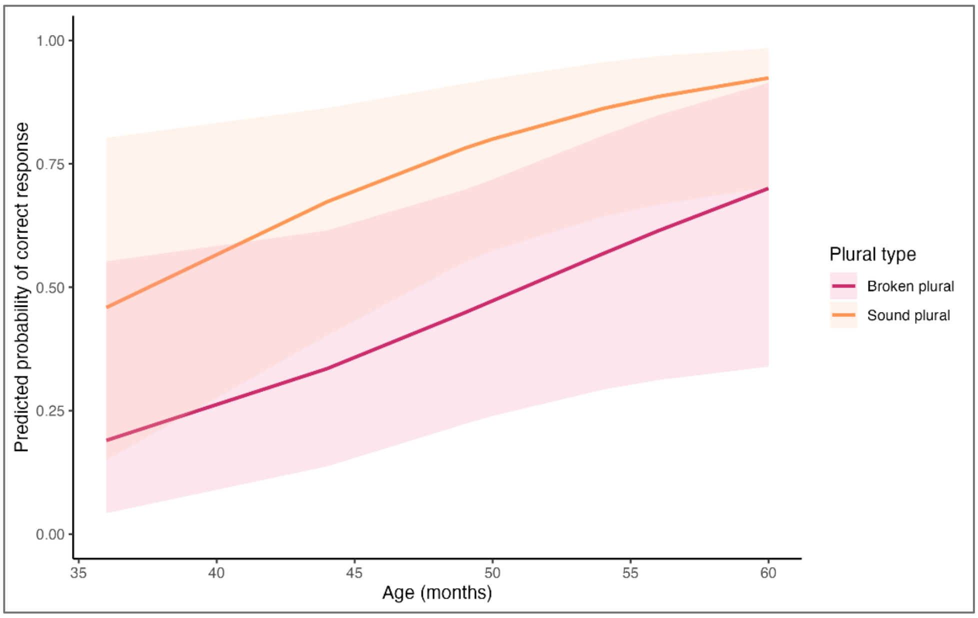


Figure 5.6 *Predicted Probability of Correct Responses Across Age in the Existing-word Plural Task*

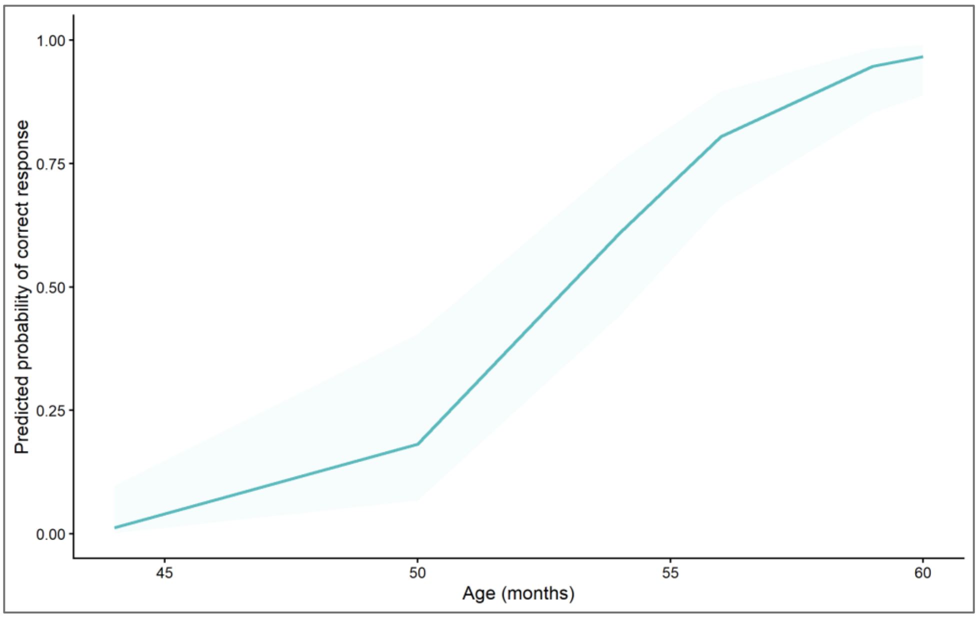


Figure 5.5 shows the model-predicted probabilities of correct responses across age for SPs and BPs. The probability of a correct response increases over the tested age range for both plural types; however, SPs consistently show a higher probability of correct production than BPs, in line with the significant main effect of plural type. Figure 5.6 presents the predicted probability of correct responses across age, averaged across plural types, and reveals a clear developmental trend, with accuracy increasing steadily with age, particularly after 50 months, reflecting the significant effect of age in the model.

In addition to target-like responses, children's non-target-like productions were coded and classified by strategy. These strategies, along with illustrative examples, are presented in Table 5.3. A detailed breakdown of one test item and the full range of responses it elicited across children are provided in Table 5.4.

Table 5.3 *Categories of Non-target-like Strategies in the Existing-word Plural Task, with Examples and Abbreviations*

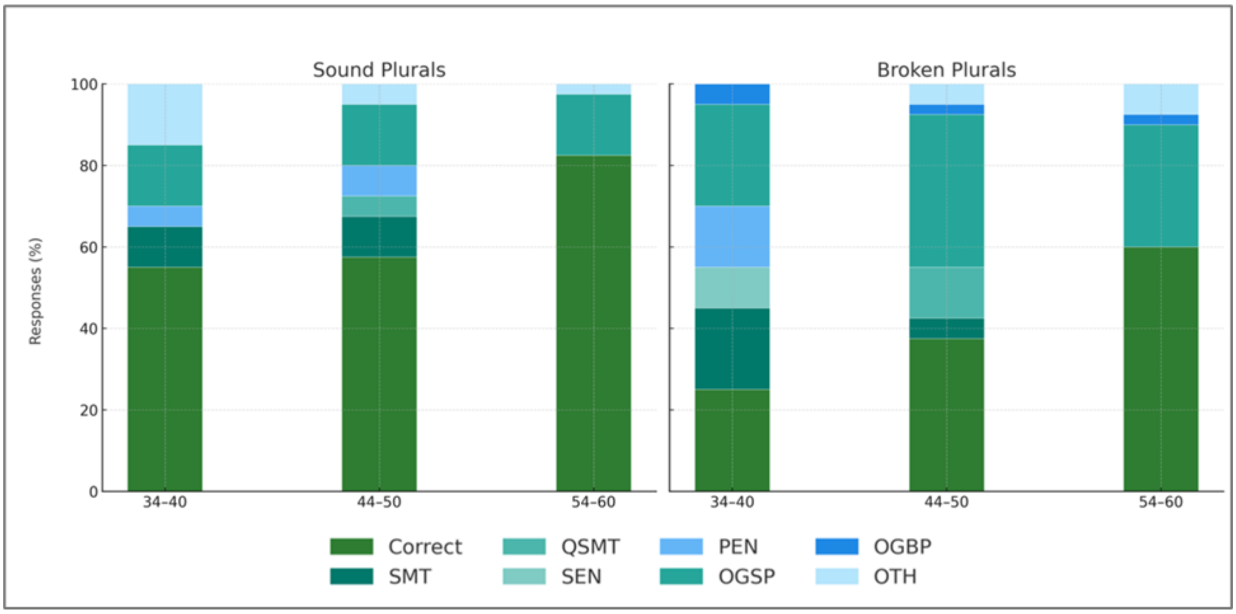
Strategy	Example	Target	Abbreviation
Repetition of test item (singular in Maltese)	*platt - platt* 'plate'	*platti*	SMT
Quantifier + singular in Maltese	*siġġu - *ħafna siġġu* 'chair'	*siġġijiet*	QSMT
English singular equivalent	*farfett - butterfly* 'butterfly'	*friefet*	SEN
English plural equivalent	*ktieb - books* 'book'	*kotba*	PEN
Overgeneralisation of SP suffixes	*mejda - *mejdi* 'table'	*imwejjed*	OGSP
Overgeneralisation of BP patterns	*ħabel - *ħbiebel* 'rope'	*ħbula*	OGBP
No/non-applicable response, unintelligible	—	—	OTH

Table 5.4 *Range of Responses Produced for a Single Test Item in the Existing-word Plural Task*

Item tested	Target	Responses						
serp 'snake'	*sriep*	*serp*	*sriep*	*snake*	*snakes*	*ħafna serp*	*serpijiet*	*sriepen*

As shown in Figure 5.7, both the diversity of non-target-like strategies and the overall number of non-target-like responses decrease with age. For SPs, the number of strategies decreases from four at 34–40 months to two at 54–60 months. For BPs, five strategies are attested at 34–40 months, reducing to three by 54–60 months.

Figure 5.7 *Proportion of All Responses Across Age Groups*

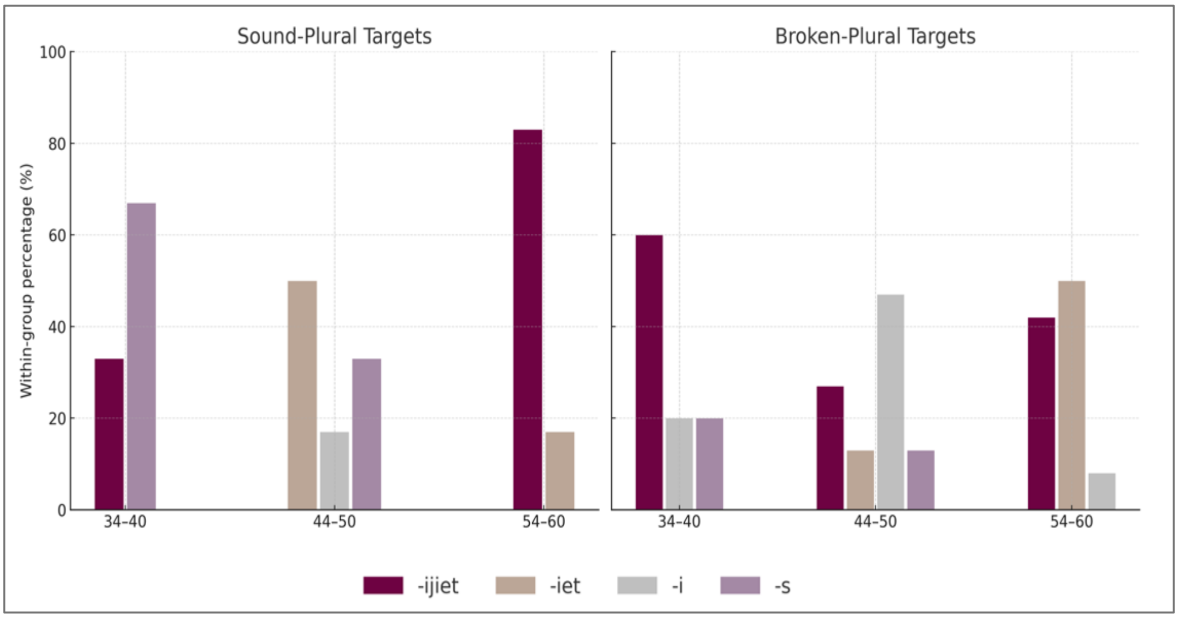


Across all ages and both plural types, the most frequent non-target-like strategy is the overgeneralisation of SP suffixes. This occurs in two main ways: a) with nouns that take a different SP suffix than the one supplied, or b) with nouns that in fact require a BP. The suffixes most often overgeneralised are *-ijiet, -i, -iet,* and *-s* (see Figure 5.8). Table 5.5 illustrates this pattern with examples from responses to *ġakketta* ('jacket').

Table 5.5 *Examples of Overgeneralisation of SP Suffixes in the Existing-word Plural Task*

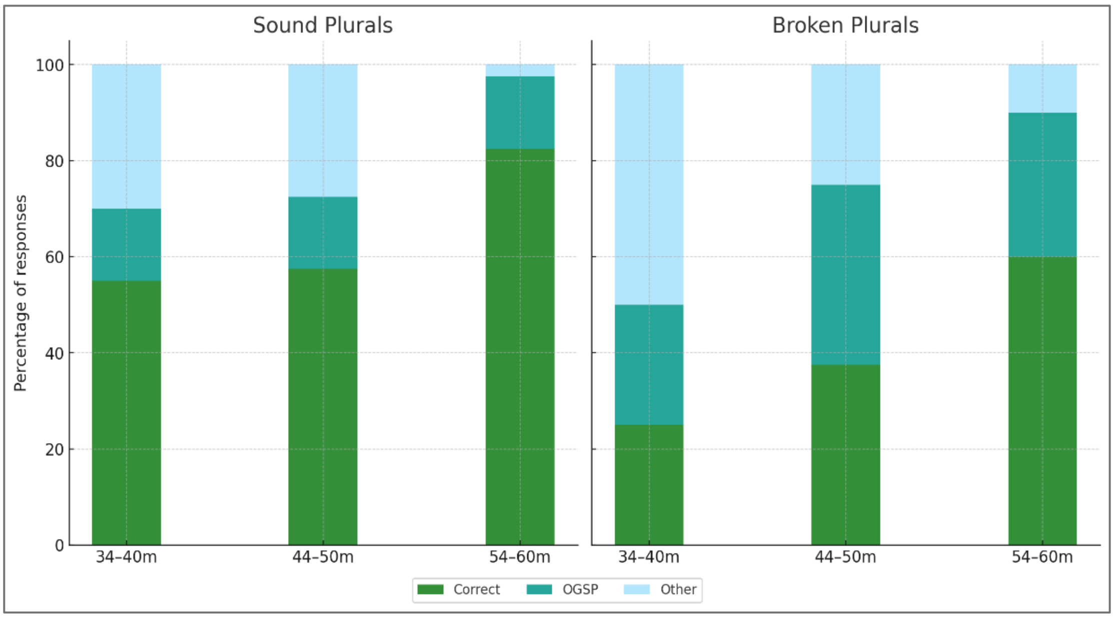
Item tested	Target	Overgeneralisation of SP suffixes			
ġakketta 'jacket'	*ġkieket*	*ġakkett**ijiet***	*ġakkett**iet***	*ġakkett**i***	*ġakketta**s***

Figure 5.8 *Overgeneralisation of SP Suffixes Across Age Groups in the Existing-word Plural Task*



To assess developmental change across age groups, Pearson's chi-squared tests of independence were conducted. These tests are well-suited to categorical data, allowing evaluation of whether the distribution of response types differ significantly across groups. For the purposes of statistical analysis, response strategies in the elicited plural task were consolidated into three categories: a) *target-like*, b) *overgeneralisation of SP suffixes* (*OGSP*), and c) *other non-target forms*. Specifically, because of the small cohort size and the resulting low cell counts, less frequent response types (e.g., singular forms, quantifier + singular, English responses) were collapsed into the *other* category. This step ensured that expected cell frequencies met the assumptions of the chi-squared test. The resulting three-way classification (*correct, OGSP, other*) permitted meaningful analysis while avoiding overinterpretation of sparse data (Figure 5.9).

Figure 5.9 *Proportion of Response Types Classified as Target-like, Overgeneralisation of the SP Suffix (OGSP), and Other Non-target Forms*



With respect to SPs, the distribution of responses differs significantly across age groups, $\chi^2(4, n = 100) = 11.35$, $p = .023$ (Table 5.6). Pairwise comparisons indicate no difference between 34–40 and 44–50 months; however, both younger groups differ significantly from the 54–60-month group. Thus, performance remains stable until around 50 months, after which a marked shift toward target-like production emerges.

A parallel analysis of BPs also reveals a significant overall effect of age, $\chi^2(4, n = 100) = 14.00$, $p = .007$ (Table 5.6). The developmental trajectory, however, differs in the production of BPs as compared to that of SPs. Pairwise tests show no reliable difference between 34–40 and 44–50 months, and no significant difference between 44–50 and 54–60 months. Only the youngest and oldest groups differ significantly, $\chi^2(2) = 12.64$, $p = .002$. This suggests gradual improvement over time, with a less consistent pathway and persistent overgeneralisations even for the oldest children. The results for both the SP and BP are tabulated in Table 5.6, below.

Table 5.6 *Response Distributions of SPs and BPs Across Age Groups, with Chi-square Comparisons*

Plural type	**Age (m)**	**Target (%)**	**OGSP (%)**	**Other (%)**	**χ^2 (2) vs 34–40m**	***p***	**χ^2 (2) vs 44–50m**	***p***
Sound	34–40	55.0	15.0	30.0	—	—	—	—
	44–50	57.5	15.0	27.5	0.04	.98	—	—
	54–60	82.5	15.0	2.5	10.02	.007	10.12	.006
Broken	34–40	25.0	25.0	50.0	—	—	—	—
	44–50	37.5	37.5	25.0	3.75	.15	—	—
	54–60	60.0	30.0	10.0	12.64	.002	4.98	.083
Global tests								
Sound	χ^2(4, n = 100) = 11.35, p = .023							
Broken	χ^2(4, n = 100) = 14.00, p = .007							

Together, the analyses demonstrate significant age-related change in the production of both plural types. SPs undergo a late yet decisive shift toward accuracy by 54–60 months, whereas BPs display a more incremental and inconsistent trajectory. While the limited cohort size and the need to collapse categories warrant cautious interpretation, the overall pattern supports Hypothesis 2A: concatenative plural forms are acquired with greater accuracy and stability than templatic forms. These findings suggest that TD, MD children follow distinct developmental pathways in mastering the two plural systems.

5.2.1.3. Novel-word plural task

The novel-word task, which required greater cognitive maturity, namely owing to its metalinguistic demands, was administered only to children aged 44 months or older. Consequently, the task was not conducted with the entire cohort. A total of six children completed the task, which comprised 20 items, yielding a total of 120 responses. Details of the task design and administration are provided in Section 3.2.2.2.

In addition to spontaneous speech and elicited plural production data, the novel-word task provides a stringent test of morphological productivity. Whereas spontaneous and elicited data may be influenced by rote-learned lexical knowledge, novel items require children to extend their morphological knowledge to novel forms. In this sense, the novel-word task offers insights into the extent to which the observed plural patterns reflect generalisable rules rather than memorised forms.[76]

Responses were categorised in line with the elicited plural task in order to facilitate comparison. Five categories were distinguished: a) *application of a SP suffix*, b) *use of a BP pattern*, c) *repetition of the singular*, d) *other* (encompassing no answer, unattested forms and altered constituents), and e) *final vowel drop*. The latter was introduced as a separate category, as it occurred with notable frequency in both age groups and thus warranted independent consideration. The response strategies, along with illustrative examples from the data and abbreviations, are listed in Table 5.7. A detailed breakdown of one test item and the full range of responses it elicited across participants are presented in Table 5.8.

[76] While novel items are especially informative for assessing morphological productivity, non-target-like productions in familiar lexical items may also reflect productive use of morphological knowledge, insofar as they extend beyond forms available in the input.

Table 5.7 *Categories of Strategies in the Novel-word Plural Task, with Examples and Abbreviations*

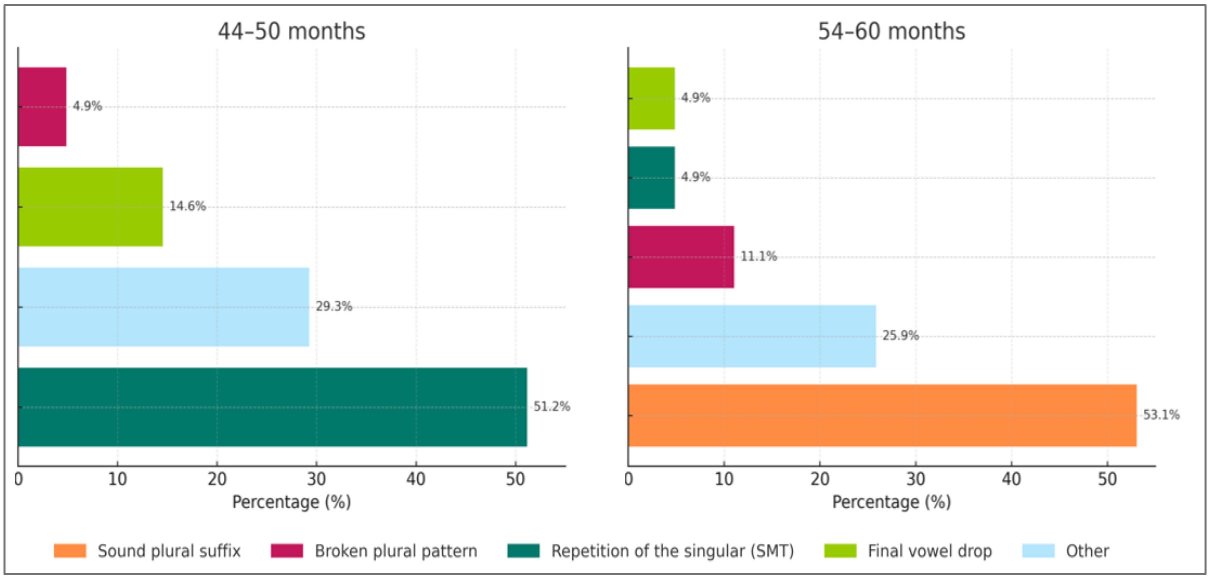
Strategy	Example	Abbreviation
Repetition of test item (singular)	*tekkina - tekkina*	SMT
Quantifier + singular in Maltese	*qetties - ħafna qetties*	QSMT
SP suffixes	*peġġa - peġġijiet*	SP
BP patterns	*torħa - toroħ*	BP
Vowel drop	*lepa - lep*	VDROP
Unattested plural form	*pannat - pannato*	UNA
Changed material	*rulipa - ritz*	CNG
No/non-applicable response, unintelligible	*ma nafhiex 'I don't know it'*	OTH

Table 5.8 *Range of Responses Produced for a Single Test Item in the Novel-word Plural Task*

Item tested	Responses					
qetties /ʔɛt'tɪːs/	qetties /ʔɛt'tɪːs/	ħafna qetties /'hɐfnɐ ʔɛt'tɪːs/	qettiesi /ʔɛt'tɪːsɪ/	qettisi /ʔɛt'tiːsɪ/	qties /ʔtɪːs/	qetti /ʔɛt'tɪ/

The results are displayed in Figure 5.10. Several developmental patterns are evident. First, repetition of the singular emerges as the most frequent response strategy at 44–50 months; however, its prevalence decreases sharply by 54–60 months, where it became the least common category. Secondly, the use of the SP suffix, absent altogether at 44–50 months, becomes the dominant strategy in the older group (at 53.1%), indicating the emergence of productive application of concatenative morphology. Among the suffixal responses, the most frequent are *-i* (n = 26) and *-ijiet* (n = 10), with only a handful of cases of *-(i)s* (n = 5), *-iet* (n = 1), and *-ien* (n = 1). Thirdly, although the application of BP patterns is relatively infrequent overall, it increases with age, suggesting that children begin to generalise templatic morphology, albeit less robustly. Finally, while the *other* category remains relatively stable, the *final vowel drop* strategy decreases, suggesting that alternative strategies are still in use yet on the decline.

Figure 5.10 *Percentage Distribution of Response Strategies in the Novel-word Plural Task Across Age Groups*



To analyse children's responses to the novel nouns, a logistic regression model without random effects was fitted, as this provided stable parameter estimates. An initial model including interactions between plural contrasts and age resulted in unstable estimates and extremely large standard errors; therefore, the final model retained only main effects of plural type and age.

The final model (correct ~ contrast_sound_vs_broken + contrast_SB_vs_rest + age_c) specified correctness as the dependent variable and plural-type contrasts and centred age as fixed predictors. Here, contrast_sound_vs_broken compared SPs and BPs of the base nouns,[77] while contrast_SB_vs_rest compared sound–broken items with the average of the sound and broken conditions. For the purposes of the regression analysis, responses were coded as correct if they involved the application of an existing plural suffix or pattern, and incorrect otherwise.

The analysis revealed a significant effect of age ($\beta = 0.49$, $SE = 0.10$, $z = 4.77$, $p < .001$), providing evidence that children became increasingly likely to apply an existing plural suffix or pattern to novel nouns as they grew older. In contrast, neither plural-type effect was significant. The comparison between SP and BP bases was not significant ($\beta = -0.32$, $SE = 0.82$, $z = -0.39$, $p = .694$), indicating that children did not show a systematic preference for either plural type based on the base noun. The result therefore suggests that children are not systematically guided by the plural patterns of the base nouns when pluralising novel items. However, this may in part reflect the fact that the base nouns were not drawn from child-directed vocabulary as they were taken directly from Nieder et al. (2020) and may therefore have been unfamiliar to the children (see Section 3.3.2.2).

The comparison between sound–broken items and the average of the sound and broken conditions was likewise not significant ($\beta = 0.18$, $SE = 0.99$, $z = 0.18$, $p = .859$). This interpretation is further supported by descriptive patterns: the use of an existing plural suffix or pattern was similar across conditions (sound: 19/33 ≈ 58%; broken: 17/30 ≈ 57%; sound–broken: 18/28 ≈ 64%).

[77] Base nouns refer to the existing lexical items from which the novel forms were derived.

Including the phonological manipulation variable (change_cv) did not improve model fit and was also not significant ($p = .442$), reflecting the fact that the type of segmental modification used to create the novel words did not influence children's responses. This contrasts with findings reported for Maltese adults (Nieder et al., 2020), where such manipulations affected performance. One possible explanation is that children's relative insensitivity to these phonological manipulations, like their lack of reliance on base noun plural patterns, reflects limited familiarity with the underlying lexical items. This interpretation is further supported by descriptive patterns: the use of an existing plural suffix or pattern was similar across conditions (sound: 19/33 ≈ 58%; broken: 17/30 ≈ 57%; sound–broken: 18/28 ≈ 64%).

To further examine whether children's responses in the novel-word plural elicitation task differed across age groups, Pearson's chi-squared tests of independence were conducted. Given the limited cohort size, Fisher's exact tests were also applied to contrasts of particular interest (SP and BP usage), as these are more robust with sparse data. The omnibus chi-squared test reveals a significant effect of age group on response distribution, $\chi^2(4, n = 120) = 54.63$, $p < .0001$. Follow-up analyses indicate a sharp increase in SP usage among the older group and Fisher's exact test confirms this difference as highly significant. In contrast, BP usage does not differ significantly between groups. The younger children rely heavily on repetition of the singular and *other* responses, whereas the older children shift toward productive pluralisation strategies, highly favouring SP suffixation. These results, together with the percentage proportions corresponding to Figure 5.8 above, are tabulated in Table 5.9.

Table 5.9 *Response Distributions in the Novel-word Plural Task Across Age Groups, with Chi-squared and Fisher's Exact Test Results*

Age (m)	SP %	BP %	Repetition of singular %	Final vowel drop %	Other %
44–50	0.0	4.9	51.2	14.6	29.3
54–60	53.1	11.1	4.9	4.9	25.9
Omnibus test					
$\chi^2(4, n = 120) = 54.63$, $p < .0001$					
Pairwise contrasts (Fisher's exact tests)					
Sound	44–50m vs 54–60m, $p < .0000000002$				
Broken	44–50m vs 54–60m, $p = .33$				

In sum, although the sample size is small, the pattern is clear: by 54–60 months, children shift from reliance on singular repetition toward productive use of the SP suffix, while BP generalisations remain relatively infrequent. This developmental trajectory mirrors the existing-word plural task, where SP formation consolidates into a productive process by 54–60 months, whereas BPs lag behind. The novel-word task thus provides limited but emerging evidence of BP generalisation, alongside clearer productivity in the SP domain, supporting Hypothesis 2A. The implications of the vowel-drop strategy and the asymmetry between concatenative and templatic forms are discussed in Section 5.3.1.

Taken together, the findings from the naturalistic, elicited, and novel-word tasks converge on a consistent asymmetry between SPs and BPs. While children demonstrate early and increasingly productive use of SPs, BPs remain less stable in production and show limited generalisation to novel forms. This convergence across datasets suggests that concatenative morphology becomes productive earlier, whereas templatic morphology develops more gradually and remains comparatively constrained.

5.2.2. Hypothesis 2B: Nominal agreement

Hypothesis 2B concerns target-like accuracy in gender and number agreement, investigated in both spontaneous speech and through a targeted elicitation task focusing on gender agreement. As outlined earlier, the analysis of agreement in both naturalistic and experimental data is restricted to relations among nouns, pronouns, demonstratives, adjectives (and adjective-like elements), and quantificational expressions (including numerals). Other types of agreement, such as subject–verb agreement, are excluded. Throughout this section, the term *agreement* therefore refers exclusively to these domains unless otherwise specified.

5.2.2.1. Naturalistic data

Across the eight transcriptions, 134 obligatory contexts requiring nominal gender or number agreement were identified and coded. Of these, 31 were excluded from the calculation of formal grammatical agreement scores. Two types of non-standard agreement motivated exclusion. The first involved invariant targets, namely adjectives lacking a full inflectional paradigm, such as invariable colour terms (e.g., *roża* 'pink') and English adjectives, both common in children's speech and CDS in Maltese. The second involved semantic agreement, in which the form of the agreeing element was determined by the semantic rather than grammatical properties of the noun. In this corpus, semantic agreement is limited to the noun *pulizija* ('police.SGM/SGF/PL'), where the target adjective inflects according to the real-world referent rather than the noun's grammatical features. The majority of exclusions (22/31) involved English adjectives. Detailed figures, percentages, and illustrative examples of each exclusion type are provided in Table 5.10.

The remaining 103 contexts involved agreement phenomena, comprising 74 instances of gender agreement and 29 instances of number agreement. Gender agreement occurred more frequently than number agreement, accounting for 3.4% of all utterances compared to 1.3%. In both domains, contexts in which the controller[78] was an English lexical item (and therefore morphologically invariant) were included if the target was in Maltese; these were coded as *code-switched utterance agreement*. For gender agreement, 14 such cases were identified. These correspond to what the literature describes as *agreement by analogy*, in which an English controller noun appears to be assigned the grammatical gender of its closest Maltese translation equivalent (e.g., *sock kbira* 'a big sock', with *sock* showing feminine agreement consistent with Maltese *kalzetta*). In this dataset, only one non-target-like production was recorded within this category (see Table 5.11), which otherwise reflects assignment of grammatical gender based on the closest Maltese translation equivalent.

For number agreement, 29 cases were identified, of which 10 were code-switched utterances. Unlike gender agreement, these were not instances of agreement by analogy, rather most involved direct number agreement within a code-switched phrase. Such cases occurred when the English controller carried an overt plural marker (e.g., *cars*), which was matched with a Maltese plural target (e.g., *dawk cars* 'those are cars'). Within this category, two non-target-like productions were observed, both involving an English numeral as controller combined with a singular Maltese noun as target, which would otherwise take the BP (e.g., *two mejda*

[78] *Controller* refers to the element that determines the agreement features (e.g., gender or number) expressed on another element.

'two table'; *mejda* – *imwejjed* 'tables', see Table 5.11). In these cases, the numeral provided a [+PLURAL] feature, but the corresponding plural form of the noun was not applied.

Excluding the code-switched items, 79 fully Maltese cases remained. Of these, 60 involved gender agreement, comprising 35 with a feminine controller and 25 with a masculine controller. A total of 10 non-target-like productions were observed: 4 involved a feminine controller with a masculine target, and 6 involved a masculine controller with a feminine target (see Table 5.11 for examples). For number agreement, 19 fully Maltese cases were identified, of which 3 contained errors, all involving a singular quantifier combined with a plural noun (see Table 5.11 for an example). One additional case is noteworthy: **għajnejh ħamra* 'his eye is red'. Here, the noun *għajn* 'eye' is grammatically feminine but forms the dual with the suffix *-ejn* (requiring plural agreement) and appears with a masculine singular clitic pronoun (*-h* 'his'). The adjective *ħamra* 'red.SGF' was used, yielding a mismatch between the noun in the dual form and the adjective in the feminine singular.

Table 5.10 *Response Distributions in the Novel-word Plural Task Across Age Groups, with Chi-squared and Fisher's Exact Test Results*

Category	*n*	% of total	Example
Total contexts	**134**	100.0	
Excluded	**31**	23.1	
Invariant adjectives	7	5.2 22.6 (of excluded)	*ħanżir roża* pig.SGM pink 'a pink pig'
English adjectives	22	16.4 71.0 (of excluded)	*lumija /yellow/* lemon.SGF yellow 'a yellow lemon'

Semantic agreement	2	1.5 6.5 (of excluded)	*dak* *pulizija* that.SGM police 'that's a policeman'
Analysed contexts	**103**	76.9	
Gender agreement	74	55.2 71.8 (of analysed)	*il-mamà* *rieqda* DEF-mum-SGF asleep.SGF 'mum is asleep'
Number agreement	29	21.6 28.2 (of analysed)	*karozzi* *sbieħ* car.PL nice.PL 'nice cars'

Note. Percentages are calculated relative to the total number of contexts (*n* = 134) unless otherwise specified in parentheses.

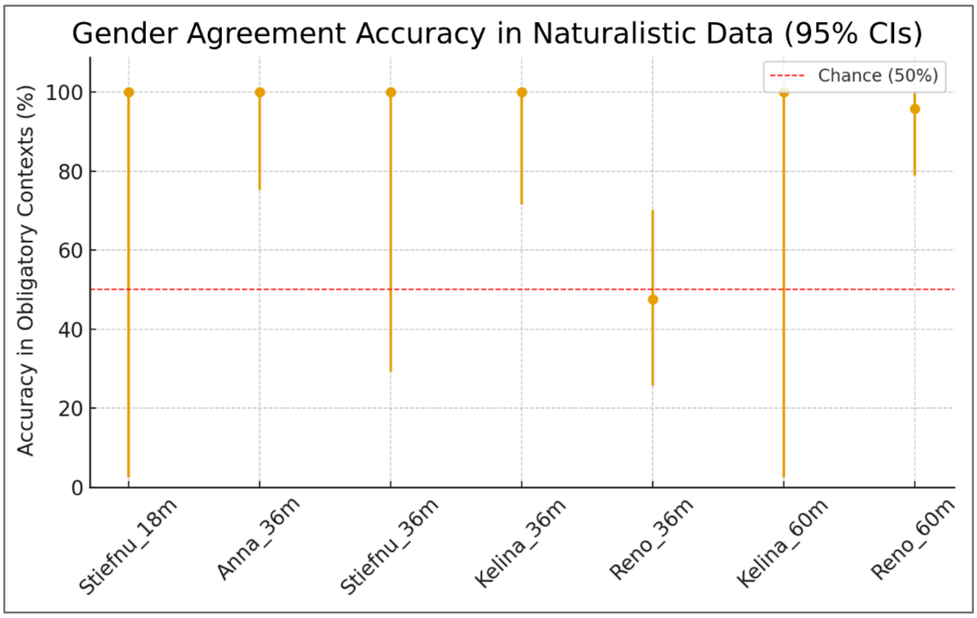
Table 5.11 *Gender and Number Agreement Errors*

Category	Error Rate (%)	Error Example
Gender Agreement	**14.9**	
Mt. controller, Mt. target	13.5	**dak* *ħabiba* that.SGM friend-SGF 'that is a friend'
En. controller, Mt. target	1.4	**dak* */beach/* that.SGM beach (Mt. beach=bajja.SGF) 'that is a beach'
Number Agreement	**17.2**	
Mt. controller, Mt. target	10.3	**dak* *annimali* that.SGM animal.PL 'that is animals'
En. controller, Mt. target	6.9	**/two/* *mejda* two table.SGF 'two table'

Overall gender and number agreement error rates are similar (14.9–16.7%). However, different developmental patterns were observed: gender agreement is attested in most children from 36 months onward, whereas number agreement remains infrequent and less consistent. These results suggest that gender agreement becomes productive earlier, while number agreement emerges later and with lower frequency in spontaneous speech. To examine developmental patterns more closely, accuracy rates in obligatory contexts were then analysed by age. Percentages are reported at 95% CIs to account for the small number of tokens available in many cases.

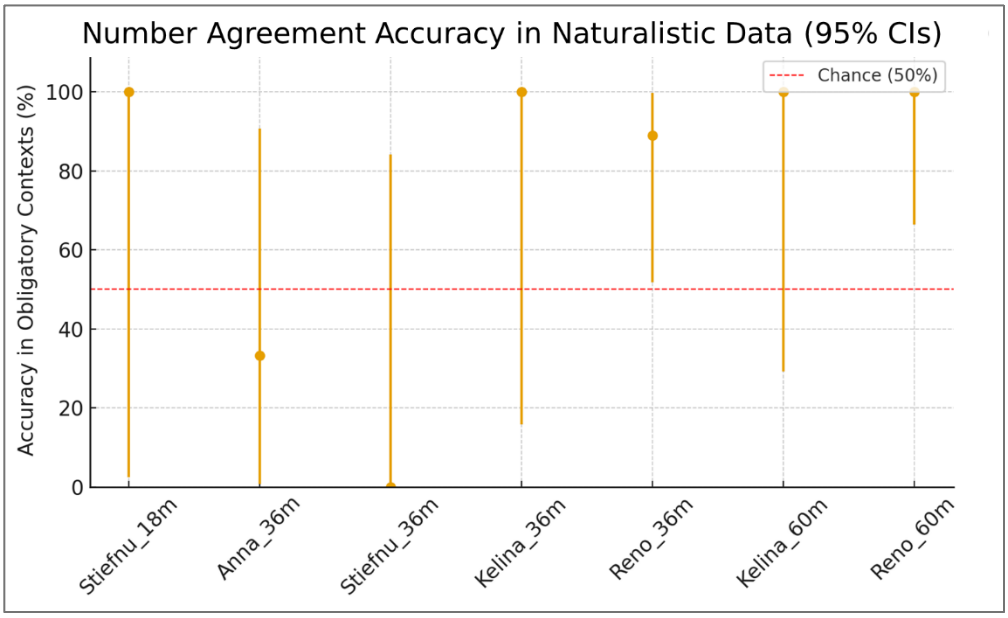
Accuracy in obligatory gender agreement contexts is highly variable in the early data but shows clear developmental progression. At 18 months, only one child, Stiefnu, produced a single instance of gender agreement, and the wide CI renders this estimate uninformative. By 36 months, three of the four children (Anna, Stiefnu, and Kelina) produced gender agreement consistently whenever attempted though low token counts again limit the stability of these estimates. Reno, however, displayed marked variability at this age, producing fewer than half of his tokens with correct agreement. By 60 months, Reno's accuracy reaches near ceiling, supported by a large number of tokens, providing the clearest evidence of stabilisation in this domain. In contrast, Kelina produced only one token at this age, which precludes interpretation of her performance. Figure 5.11 presents these results, with error bars representing 95% Clopper–Pearson exact CIs. Results are shown separately for each child to illustrate individual variability and developmental progression.

Figure 5.11 *Accuracy in Obligatory Gender Agreement Contexts Across Ages*



Accuracy in obligatory number agreement contexts was virtually absent at 18 months, with only one token produced. By 36 months, all four children attempted number agreement, but production remained limited and variable: Anna and Stiefnu showed low accuracy, Kelina produced only a small number of correct tokens, and Reno showed relatively high accuracy but on the basis of few data points. By 60 months, Reno produced nine correct instances with no errors, providing the clearest evidence of stability, while Kelina produced three correct instances, too few to interpret with confidence. Figure 5.12 summarises these results, illustrating the limited frequency of nominal number agreement, its variability at 36 months, and the more stable, yet still variable, pattern that emerges by 60 months.

Figure 5.12 *Accuracy in Obligatory Number Agreement Contexts Across Ages*



To complement the descriptive and age-based analyses presented above, a regression analysis was conducted to test whether agreement accuracy differed systematically between gender and number agreement in the naturalistic data. A logistic regression model without random effects was fitted because the corresponding mixed-effects model yielded a boundary (singular) fit with a random intercept variance of zero, indicating that the inclusion of random effects was not supported by the data.

The final model (*correct ~ agreement_c*) specified agreement accuracy as the dependent variable and agreement type as the sole fixed predictor. Agreement type was contrast-coded (gender = +0.5, number = −0.5), such that the intercept represents the grand mean accuracy and the coefficient reflects the difference between gender and number agreement. For the purposes of the regression analysis, responses were coded as correct if the target agreed appropriately with the controller in the relevant feature (gender or number), and incorrect otherwise.

The analysis revealed no significant effect of agreement type (β = 0.01, SE = 0.08, t = 0.13, p = .901), indicating that children did not differ reliably in their accuracy across gender and number agreement. The intercept was significant (β = 0.83, SE = 0.04, t = 20.28, p < .001), reflecting a high overall level of agreement accuracy across the dataset.

This result contrasts with the descriptive patterns reported above, which suggest earlier emergence and stabilisation of gender agreement relative to number agreement. One possible explanation for this discrepancy is that the regression model captures accuracy differences, whereas the descriptive analysis reflects frequency and distribution across developmental stages. In addition, the overall high accuracy rate (86/103 ≈ 83.5%) suggests a potential ceiling effect, which reduces the variability necessary to detect differences between agreement types. The relatively small and imbalanced dataset (gender: 74 tokens; number: 29 tokens) further limits statistical power.

The findings from the naturalistic data point a pattern in which gender agreement emerges around age three and reaches stabilisation by age five, whereas number agreement remains less frequent and more variable even at later ages. Framed against Hypothesis 2B, this asymmetry points to a developmental trajectory in which gender agreement stabilises earlier than number agreement. However, the regression analysis does not reveal a statistically reliable difference in accuracy between the two domains. Taken together, the findings support the first component of Hypothesis 2B, namely that accuracy in gender and number agreement increases with age. The second component, that gender agreement stabilises earlier than number agreement, is not confirmed by the present analysis and remains a possibility for future research to investigate using larger and more balanced datasets. The following section turns to an elicitation task targeting gender agreement only, providing a denser dataset to evaluate children's productive accuracy and to test whether its developmental course reflects the trajectory suggested by the naturalistic findings.

5.2.2.2. Gender agreement elicitation task

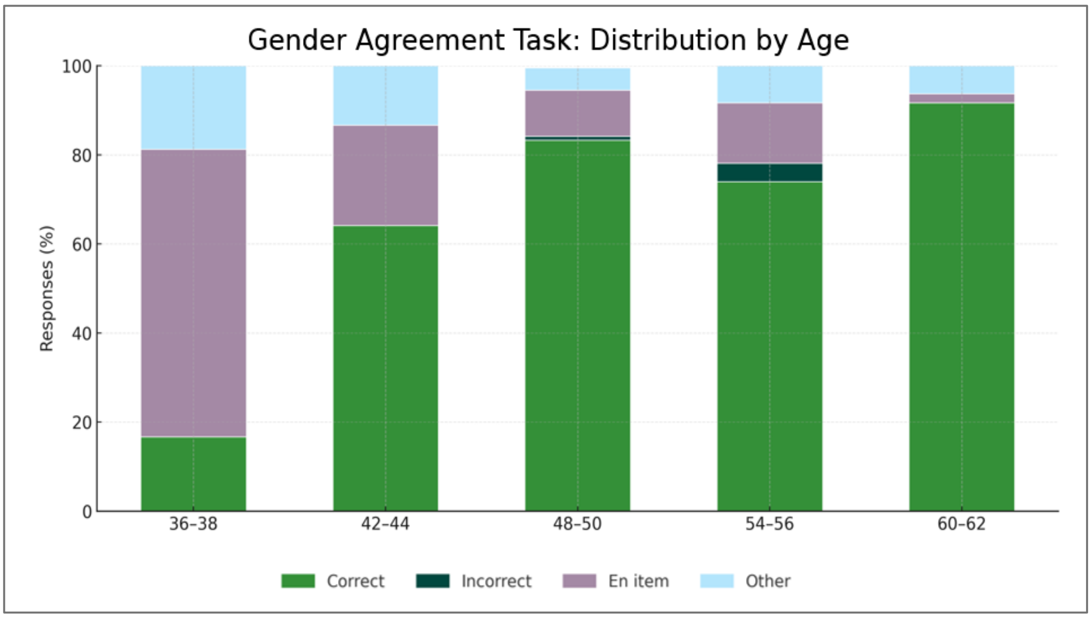
While the naturalistic corpus provides valuable insights into children's use of agreement in everyday contexts, the limited number of obligatory contexts restricts the conclusions that can be drawn. To complement the spontaneous speech findings, an elicitation task was designed to probe gender agreement more systematically. This task was administered 18 times in total, with each child participant in this study completing it on two or three occasions at least six months apart. The task was conducted with children aged 36–60 months, grouped into five age bands. It was designed to elicit six pairs of opposite adjectives (e.g., *kbir* 'big' vs *żgħir* 'small') in both masculine and feminine contexts, yielding 24 adjectival responses per administration. Details of task design and administration are provided in Section 3.2.2.2.

Responses were coded into four categories: a) *target-like* (grammatically correct, whether or not the precise target adjective was produced), b) *non-target-like*, c) *non-applicable* (e.g., no response, invariant adjective), and d) *English responses*. Examples of each category are provided in Table 5.12. Figure 5.13 presents the distribution of responses across these categories by age group. As illustrated, target-like responses increase steadily with age, while non-target-like and English responses decrease.

Table 5.12 *Examples of Response Categories in the Opposites Task*

Category	Example
Target-like	*ballun kbir u ballun żgħir* ball.SGM big.SGM and ball.SGM small.SGM 'a big ball and a small ball'
Non-target-like	**ħanżir nodfa* pig.SGM clean-PL 'a clean pig'
Non-applicable	*baqra iżgħar* cow.SGF smaller.CMP 'a smaller cow'
English response	*/clean/ u /yucky/* 'clean and yucky'

Figure 5.13 *Distribution of Response Categories in the Gender Agreement Elicitation Task by Age Group*



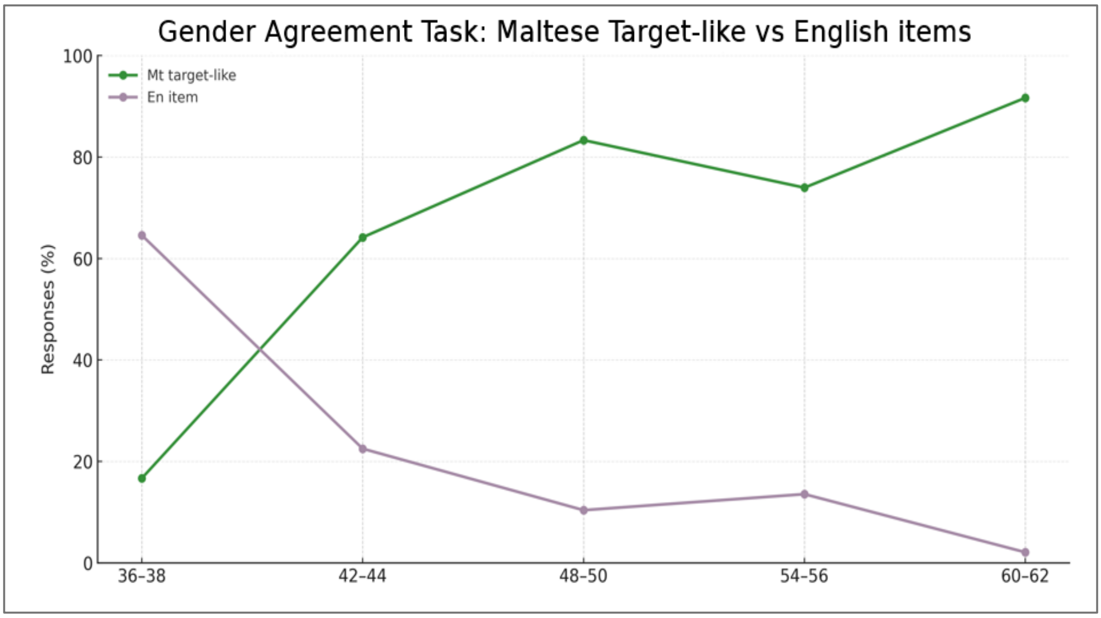
Due to the relatively small number of participants per group (n = 2–5), for the purposes of statistical analysis, categories (b–d) were collapsed and accuracy was analysed as target-like versus other. Accuracy varied significantly with age, $\chi^2(4, N = 431) = 88.87$, $p < .0001$. As shown in Table 5.13, the sharpest developmental gains occurred between 36–38 months and 42–44 months, where accuracy rose from low to moderate levels. Performance continued to improve in the 48–50-month group, which outperformed the 42–44-month group. In the middle period, accuracy plateaued, with no significant differences between 48–50 and 54–56 months. By 60–62 months, children reached near-ceiling accuracy, significantly higher than at 42–44 and 54–56 months. Note that accuracy in Table 5.13 is expressed as the proportion of target-like responses. Pairwise comparisons were conducted using Fisher's Exact Test.

Table 5.13 *Accuracy in Gender Agreement Elicitation Task by Age Group*

Age group (m)	% Target-like	Pairwise contrasts
36–38	16.7%	Lower than all older groups (p < .0000001)
42–44	64.2%	Higher than 36–38 (p < .00000002); lower than 48–50 (p = .0006)
48–50	84.0%	Higher than 42–44 (p = .0006); not different from 54–56 (p = .089)
54–56	74.0%	Not different from 42–44 (p = .14) or 48–50 (p = .089)
60–62	91.7%	Higher than 42–44 (p = .0002) and 54–56 (p = .015)

A further pattern of interest concerns the relationship between target-like and English responses (see Table 5.12). Figure 5.14 plots the proportion of these two categories across age groups, showing a clear inverse trend: as accuracy in gender agreement increases across age groups, the proportion of English responses decreases. This trade-off suggests that developmental gains in target-like agreement forms are closely tied to a reduction in reliance on English lexical items. The implications of this pattern are discussed in Section 5.3.2.

Figure 5.14 *Proportion of Target-like and English Responses in the Gender Agreement Elicitation Task by Age Group*



Collectively, these findings indicate that accuracy in gender agreement increases steadily with age, with the sharpest gains between three and four years, followed by stabilisation and near-ceiling performance by around five years.

Across the naturalistic and elicited data analysed, the first component of Hypothesis 2B is confirmed, as accuracy in both gender and number agreement increases with age. For number agreement, however, this is supported only by the naturalistic data, as no elicitation task targeting number agreement was conducted. Regarding the second component of Hypothesis 2B, that gender agreement stabilises earlier than number agreement, descriptive patterns suggest that number agreement remains less frequent and more variable in spontaneous production. However, the regression analysis does not provide evidence for a statistically reliable difference in accuracy between gender and number agreement. As such, this component of the hypothesis is not confirmed and remains a question for future research. This concludes the examination of nominal agreement. The following section considers these findings in the wider context of overall grammatical development.

5.2.3. Hypothesis 2C: Developmental timing

This final hypothesis integrates the four main strands of investigation pursued in this thesis: the relationship between MLUm and chronological age, together with the noun bias over time (Hypothesis 1), and the acquisition of nominal inflection and agreement (Hypotheses 2A and 2B). In doing so, it closes the empirical loop and brings the work to its conclusion. Hypothesis 2C predicts that the progression of nominal inflection and agreement will coincide with MLUm levels that support multi-word utterances and with a lexical shift from a noun-dominant vocabulary to one with greater POS diversity.

To evaluate this hypothesis, data from across domains were combined. For MLUm, I used the final figures from both the longitudinal and cross-participant analyses. For the noun bias, or its attenuation, I aggregated the proportions of non-nominal items, that is, verbs, adjectives, and the *other* category, which includes function words, adverbs, auxiliaries, pseudoverbs, and active participles, while excluding nouns and child-language-specific categories. In both cases, analyses drew on the full dataset spanning 18–60 months.

For nominal inflection and agreement, however, the analysis relies on experimental data rather than spontaneous production, owing to the greater robustness of the former and the limited frequency of relevant tokens in the latter. Consequently, the age range is narrower (34–60 months) as elicitation tasks were not administered to younger children. Moreover, the focus within nominal inflection is on number marking since this is what the elicitation task targeted, while the analysis of nominal agreement is restricted to gender agreement, primarily involving nouns, adjectives, and demonstratives, with verbs and clitics excluded. For number inflection, the reported values represent the mean of the SP and BP results, which were elicited in equal numbers.

The results of the integrative analyses are shown in Figures 5.15 and 5.16. MLUm displays a steady upward trajectory, beginning at 1.3 in the 18–24-month range, consistent with single-word or transitional two-word productions. By 30–36 months, MLUm reaches 3.0, a level generally associated with productive multiword syntax. A steeper rise occurs by 42–48 months (4.56), followed by a plateau (4.68 at 56–60 months), suggesting that structural growth stabilises in the fourth year of life.

Lexical composition follows a complementary trajectory over this period. At 18 months, non-nominals constitute 27.5% of lexical items, reflecting the well-attested early noun bias. By 36 months, this balance shifts, with non-nominals accounting for 78.7%, and by 60 months this proportion stabilises at 83.8%. Thus, the transition from a noun-dominant vocabulary to one with greater lexical diversity coincides with the same developmental window in which MLUm crosses into the multiword range.

Figure 5.15 *Developmental Trajectories of MLUm, Non-nominals, Gender Agreement, and Number Inflection*

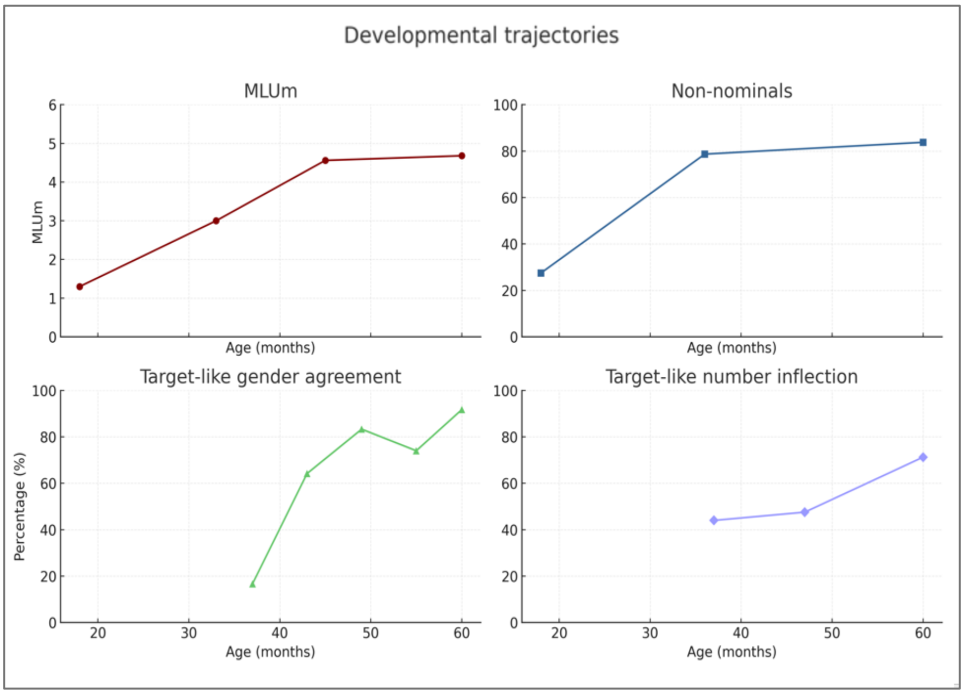
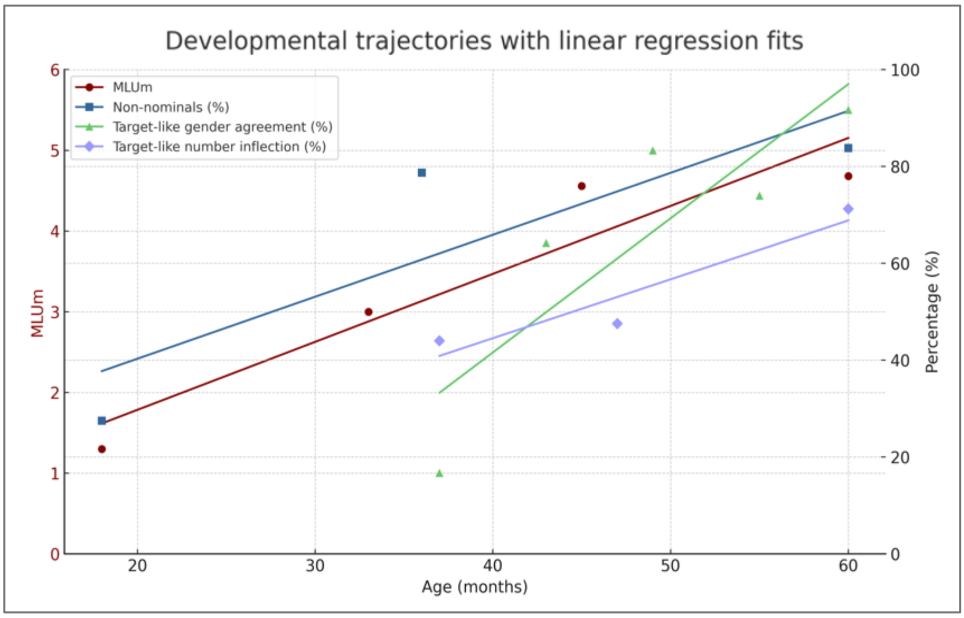


Figure 5.16 *Developmental Trajectories of MLUm, Non-nominals, Gender Agreement, and Number Inflection with Linear Regression Fits*



Nominal morphology shows a partially overlapping, but not identical, profile. Gender agreement accuracy is still emergent at 36–38 months (16.7%) but increases sharply between 42–44 months (64.2%) and 48–50 months (83.3%), with near-ceiling performance by 60 months (91.7%). Number inflection, by contrast, progresses more slowly: accuracy remains limited at 34–40 months (44.0%) and 44–50 months (47.6%), with substantial gains only evident from 54 months onward (71.3% at 54–60 months). Importantly, the aggregate figures presented here mask a divergence between the SP and BP: accuracy for SPs approaches ceiling levels, while BPs remain a locus of error. This lack of accuracy in the full plural system (both SP and BP) may contribute to the developmental course of number agreement, which depends on a stable base of inflectional morphology, although the present findings do not provide statistical evidence for a delayed stabilisation relative to gender agreement.

No additional inferential analyses were conducted for this integrative hypothesis. Since each domain, i.e., MLUm, POS category distribution, gender agreement, and number inflection, was examined statistically in the preceding sections, the aim here is to synthesise these established findings. Accordingly, the convergence across domains is illustrated descriptively and by means of regression-based trend lines (see Figure 5.16) which capture the developmental trajectories over time. It should also be noted that the dataset is limited to eight children, and the present synthesis should therefore be viewed as a springboard for further research.

These results indicate a convergence of structural, lexical, and morphosyntactic development in the 36–48-month window. Specifically, they highlight the transition to MLUm values supporting productive multiword syntax, the dissolution of the noun bias, and robust gender agreement cluster within this period, whereas number inflection lags slightly behind. This alignment supports Hypothesis 2C: the attainment of consistent, target-like nominal inflection and agreement emerges in concert with both increases in MLUm and a shift away from a noun-bias concentration. As such, the integrated patterns reported here serve as the empirical bridge to the discussion (Section 5.3) in which the implications of this convergence for theories of acquisition are considered.

Taken together, the findings from Hypotheses 2 trace the developmental trajectory of Maltese nominal morphology. Converging evidence from elicited and novel-word tasks indicates that morphological productivity consolidates around four to five years of age, coinciding with the transition to multiword syntax and a lexicon increasingly characterised by non-nominal items. In this way, the findings across domains reveal an integrated pattern of grammatical development: as children's utterances lengthen and diversify, their control over nominal morphology becomes both more systematic and more target-like. Despite the limited cohort and uneven form frequencies, the results point to a principled alignment between lexical, morphological, and structural growth. The next section considers the theoretical implications of this convergence for models of bilingual acquisition and morphological complexity.

5.3. Discussion

5.3.1. Hypothesis 2A: Nominal inflection

Hypothesis 2A predicts that children will acquire concatenative plural forms with greater accuracy and regularity than templatic formations. The findings from both spontaneous speech and two elicited production tasks (real-word and novel-word) provide consistent support for this claim. The clearest evidence comes from elicitation: by 54–60 months, children achieve 82.5% accuracy on SP targets compared with only 60% on BPs. Across tasks and naturalistic data, overgeneralisation of SP suffixes emerges as a dominant strategy, further underscoring the accessibility of concatenative morphology in early acquisition.

Developmental accuracy and cross-linguistic comparisons

Accuracy patterns across the developmental trajectory mirror those reported in earlier Maltese work and in other Semitic languages. Although Maltese-specific research in this domain is mostly limited to a couple of unpublished dissertations, the insights on MFLA in these studies are still of great relevance to this research. Coppini (2002) shows steadily increasing accuracy in three small cohorts (3-year-olds: 37.5%, 4-year-olds: 60.4%, and 5-year-olds: 82.3%), with remaining non-target-like productions in the oldest group largely restricted to BPs. Portelli (2005), in an MLUm-based study, similarly traces a gradual progression, with SPs appearing at MLUm 1.5–2.0, BPs at MLUm 2.5–3.0, and both categories generally mastered by MLUm 3.0–3.5. Case-study data also echo these trends. For instance, Vassallo (2006) documents the emergence of plural forms at 1;6, with productive though still inconsistent use by 2;6.

Cross-linguistically, the results align with those for Palestinian Arabic (Saiegh-Haddad et al., 2012), where SPs show high early accuracy, while BPs remain inaccurate, with accuracy reaching 75% by age 7–8. Comparable difficulties are attested in Jordanian Arabic (Mashaqba et al., 2020) and Hebrew (Ravid & Schiff, 2009), where stem-internal alternations prove significantly slower to acquire than suffixed ones. Plateau effects in SP development, such as those observed in the current work between 34–50 months, also parallel findings from Palestinian Arabic (Ravid & Farah, 1999). The cumulative evidence suggests that the transparency and regularity of concatenative morphology afford children a clear processing advantage, while stem-internal modifications entail protracted learning.

Strategies in elicited production

Analysis of children's responses in the elicited production tasks sheds light on responses produced when the target plural was not used. Four processes dominated: a) *repetition of the singular*, b) *overgeneralisation of SP suffixes*, c) *reliance on English plurals*, and d) *final vowel deletion*.

Repetition of the singular. Younger children frequently repeat the singular instead of producing a plural, particularly in the novel-word plural task (e.g., 51.2% at 44–50 months, dropping to 4.9% by 54–60 months). This mirrors earlier reports on acquisition in Maltese: Murray Curtis (1997) finds 34% singular use at age 3 compared to only 9% by age 5 in her study on the emergence of the plural, while Zammit (2008) documents similar patterns in younger groups. Such omissions are well attested in classic elicitation work (e.g., Berko, 1958; Ramscar et al., 2013) and are generally interpreted as reflecting incomplete mastery of plural formation. They may also indicate an early stage in which singular–plural contrasts have not yet been consistently differentiated (e.g., Clark & Nikitina, 2009).

Overgeneralisation of SP suffixes. Across all tasks, children extend a small set of suffixes (*-ijiet, -iet, -i, -s*) to one of two contexts: either attaching the wrong SP suffix to a given noun (e.g., *plattiet* for *platti* 'plates'; singular *platt*) or applying SP suffixation to nouns that take the BP (e.g., *mejdijiet* for *imwejjed* 'tables'; singular *mejda*). Previous research on Maltese preschool and school-aged children (Cachia, 1982; Murray Curtis, 1997; Vassallo, 2006; Zammit, 2008; Zammit, 2020) highlights the overgeneralisation of SP suffixes, with *-ijiet, -iet, -i* and *-s* identified as the most frequently overgeneralised forms.[79] Bartolo (2014) likewise reports a 6-year-old who rarely produced BPs, instead overgeneralising SP suffixes.

This recurring pattern may be linked to the fact that *-iet*, *-ijiet*, *-i* and *-s* are among the most common plural markers in the input, leading children to overgeneralise them in a manner comparable to English forms such as *childs* instead of *children*. However, research is yet to be conducted to determine the frequency and distribution of pluralisation patterns in spoken Maltese and in Maltese child language. Importantly, Nieder et al. (2020) show that the suffixes *-ijiet, -iet* and *-i* are also overgeneralised by Maltese adults, both with existing but infrequent nouns and with novel forms. This suggests that the productivity of these SP rules facilitates their extension beyond child language alone. In addition, the literature notes that

[79] A more detailed overview of the existing literature on the acquisition of the Maltese plural system, with extended examples, is provided in Vella (2025).

frequent nouns taking the BP, such as *kotba* 'books', pose less of a challenge than infrequent BP nouns, plausibly due to their high input frequency (Azzopardi-Alexander, 2016; Zammit, 2020). In the same vein, lower performance on some BP items may partly reflect the lower frequency of individual lexical items, indicating that frequency effects contribute to the observed asymmetry alongside morphological factors.

Saiegh-Haddad et al. (2012) report that in Palestinian Arabic, BP targets elicited non-target-like productions involving FSP morphology across all age groups, suggesting that FSP functions as the default pluralisation strategy. Notably, BP overgeneralisations appeared primarily among the oldest children, indicating that BP patterns consolidate only later in development. In the present study, BP overgeneralisations are rare but not confined to the oldest group, reflecting a similar but less sharply delineated trajectory. Mashaqba et al. (2020) likewise show that in Jordanian Arabic the FSP is acquired before BPs and overgeneralised to items requiring templatic morphology. They argue that children's choices are shaped by productivity and predictability rather than lexical familiarity, with overgeneralisation spanning morphological, phonological, and semantic dimensions. These findings reinforce the view that concatenative suffixation is both the earliest and most resilient plural strategy, while BPs emerge only with broader lexical growth and sensitivity to distributional patterns in the input.

In the novel-word plural task, while the cohort size is small, the contrast between the two age groups is nonetheless indicative. The absence of SP suffixation at 44–50 months implies that, at this stage, children are not yet applying a productive suffixal rule and instead rely on strategies such as singular repetition. By 54–60 months, however, SP suffixation emerges as the most common strategy, indicating that older children have begun to extend the suffixal rule beyond memorised forms to novel items. This developmental shift, though based on limited data, aligns with Hypothesis 2A in suggesting that concatenative morphology is acquired with greater accuracy and consistency than templatic morphology, a pattern often attributed to the greater transparency and regularity of affixation processes in acquisition (e.g., Clark, 1993; Dressler, 2005; Slobin, 1985).

Use of -*s* and English plurals. The prominence of the suffix -*s* in child productions is striking, given its restricted status in Maltese, where it appears with English loanwords.[80] In the elicitation task data, it occurs with just two SP nouns: *papras* for *papri* 'ducks' (singular *papra*) and *nannas* for *nanniet* 'grandmothers' (singular *nanna*). In contrast, children extend the suffix -*s* productively, and often rampantly, to BP nouns (e.g., *balluns* for *blalen* 'balls'; singular *ballun*, and *kelbs* for *klieb* 'dogs'; singular *kelb*). In many cases, they also replace Maltese nouns with their English equivalents altogether (e.g., *stars* for *stilel* 'stars'; singular *stilla*) (see also Cachia, 1982; Zammit, 2008; Zammit, 2020). Unlike many SP patterns in Maltese, where suffixation involves stem alternation or replacement (e.g., *fjura* – *fjuri* 'flower–flowers'), -*s* is added linearly without modifying the stem, making it morphophonologically simple and salient.

[80] Until relatively recently, comprehensive inventories of Maltese SP suffixes did not include -*s* among the attested suffixes (e.g., Schembri, 2006).

Such reliance on English should not be dismissed as confusion, since simultaneous bilinguals distinguish their languages from around age two (Genesee & Nicoladis, 2007). Rather, the pattern suggests that the suffix *-s* may be incorporated into children's developing morphological system as a general plural marker. Given its high frequency and transparency across the input, children readily extend it to contexts where a Maltese plural form is unavailable or not yet acquired. In spontaneous code-switched utterances, children often alternate seamlessly, e.g., *mhux il-/books/* 'not the books' (Lippu, 27 months), using the English plural while otherwise speaking Maltese. This behaviour clearly demonstrates how young bilinguals dynamically leverage their full linguistic repertoire to communicate successfully and reflects the pervasive presence of English in Malta and its significant, though variable, role in the linguistic environments of Maltese children (National Statistics Office, 2023a).

Final vowel deletion. Children also exploit morphophonological cues by dropping final vowels (*-a, -u*) in novel items (e.g., *lepa – lep*; Duminku, 50 months), thereby approximating frequent plural forms, namely suffixal *-i, -iet, -ijiet,* and *-s* forms as well as Types A–D templatic patterns, all of which do not end in *-a* or *-u*. In Maltese, final vowel loss sometimes accompanies plural formation (e.g., *sena – snin* 'years', *fenka – fniek* 'rabbits') and can also result in forms resembling collectives, which, as discussed earlier, can approximate to the plural semantically as they denote groups (e.g., *dudu – dud* 'worms'). While not target-like, such truncation demonstrates children's early awareness of frequent distributional patterns.[81]

[81] A subset of SP formations and less common BP types (Types E–G, J–K) do end in *-a* (e.g., *raħal – rħula* 'villages').

The role of collectives and lexical semantics

Certain non-target-like productions raise the possibility that collective nouns complicate plural acquisition. For instance, *tadamiet* ('tomatoes') and *brokk(o)li* ('broccoli')[82] were mis-produced with SP suffixes, despite requiring SP morphology in the contexts in which they appeared. Collective nouns in Maltese often involve complex paradigms, consisting of a three/four-way distinction, as discussed in Section 2.5.1.1. It is plausible that the additional morphological layer of the collective complicates children's mapping of plural forms, though this possibility is only speculative and requires empirical testing. While Abdilla and Spagnol (2016) investigate the use of collectives in Maltese-speaking adults, to date no comparable studies have been conducted with children.

In the case of *brokk(o)li*, the child produced *brokkolis*. One explanation is cross-linguistic influence: although embedded in a Maltese phrase, the noun may have been conceptualised in English (*broccoli*) and pluralised with the English suffix -*s*. The production could therefore reflect a reinterpretation of *broccoli* as a count rather than a mass noun, comparable to other child innovations in English such as *rices* or *spaghettis*.

[82] The status of *brokkli* as a collective is somewhat variable among speakers. While it may function as a collective noun for some, others prefer to treat it as a count noun (e.g., *żewġ brokkli* 'two broccolis') and find the derived plural *brokkliet* uncommon or unnatural. The example is therefore best considered a peripheral or marginal case.

Interpreting non-target-like productions

As noted earlier, many non-target-like productions should be viewed not only as developmental slips but as evidence of children's rule-based productivity. The overgeneralisations documented here resemble Berko's (1958) classic *Wug Test* results, demonstrating that children abstract morphological patterns beyond memorised forms. At the same time, some deviations may reflect creative or playful language use. Clark (2009) and Crystal (1998) emphasise children's lexical innovation and language play, and case studies confirm that invented forms often arise during pretend play (e.g., Nwokah & Graves, 2009).

The present data contain similar made-up forms, such as *socketti* /sɔkˈkɛttɪ/(used to refer to socks, blending the word *sock* and *kalzetti* /kɐlˈtsɛttɪ/ 'socks' – Anna, 36 months) or *sockina* /sɔkˈkiːnɐ/('small sock', employing the diminutive suffix -*ina*, commonly used in Maltese babytalk items such as *patatina* 'a small bum' – Anna, 36 months), and creative phrases like *scarf t'idejna* (literally 'scarf of our hands' – Lippu, 33 months) for 'gloves'. Such productions blur the line between error and invention, illustrating that acquisition is simultaneously a process of grammatical consolidation and imaginative experimentation.

Self-priming and task effects

Several non-target-like productions in the existing-word plural task suggest that children may have been primed by their own previous responses. For instance, a child who correctly pluralised *ġakketta* ('jacket') and *farfett* ('butterfly') with Type

A BP (*ġkieket, friefet*) subsequently produced *ħbiebel* (rather than target *ħbula*) for *ħabel* ('rope'), possibly extending the earlier pattern. Similarly, another child's repeated use of -*ijiet* with 11 of 20 items, only two of which required this suffix, points to self-priming effects. Although this study does not explicitly address priming, the possibility emphasises how task design and item ordering may shape children's morphological choices and constitutes an interesting avenue for future research.

Broader typological parallels

The asymmetry between concatenative and non-concatenative morphology resonates beyond Semitic. In German, for example, children struggle particularly with umlaut plurals that require stem-internal vowel change without an overt suffix such as *Apfel – Äpfel* 'apple – apples', *Mutter – Mütter* 'mother – mothers'. Here too, overgeneralisation of suffixal patterns is common, and mastery of the full system typically extends beyond age six (e.g., Kauschke et al., 2011). Thus, the Maltese data align with a broader typological pattern in which regular, segmentable concatenative operations are favoured in early morphology, whereas stem-internal or non-transparent alternations are typically acquired later.

Conclusion

All the evidence, across tasks and data sources, supports Hypothesis 2A since the children in this study acquire SPs earlier, more accurately, and more productively than they do BPs. Their reliance on overgeneralisation of SP suffixes demonstrates that concatenative morphology is accessible and extendable even when not fully mastered. In contrast, BPs, with their structural complexity and limited predictability, develop more slowly and trigger workaround strategies, mirroring findings from other languages.

The frequent use of English plurals suggests that bilingual children may have incorporated the more transparent suffix *-s* into their developing morphological system. Such behaviour should be seen as evidence of bilingual resourcefulness, shaped both by internal cognitive economy and external sociolinguistic pressures. The occasional influence of collectives and the possibility of priming effects highlight additional variables worthy of future investigation. Finally, as suggested earlier, the creativity evident in children's playful coinages serves as a reminder that acquisition is not a linear accumulation of rules but a dynamic process of innovation and flexible use of the linguistic repertoire.

Overall, the data confirm that concatenative processes are acquired more rapidly than templatic ones. The regularity and productivity of SP suffixation render it both an early-emerging and an overgeneralised default, while templatic morphology remains challenging well into the school years. This asymmetry situates Maltese within the broader typological pattern whereby transparency and segmentability accelerate morphological development, a finding of relevance not only for the acquisition of Semitic languages but also for cross-linguistic theories of morphological learnability.

5.3.2. Hypothesis 2B: Nominal agreement

Hypothesis 2B predicts that accuracy in gender and number agreement increases with age, with gender stabilising first. The spontaneous and elicited production data analysed provide support for the first component of this hypothesis, while the second component is not confirmed and remains a possible avenue for future research. Across both datasets, accuracy improves developmentally. Elicited production confirms steady gains in gender, with the sharpest increases between 36–44 months, followed by near-ceiling performance at 60 months.

Developmental Patterns of Gender and Number

Descriptive patterns in the naturalistic data suggest earlier consolidation of gender agreement relative to number agreement, and these patterns accord with findings reported for Semitic languages and beyond. In Hebrew, gender marking is acquired earlier and more consistently than number, reflecting its frequency and relative transparency (Berman, 1985; Ravid & Schiff, 2012). Similar developmental sequencing is reported for Saudi Arabic (Moawad, 2006) as well as for Spanish (Socarrás, 2011), among others. In Maltese, Coppini's (2002) work on nominal agreement in 3–5-year-olds likewise finds that plural targets elicit more non-target-like productions than feminine targets, suggesting earlier mastery of gender. She argues that number, unlike gender, is not lexically arbitrary but conceptually grounded in plurality, and that non-target-like productions with plural forms are attributable not to semantic ambiguity but to morphophonological complexity. More generally, monolingual and bilingual acquisition research shows that features characterised by regularity, transparency, and salience are mastered earlier than those with irregular or opaque mappings (e.g., Demuth & Weschler, 2012; Paradis, 2010). Bilingual studies further confirm that acquisition trajectories are strongly language-specific, with distinct sequencing and non-target-like patterns (e.g., Thordardottir, 2015).

The current study tests variability in number agreement in spontaneous speech, where the scarcity of obligatory contexts limits systematic analysis. Nevertheless, the available data suggest challenges with number marking. Under usage-based or emergentist models, acquisition proceeds through generalisation over frequent patterns and exploitation of multiple cues (Bybee, 2007; Tomasello, 2003). In Maltese, as in other Semitic languages, both gender and number are expressed across multiple elements (e.g., on determiners, adjectives, and verbs). However, gender agreement marking tends to be more regular and segmentable, whereas plural morphology exhibits greater diversity. This asymmetry may facilitate earlier stabilisation of gender. Input frequency is not tested here, and therefore, further work is required to examine frequency effects in Maltese caregiver speech.

Evidence from Spontaneous and Elicited Data

This section discusses the spontaneous and elicited data in relation to Hypothesis 2B, illustrating how children's early innovations and repairs shed light on the development of gender and number. In (1), a 27-month-old first produces a masculine demonstrative with a feminine noun but immediately self-corrects, producing the feminine form in the same breath. Such online repair demonstrates not only awareness of the mismatch but also access to the relevant agreement rule, suggesting that gender is already being monitored in real time at this age. The noun *mamà* 'mum' is among the most frequent words in CDS and refers to a salient animate entity whose natural gender is unambiguous. Both its frequency and conceptual transparency likely support the child's rapid detection and correction of the agreement error, making this an especially clear case of early gender sensitivity.

(1) *Dal-mamà... dil-mamà u da' t-tifel.*
this.SGM-DEF-mum.SGF this.SGF-DEF-mum.SGF and this.SGM DEF-boy.SGM
'This is the mum and this is the boy.' (Lippu, 27 months)

In (2), a 36-month-old produces a mismatch between demonstrative and participle (adjective-like). This reflects difficulty in coordinating agreement across multiple targets within a single utterance, a pattern observed cross-linguistically when features must be realised redundantly across several morphosyntactic domains.

(2) *Din[a] mimli kokò?*
this.SGF full.SGM poop.SGM
'Is this one full of poop?' (Toninu, 36 months)

In (3), a 60-month-old alternates between masculine and feminine demonstratives when referring to a toy excavator. The lexical item *digger* functions here as a code-switched or loan element. Initially marked feminine by the child, it is subsequently shifted to masculine, consistent with both phonological form (consonant-final, defaulting to masculine) and the investigator's input.

(3)	CHI: *ma da' x'jgħidulu?*	'mum, what's this one called?'
	MOT: *x'inhu ħi?*	'what's that?'
	INV: *xi digger* ***<u>dak</u>*** *x'inh****<u>u</u>*** *għid?*	'is it a digger? what would you say?'
	CHI: *iwa digger ta*	'yes, it's a digger'
	CHI: *ee* ***<u>dik</u>*** *digger*	'yeah, that's a digger'
	INV: ***<u>j</u>****agħmel ħafna ħoss hux?*	'it makes a lot of noise, right?'
	CHI: *mm* ***<u>dak</u>*** *digger*	'mm, that's a digger'

(Reno, 60 months)

This example illustrates how children balance multiple cues (e.g., morphological endings, conceptual features, and adult modelling) in the assignment of grammatical gender. The fact that the child ultimately aligns with masculine agreement suggests flexibility in applying analogical strategies, especially for low-frequency or novel items. The capacity for self-correction and cue-shifting reveals gender as an emerging but productive domain in the grammar. This is consistent with Farrugia's (2018) findings from a novel-word task with adults, which show that word-final segments strongly influence gender assignment in Maltese, a tendency also noted by Stolz (2009). In such cases, formal cues may override conceptual ones, and children adjust their productions accordingly.

Turning to elicited production, patterns of lexical avoidance and substitution emerge. For instance, a number of children resist using *oħxon/ħoxna* 'fat' with animals in the elicitation task, with parents explaining that this reluctance may stem from reprimands when the adjective had been used for human referents (4). This underscores the role of pragmatic and sociolinguistic filters in acquisition.

(4) MOT: *għax dik ngħidilha ma toqgħodx tgħid hekk fuq in-nies ħalli ma tweġġgħalhomx il-/feelings/*

'because I tell her not to talk about people like that, so she doesn't hurt their feelings' (Kelina, 41 months)

The adjective pair *twil/qasir* ('tall/short') is likewise never produced. Instead, children consistently employ *kbir/żgħir* ('big/small'). This pattern reflects the developmental sequence whereby general size adjectives, being conceptually broader and more perceptually salient, are acquired prior to dimension-specific terms (e.g., Ferry et al., 2025). Although the inclusion of *oħxon/irqiq* and *twil/qasir* constitutes a minor oversight in task design, it nevertheless yields valuable insights. As illustrated in (5–7), when the target adjectives is not produced, children used descriptive paraphrases or morphological innovations, demonstrating communicative flexibility and creativity. Example (5) corresponds to the *twil/qasir* pair, and examples (6) and (7) to *oħxon/irqiq*.

(5) *dan il-pa[j]ent u dan iċ-child* (Toninu, 56 months)
'this is the parent and this is the child'

(6) *għax dik b'għajnejha miftuħin u dik magħluqin* (Kelina, 48 months)
'because that one has her eyes open and the other one has them closed'

(7) *naqra żaqqieq* (Reno, 48 months)
'a bit of a glutton'

Finally, the adjective pair *nadif/maħmuġ* 'clean/dirty' elicited particularly striking results. *Nadif* was used by only three children, two of whom misinflected it (*nodfa* 'clean.PL' was used with singular referents; see Table 5.12). These non-target-like productions indicate difficulty with the lexical form itself rather than the referent. Most children relied on English *clean*, likely due to its high frequency in educational settings such as childcare routines and songs, or produced innovative Maltese forms (see Table 5.14). Of particular note is the production of *imħammeġ/imħammġa* for standard *maħmuġ/maħmuġin*, seemingly derived from the derived second form causative verb *ħammeġ* 'to make dirty' rather than the first form base **ħameġ*, which does not occur. These forms possibly reveal reliance on a highly productive derivational schema (Verb → Participle → Adjective). Although non-standard, they highlight how children exploit transparent morphological patterns in building their lexicon, resulting in overgeneralisation (e.g., *kisser* – *imkisser* 'he broke – broken'). Table 5.15 summarises child innovations in adjective production found in the data.

Table 5.14 *Alternatives Provided for the nadif/maħmuġ Pair*

<table>
<tr><th>Target Adjective</th><th colspan="5">Alternative Child Responses</th></tr>
<tr><td>nadif 'clean'</td><td>/clean/</td><td>mhux *imħammeġ
'not dirty'</td><td>/nice/</td><td>sabiħ
'nice'</td><td>għadu kif inħasel
'he just had a shower'</td></tr>
<tr><td rowspan="2">maħmuġ 'dirty'</td><td>/dirty/</td><td>*imħammeġ
'dirty'</td><td>mhux /clean/
'not clean'</td><td colspan="2">mimmi
'pain/injury' – Maltese babytalk item</td></tr>
<tr><td>/messy/</td><td>/yucky/</td><td>dan għandu l-/paint/
'he has paint'</td><td colspan="2">dan il-/pig/ iħobb jagħmel hekk ġol-/muddy puddles/
'this pig likes to do this in muddy puddles' + jumping and hand-waiving gesture</td></tr>
</table>

Table 5.15 *Child Innovations and Avoidance Strategies in Adjective Production*

Target Adjective Pair	Child Responses	Linguistic Process
nadif/maħmuġ 'clean/dirty'	*/clean/*, *imħammeġ*, *mhux imħammeġ; alternative words*; descriptive multi-word expressions	Code-switching; overgeneralisation of participial derivation; avoidance
twil/qasir 'tall/short'	*kbir, żgħir* 'big, small'; descriptive multi-word expressions	Conceptual substitution; preference for general size adjectives
oħxon/ħoxna 'fat'	Avoidance; explicit metalinguistic rejection	Sociolinguistic awareness; taboo avoidance

Together, the spontaneous and elicited data show that gender agreement, though not fully stable in the youngest speakers, is already available for monitoring, repair, and extension. Children demonstrate not only early awareness of agreement constraints but also flexibility in navigating morphophonological complexity, lexical gaps, and sociolinguistic norms.

Code-Switching and Agreement by Analogy

CS data show that English nouns serving as controllers are typically assigned gender by analogy onto the Maltese equivalent, with only one non-target-like production (7.1%). This reflects cross-linguistic patterns whereby bilingual children extend agreement categories from one language to invariant forms in another (e.g., Kupisch, 2007). The children's reliance on translation equivalents indicates use of abstract lexical knowledge of Maltese gender categories rather than surface-form strategies. Such cases demonstrate flexible, productive deployment of the Maltese system across languages. This pattern is also consistent with previous observations on Maltese gender agreement. For instance, Desira (2002) reports the production of *dan /plane/* 'this.SGM is a plane' (*ajruplan* 'plane.SGM') by the 2-year-old boy in her case study, suggesting that children extend Maltese agreement rules to English lexical items.

For number, English plurals are directly mapped to Maltese plural targets, with relatively high accuracy. Interestingly, an inverse relationship was observed between reliance on English lexical items and target-like Maltese agreement, that is, as children's accuracy in Maltese increased, use of English decreased. This suggests that English serves as a communicative scaffold in contexts of morphological uncertainty, particularly where Maltese plural formation poses difficulties. These findings resonate with bilingual acquisition research highlighting the interaction of dominance, structural transparency, and communicative strategies (e.g., Genesee & Nicoladis, 2007; Meisel, 2011).

Theoretical Implications

The descriptive asymmetry observed between gender and number agreement is compatible with gradualist accounts of morphosyntactic development. From a usage-based perspective, gender may become entrenched earlier due to its frequency and relative morphological transparency, whereas plural formation remains more vulnerable to opacity and irregularity (e.g., MacWhinney, 2005; Ravid, 2019). From a generative standpoint, the findings are also consistent with proposals that feature-checking mechanisms emerge in a staged manner, although the present results do not provide statistical evidence that gender stabilises before number (e.g., Radford, 1990; Snyder, 2021).

More broadly, the Maltese data align with cross-linguistic observations that agreement features do not necessarily become productive simultaneously. Rather, they develop incrementally and are shaped by their morphological and distributional properties.

Conclusion

The findings support the first component of Hypothesis 2B since gender and number agreement both improve as children grow older. The second component of the hypothesis, that gender stabilises earlier than number, is not confirmed and requires further investigation. Children's use of analogy and CS highlights agreement as a staged process within wider grammatical development. This incremental consolidation sets the stage for examining how the timing of target-like agreement aligns with broader developmental milestones such as MLUm growth and the lexical shift toward more varied POS category proportions.

5.3.3. Hypothesis 2C: Developmental timing

The integrative analysis of MLUm, POS distribution, and nominal morphology provides an empirical test of Hypothesis 2C. The results show that children's attainment of target-like gender agreement coincides with the point at which MLUm crosses into the multiword stage, and the lexicon undergoes a restructuring away from noun dominance. Number inflection, by contrast, develops more slowly and remains less stable across the observed age span. This profile highlights that morphosyntactic acquisition must be understood in relation to broader developmental trajectories of structure and lexicon.

Domain convergence and theoretical implications

The synchronisation, albeit imperfect, across MLUm, non-nominal proportions, and gender agreement between 36–48 months parallels a widely accepted view in CLA: morphosyntax emerges in concert with lexical and structural growth (e.g., Bybee, 1985; Ravid, 2019). This pattern echoes usage-based and emergentist accounts, which stress the coupling of morphosyntactic development with lexical expansion and utterance complexity (e.g., MacWhinney, 2005; Tomasello, 2003). The sharp increase in non-nominal lexical categories around 36 months not only signals the erosion of the early noun bias (e.g., Bornstein et al., 2004; Gentner, 1982) but also provides the lexical classes necessary to support agreement processes. In this sense, lexical restructuring may serve as a prerequisite for subsequent morphological breakthroughs.

The plateau of MLUm around 4.6–4.7 resonates with earlier findings for other languages (e.g., Brown, 1973; de Villiers & de Villiers, 1978) which show that, once children routinely produce multiword constructions, further increases in utterance length contribute less to the acquisition of new morphology. While MLUm continues to be informative as a measure for language development in older children (e.g., Portratz et al., 2022), lexical and morphological diversification become increasingly important predictors of morphosyntactic competence (e.g., Frizelle et al., 2018). These results thus provide further support for lexicon–syntax coupling, that is, agreement is not acquired independently of structural and lexical growth but emerges when utterances reach sufficient complexity and POS proportions shift (e.g., Pine et al., 2013; Valian, 1986).

From a generative perspective, one might interpret this synchronisation as signalling the acquisition of feature-checking mechanisms (e.g., gender, number) once multiword syntax is in place. The fact that gender appears to stabilise earlier than number in descriptive terms suggests that features do not become fully productive simultaneously but emerge in a staged fashion. This developmental lag aligns with Snyder's (2021) P&P characterisation of acquisition as a deterministic but additive process, in which grammatical changes are interconnected yet may unfold at different rates. The findings also echo reports from Semitic languages, where gender agreement tends to stabilise before number agreement, with plural morphology remaining a locus of difficulty well into the preschool years (e.g., Ravid & Schiff, 2009). Similar sequencing effects have been observed in bilingual contexts, where morphosyntactic features associated with higher frequency and greater transparency consolidate earlier than those with irregular mappings (e.g., Paradis, 2010).

Morphological complexity and asynchrony

Morphological complexity compounds this asymmetry between nominal gender and number agreement. In the present data, aggregate measures of number inflection obscure the relative ease of SPs compared to BPs. The persistence of non-target forms in the latter may delay the stabilisation of number marking, with potential effects for number agreement. This pattern mirrors broader cross-linguistic evidence which reveals irregular or non-transparent plural morphology as a developmental bottleneck in numerous languages such as Romansh Tuatschin and Palestinian Arabic, among many others (e.g., Mazara et al., 2024; Saiegh-Haddad et al., 2012). In addition, these findings argue against a monolithic "feature activation" model (e.g., Radford, 1990) and instead support gradualist accounts in which morphosyntactic features become productive at different points, depending on their morphological and distributional profiles.

Conclusion

In sum, the findings from Hypothesis 2C demonstrate that the attainment of productive nominal agreement is not an isolated achievement but one that emerges at the intersection of structural growth, lexical diversification, and morphological consolidation. The descriptive asymmetry between gender and number underscores the role of morphological complexity and irregularity, while the broader convergence across domains provides clear evidence that acquisition proceeds through dynamic interactions between the lexicon, morphology, and syntax. These results not only align with patterns observed in other Semitic and bilingual contexts but also set the stage for future work on how cross-domain interdependencies shape the acquisition trajectory.

5.3.4. Concluding remarks

Across the discussion, a developmental pathway emerges of how TD, MD bilingual children assemble nominal inflection and agreement, with productivity first consolidating in transparent, segmentable patterns and more opaque templates trailing behind. The asymmetry between concatenative and templatic pluralisation, coupled with the earlier emergence and consolidation of gender relative to number in descriptive patterns, points to a layered acquisition path in which form–meaning mappings that are frequent, regular, and perceptually tractable provide the scaffolding for later generalisations.

The evidence from spontaneous and elicited data suggests that non-target-like productions, rather than reflecting failure, highlight children's developing morphological systems and strategies for managing complexity. CS and overgeneralisation emerge as productive means of negotiation within a bilingual repertoire, revealing adaptive use of linguistic resources rather than interference.

Theoretically, the Maltese data contribute to refining cross-linguistic claims about learnability and feature timing. They support accounts in which the distributional transparency of exponents and redundancy across targets accelerate agreement, while irregularity and stem alternations delay it. They also illustrate how bilingual repertoires are exploited as economical problem-solving resources: CS by analogy and selective reliance on the English plural *-s* are principled adaptations to morphological complexity. Taken together, these patterns substantiate the developmental trajectory outlined in Hypothesis 2, confirming that the acquisition of gender and number features proceeds in a systematic, growth-linked manner. However, while descriptive evidence points to earlier consolidation of gender agreement, the present findings do not provide statistical confirmation that gender stabilises before number. While the scope of the dataset and task design impose limitations, the converging results highlight the interdependence of lexical, morphological, and structural growth.

6. Conclusion

This study has sought to chart the emergence and development of morphosyntax in TD, MD bilingual children aged 18–60 months, tracing their progression from use of single-word utterances to the productive use of multiword syntax, with particular attention to the nominal domain. Drawing on longitudinal naturalistic data from eight children, complemented by elicited and novel-word production tasks with the same participants, it has offered an account of early grammatical growth within a bilingual context.

This concluding chapter begins by synthesising the principal findings in relation to the study's hypotheses (Section 6.1) and then discusses their broader theoretical and methodological implications (Section 6.2). It proceeds to outline the study's main contributions (Section 6.3) and concludes with a reflection on its limitations and suggestions for future research (Section 6.4).

6.1. Summary of findings

This study investigated the relationship between structural growth (indexed by MLUm), POS category distribution, and the acquisition of nominal morphosyntax, encompassing inflection and agreement, and their developmental timing. Collectively, these results reveal a coherent pathway in which structural, lexical, and morphosyntactic systems progressively converge between 18–60 months, portraying development of nominal systems as gradual and interdependent. Table 6.1 offers an integrated overview of the principal findings and their theoretical interpretations, summarising for each hypothesis the analytical focus, key empirical outcomes, and their implications for understanding language development in a TD, MD context.

MLUm increased with age, marking steady structural expansion, while the early noun bias gave way to a more functionally diverse lexicon. Nominal inflectional accuracy followed a clear progression in which transparent concatenative forms were acquired earlier and more productively than templatic ones, and agreement stabilised earlier in gender than in number. The timing of these achievements coincided with the transition to multiword utterances and the attenuation of noun dominance, indicating close coupling between structural, lexical and morphosyntactic domains. Overall, the findings confirm the predicted developmental patterns: measures of MLUm and POS distribution index grammatical growth, and the acquisition of nominal inflection and agreement unfolds in tandem with, though not identical to, broader structural advances.

Together, these outcomes highlight a developmental continuum driven primarily by morphosyntactic transparency while recognising that other factors not examined in this study, such as input frequency, may also contribute to the course of development. The longitudinal design, spanning 18–60 months, made it possible to trace not only the order in which these systems emerge but also their evolving interdependence. The following sections draw these strands together, discussing the broader theoretical implications for models of morphosyntactic learnability, bilingual acquisition, and early grammatical development.

Table 6.1 *Summary of Main Findings and Theoretical Interpretations*

Hypothesis	Focus	Empirical Findings	Interpretation and Theoretical Implications
1A	MLUm developmental trajectory and chronological age	MLUm increased between 18–60 months and showed a strong positive association with age ($r = .91$, $p < .001$), with development comparable to Brown (1973), Portelli (2005), and Tallas-Mahajna and Dromi (2023)'s findings.	Supports MLUm as a robust index of morphosyntactic development in Maltese and situates the findings within both Maltese-specific and cross-linguistic research.
1B	Noun bias over time	Noun dominance at 18 months (84%) decreased steadily; verbs and function words increased and predominated by 60 months.	Demonstrates POS category restructuring; supports interaction between lexical expansion and morphosyntactic elaboration.
1C	MLUm and noun proportions	Strong negative correlation between MLUm and noun proportions ($r = -.81$, $p < .001$).	As utterances lengthen, POS diversity increases; grammatical complexity emerges alongside a reduced noun bias.
2A	Nominal inflection	SPs are acquired earlier and with greater accuracy than BPs. Overgeneralisations were systematic.	Morphological transparency and regularity facilitate acquisition; supports cross-linguistic hierarchy of morphosyntactic learnability.
2B	Nominal agreement	Gender agreement shows earlier emergence in descriptive patterns, while number agreement remains more variable and less frequent.	Gender features may be acquired earlier due to higher morphological predictability, though not statistically confirmed.
2C	Developmental timing	Productive morphology coincided with MLUm exceeding ~3.0 and with attenuation of the noun bias.	Indicates convergence of structural, lexical and morphosyntactic growth; supports domain-interaction models of morphosyntactic acquisition.

6.1.1. Structural and lexical growth

Across both longitudinal and cross-participant analyses, MLUm was strongly and positively correlated with age and aligned with established MLUm models (e.g., Brown, 1973; Portelli, 2005; Tallas-Mahajna & Dromi, 2023), particularly within the 30–48-month period, therefore confirming Hypothesis 1A. Children's utterances lengthened steadily from the single-word stage to complex multiword constructions. The longitudinal perspective further captured dynamic individual paths, revealing that morphosyntactic growth cannot be reduced to static group means.

Analyses of POS categories revealed the expected early noun bias, which diminished as verbs and function words gained prominence (Hypothesis 1B). The robust negative correlation between MLUm and noun proportion indicates that as utterances lengthen, POS repertoires diversify, highlighting the interdependence between POS diversification and grammatical elaboration (Hypothesis 1C). These findings align with usage-based accounts, in which shifting proportions of POS categories and an expanding lexicon scaffold the emergence of morphosyntactic organisation.

Finally, variation in CS across children underscored the dynamic nature of the bilingual environment, with a tendency for higher CS rates among girls in the developmental group. Sex-based differences in MLUm likewise remained inconclusive, reflecting individual variability and the small sample size. While exploratory, these observations suggest that alternating between Maltese and

English constitutes a typical aspect of the developmental ecology rather than a deviation from it.

6.1.2. Nominal inflection and agreement

Analysis of nominal inflection confirmed that concatenative plural forms were acquired earlier and with greater stability than templatic forms (Hypothesis 2A). Children's reliance on suffixal strategies, often overgeneralised beyond target contexts, illustrates both the productivity and the limits of developing morphological rules. In contrast, BPs remained error-prone and infrequent, reflecting the higher processing demands of non-transparent morphology.

Non-target-like productions revealed active experimentation: overgeneralisation, vowel deletion, and creative coinages indicate rule formation rather than random error, with some coinages also reflecting functional descriptions used to compensate for lexical gaps. English plural morphology, particularly the productive *-s* suffix, served as an accessible alternative when Maltese morphology proved complex, showing that bilingual children recruit both systems adaptively.

In the domain of agreement, gender appeared to emerge earlier and reach stable, near-ceiling accuracy by around five years in descriptive patterns, whereas number remained more variable and less frequent (Hypothesis 2B). However, this asymmetry was not statistically confirmed and remains a question for future research. CS data revealed that children alternated between Maltese and English within nominal phrases, providing additional insight into how morphological patterns may generalise across the two languages. Although the evidence is limited, there are indications that some English nouns might have been assigned Maltese gender and number features by analogy, which could tentatively suggest an emerging level of abstraction rather than simple surface imitation.

Integrative analyses further demonstrated that the emergence of stable nominal morphosyntax coincided with the developmental window in which MLUm reached the multiword stage and the lexicon diversified beyond noun dominance (Hypothesis 2C). Morphosyntactic productivity, therefore, arises once utterances and the lexicon become sufficiently complex to support it.

6.1.3. Interconnected development

Taken together, these findings depict a coherent developmental progression. Between 18–60 months, MLUm growth, POS diversification, and consolidation of nominal morphosyntax advanced in concert, though not in lockstep. As utterances lengthen and POS categories diversify, morphological rules become productive; as morphology stabilises, it feeds back into syntactic organisation. The longitudinal scope of the dataset captures this interdependence, charting the shift from isolated lexical items to an integrated grammatical system. The MD bilingual context further underscores that such growth unfolds within a dynamic interplay of two linguistic systems, with children drawing flexibly on both systems to communicate effectively. The developmental synthesis outlined above sets the stage for evaluating the study's wider theoretical and methodological contributions.

6.2. Theoretical and methodological implications

The findings of this study point to a gradual and interdependent pathway of grammatical development, in which morphosyntactic productivity arises through the interplay of structural, lexical, and morphosyntactic growth. From one perspective, this developmental progression accords with usage-based models that view grammar as emerging from experience, pattern abstraction, and domain-general learning processes (e.g., Bybee, 2010). The incremental acquisition of transparent, concatenative forms before less regular, non-concatenative forms is consistent with accounts that emphasise the role of systematicity and cognitive accessibility in shaping early morphosyntax.

However, the patterns observed here are not exclusively compatible with usage-based explanations. The relatively early and consistent mastery of certain morphosyntactic distinctions, such as gender agreement, can also be interpreted within frameworks that posit an innate grammatical architecture guiding acquisition (e.g., Snyder, 2021). Moreover, because this study did not examine input characteristics, which are central to usage-based approaches, its explanatory reach remains circumscribed. The results therefore invite, rather than settle, broader theoretical claims. As discussed in the following section, future research incorporating measures of input frequency and distribution will be essential for adjudicating between competing explanations of bilingual morphosyntactic development.

Extending this perspective to the Maltese context more specifically, the data provide a clear illustration of how concatenative morphology, being segmentable and regular, serves as a developmental anchor, whereas templatic patterns require longer exposure and greater abstraction. Typologically, these findings align Maltese with other Semitic and non-Semitic languages that exhibit protracted acquisition of stem-internal alternations, pointing to a broader principle of morphosyntactic learnability (e.g., Mashaqba et al., 2020 – Jordanian Arabic). Within the bilingual environment, the children's flexible deployment of English morphology alongside Maltese forms demonstrates that cross-linguistic interaction is adaptive rather than disruptive. This aligns with empirical evidence that bilingual learners draw on both linguistic systems in development, modulated by exposure and task demands (e.g., Castilla-Earls et al., 2023; Paradis, 2010). These outcomes suggest that bilingual acquisition can involve convergence across systems while maintaining language-specific distinctions, highlighting the dynamic nature of early grammatical development.

Methodologically, the integration of naturalistic corpus data with elicited real-word and novel-word tasks proved particularly informative. The naturalistic corpus captured spontaneous morphosyntactic use and CS in free-play contexts, while the elicited data revealed underlying productivity not always visible in spontaneous production. Tracking the same children over two years provided developmental continuity, allowing fine-grained observation of how structural, lexical, and morphosyntactic systems become increasingly interdependent. Such triangulation addresses a long-standing methodological limitation in the field, which often relies on single-method or cross-sectional designs (e.g., Conboy & Thal, 2006).

Overall, these insights underscore both the theoretical and the methodological implications of the study. Depending on the framework adopted, the results can be interpreted as evidence of either emergent, usage-based learning or the early activation of abstract structural knowledge, while, at the same time, not excluding other alternative frameworks. This framework-neutral approach, coupled with the longitudinal, mixed-method design, provides a flexible means of capturing developmental continuity in bilingual morphosyntax. Nevertheless, future research integrating measures of input and processing will be essential to substantiate and extend these conclusions.

6.3. Contributions

This study offers significant contributions at empirical, methodological, and theoretical levels, alongside practical implications that extend beyond the local and international linguistic spheres.

Empirically, it provides the first in-depth longitudinal account spanning two years and focusing on a specific area of Maltese morphosyntactic acquisition as an L1. By documenting both spontaneous and elicited child language within a MD bilingual environment, it establishes an invaluable empirical foundation for understanding the emergence and progression of Maltese morphosyntax, particularly in the nominal domain. In doing so, it enriches an underrepresented area of research and offers a crucial reference point for future studies on typical Maltese language development.

Methodologically, the study demonstrates the value of integrating corpus-based and experimental approaches to capture both the accuracy and the productivity of developing morphosyntax. This mixed-method, longitudinal design serves as a model for future research aiming to refine methodological practices in bilingual acquisition studies, illustrating how quantitative and qualitative data can be combined to produce a more comprehensive account of linguistic development.

In addition, a central contribution of this thesis lies in the compilation of a new corpus comprising over 30 hours of Maltese child language data collected through spontaneous speech and free play interactions between the child and the investigator. This corpus, manually transcribed and systematically annotated, represents a substantial methodological and empirical resource for the present analyses and for stimulating future research on MFLA. As Ambridge and Rowland (2013, p. 150) note, “naturalistic speech analysis can be very cost-effective in the long term,” since once such data have been collected and transcribed, they can be used repeatedly to investigate a range of phenomena in CLA.

Theoretically, the research advances cross-linguistic debates on morphosyntactic learnability and the mechanisms underlying language acquisition by incorporating Maltese into these discussions and enhancing the much-needed typological diversity in the field (Allen, 2024; Kidd & Garcia, 2022). It shows that morphosyntactic transparency predicts acquisition in the Maltese hybrid system, thereby contributing new evidence to debates on typology and learnability. Furthermore, the findings reframe bilingualism as a resource rather than a constraint, demonstrating that cross-linguistic transfer can facilitate rather than hinder morphosyntactic development. In so doing, the study contributes to ongoing theoretical discussions surrounding innateness and emergentist perspectives on language acquisition and processing.

Beyond these core contributions, the research carries important practical implications. For the Maltese linguistic community, it provides a set of normative developmental data on nominal morphosyntactic acquisition, addressing the long-recognised gap that compels clinicians to rely on English-based developmental norms (Gatt, 2010). The findings will thus serve as an indispensable resource for speech and language pathologists in identifying, assessing, and managing language difficulties in MD children, as well as supporting the rehabilitation of linguistic skills in Maltese-speaking adults with acquired language disorders.

The study's insights are equally relevant for educational practice. By outlining developmental patterns of nominal morphosyntactic acquisition in relation to broader structural and lexical growth, the research can inform the design of pedagogical materials and curricula for the early instruction of Maltese as both an L1 and L2. This information will be valuable not only to educators, but also to Maltese parents, authors, and language resource developers seeking to create age-appropriate materials for young learners.

Finally, the study's theoretical openness and its integration of longitudinal, corpus-based, and experimental methodologies establish a solid foundation for future research. It is hoped that these findings will act as a springboard for further investigations into MFLA, thereby extending our broader understanding of bilingual CLA and morphosyntactic development.

6.4. Limitations and directions for future research

While this study has provided valuable insights into the development of Maltese morphosyntax in early bilingual acquisition, a number of limitations must be acknowledged. These, in turn, highlight promising directions for future research.

6.4.1. Limitations

Sample size and scope of analysis

The study was based on a small cohort of eight children, that inevitably constrains the generalisability of the findings. Although the longitudinal design allowed for detailed developmental tracking, a larger sample would enable more systematic consideration of individual variation and dominance profiles. Larger datasets would also improve the representativeness of morphosyntactic structures and features, such as templatic alternations, which were underrepresented in the present corpus. Future work could therefore broaden longitudinal sampling and extend the developmental window beyond 60 months to trace how templatic morphology and number agreement consolidate during the early school years.

Linguistic range and methodological boundaries

The present analyses focused primarily on nominal inflection and agreement in relation to broader structural and lexical growth, in line with the study's aims. The exclusion of other POS, inflectional and agreement relations not analysed here, and derivational processes[83] inevitably restricted observation of additional patterns of bilingual grammatical development. Future studies could extend analyses to these areas to provide a more comprehensive account of morphosyntactic development.

Although naturalistic data ensured ecological validity, the fact that the main interactions took place between the child and the investigator, that is, an outsider to the child's usual home environment, introduced variability related to context, recording conditions, and interactional setting. For instance, children's use of English lexical items may have been heightened during recorded sessions since many childcare and educational institutions operate in English, and the investigator may have been associated with such institutions. Recording children in their home environment, with family members and during daily routines, could enhance ecological validity further and allow parallel analysis of CDS, including the quantification of input frequency. Such naturalistic setups are encouraged by the Acquisition Sketch Project (Hellwig et al., 2023), where an acquisition sketch for Maltese is currently in progress.

While combining corpus data with controlled elicitation proved particularly effective in this study, no purely experimental tasks were conducted, as the

[83] Derivational processes are still relatively underexplored in CLA (e.g., Ravid, 2019).

receptive and production activities were carried out in a casual manner within the home environment during play visits. Future methodological advancement in MFLA could therefore come from incorporating experimental paradigms such as priming and eye-tracking, which would enable the investigation of morphosyntactic processing in real time and strengthen the empirical foundations for cross-linguistic comparison. From a technical standpoint, although statistical tests were applied, future research could adopt advanced modelling approaches to link child-level measures (e.g., MLUm and POS category diversity) more directly to morphosyntactic accuracy, allowing for formal testing of the interdependencies observed here.

6.4.2. Observations beyond the scope of the present study

Additional directions emerging from the data

Although the focus of this study is morphosyntactic, the corpus also includes observations of potential interest to adjacent domains, offering promising avenues for future research. Four such strands are briefly outlined below.

a) Pragmatic use of language

Several examples show children using language for explanatory or justificatory purposes, extending beyond simple reference or naming. At 36 months, one child explains discomfort caused by a clothing label, while at 60 months another provides a justification linked to the investigator’s toy:

(1) *qisha għandi xi /paper/ ġo saqajja* (Anna, 36 months)
'It's like I have some paper in my legs.'

(2) *kont ħa nsiblek is-sellum għax ħallejt xi sellum għandi* (Reno, 60 months)
'I was going to look for a ladder for you because you left a ladder here with me.'

Such utterances signal the emergence of explanatory and justificatory functions of language which offer valuable directions for future research on pragmatic development.

b) Psychological salience of age milestones

Around the age of three, children begin to reflect metalinguistically on their own growth and achievements (e.g., Chaney 1992; Taylor & van den Herik, 2021). Examples from the corpus illustrate how developmental milestones are woven into self-narratives:

(3) *dalwaqt jiġi l-/birthday/ tiegħi* (Kelina, 35 months)
'It's soon my birthday.'

(4) *il-/big boy bed/ għax għalaqt /three/* (Toninu, 36 months)
'The big boy bed because I turned three.'

(5) *għall-/birthday three/ armejt id-/dummies/ kollha* (Lippu, 38 months)
'For my third birthday I threw away all my pacifiers.'

These observations highlight a developmental phase in which linguistic, cognitive, and socioemotional growth intersect, an area ripe for future interdisciplinary investigation. Interestingly, several metalinguistic comments in the corpus take the form of questions, often expressing spontaneity, humour and curiosity about adult behaviour or social norms. Such interrogatives reveal a growing capacity for perspective-taking and reasoning about communicative contexts, suggesting potential for both pragmatic and psychological analyses of the corpus, as well as for dedicated studies of question formation due to their abundance in the data:

(6) *għaliex int kbira u xorta tilgħab bit-/toys/?* (Lippu, 50 months)
'How come you're a grown-up and you still play with toys?'

c) Negation and polarity

Negative constructions occur frequently across both spontaneous and elicited (e.g., spot-the-difference) tasks but were not analysed in the present work. Their formal and functional diversity makes them a valuable locus for future study of early morphosyntactic and pragmatic development. Analysing such patterns could also shed light on cross-linguistic variation in negation strategies within bilingual acquisition contexts.

d) Narrative development

The narrative tasks collected as part of this project provide a rich resource for examining structure and coherence in early storytelling. A collaborative study on *Frog Story* narratives is currently being initiated, comparing MD and Estonian-dominant children. Further work could trace how narrative devices evolve in relation to morphosyntactic competence, exposure context, and broader sociocognitive development, contributing to a more integrated understanding of early narrative ability.

Societal and Contextual Factors

All participants in this study were born soon before or during the COVID-19 pandemic, a factor increasingly recognised as relevant to early linguistic development (e.g., Zuniga-Montanez et al., 2025). Future research could compare pre-, during-, and post-pandemic cohorts to examine whether differences in socialisation or exposure have measurable linguistic effects.

Another emerging dimension concerns children's immersion in digital environments. The corpus offers glimpses of children's awareness and use of technology. During one session, upon noting that the investigator's phone was not present, a child under 3 years of age inquired about it. On another instance, at 43 months a child enthusiastically talked about the digital games he enjoys playing.

(7) *m'għandekx mowbajl int?* (Lippu, 33 months)
'Don't you have a mobile phone?'

(8) *il-papà għandu tablet għalija biex noqgħod nilgħab logħba* (Toninu, 43 months)
'My dad has a tablet for me to stay playing a game.'

Such examples underscore the pervasive role of digital experience in young children's communicative expectations and highlight the need for future research exploring the relationship between screen exposure, attention, and bilingual language development.

6.4.3. Future Directions

Building on this study's findings, future research could:

- Broaden the participant pool to capture variation across language-dominance profiles, including English-dominant children to assess how dominance influences morphosyntactic development.

- Extend the analysis to a broader range of grammatical categories, particularly the verbal domain (e.g., tense–aspect marking and agreement), to determine developmental patterns emerging beyond nominal morphosyntax.

- Integrate corpus-based analyses with elicitation and experimental techniques, such as priming and eye-tracking, to examine both production and real-time processing of morphosyntactic forms.

- Examine frequency effects in CDS, focusing on how input salience and relative frequency contribute to acquisition.

- Investigate pragmatic, narrative, and metalinguistic dimensions of MFLA, including children's ability to recount sequences of events and reflect on linguistic form.

- Undertake cross-linguistic comparisons with other bilingual pairings (e.g., Vaks et al., 2025), including languages with similar morphological typologies to identify convergent patterns, as well as typologically distinct languages to assess the generalisability of the observed findings.

The corpus compiled through this project comprises over 30 hours of spontaneous interaction and free play, thus represents a valuable empirical resource that can be used for further research and expanded with new data. Hence, it is hoped that the present work will continue to inform and inspire future linguistic, cognitive, and interdisciplinary research in the field.

6.5. Final Remarks

Viewed in its entirety, this study reveals early development of the nominal system as a gradual convergence of structure, lexicon, and morphosyntax. Between the second and fifth years of life, TD, MD children move from isolated lexical productions to richly structured morphosyntactic constructions, drawing first on transparent patterns and gradually extending these to more complex domains.

The longitudinal lens afforded by data spanning 18–60 months captures this transformation in real time. It is within this window that grammatical complexity and productivity take root and POS categories diversify, as bilingual resources interweave into a single, cohesive linguistic system.

Ultimately, this research demonstrates that CLA in the morphologically rich and bilingual context of Maltese is a dynamic and adaptive process of integration. By tracing the trajectory of nominal development in TD, MD children, the study contributes to cross-linguistic theory on morphosyntactic learnability and the emergence of bilingual grammar: an intricate linguistic architecture shaped by transparency, interaction, and the creativity of the developing mind.

References

Abdalla, F., & Crago, M. (2008). Verb morphology deficits in Arabic-speaking children with specific language impairment. *Applied Psycholinguistics, 29*(2), 315–340. doi:10.1017/S0142716408080156

Abdalla, F., Aljenaie, K., & Mahfoudhi, A. (2013). Plural noun inflection in Kuwaiti Arabic-speaking children with and without Specific Language Impairment. *Journal of Child Language*, *40*(1), 139–168. https://doi.org/10.1017/S0305000912000499

Abdilla, T. & Spagnol, M. (2016). L-għadd tal-kollettiv Malti. Leħen il-Malti, 35, 169-182.

Adone, D. (2012). *The acquisition of creole languages: How children surpass their input*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139043366

Ahn, D., & Snedeker, J. (2022). Early Acquisition of Plural Morphology in a Classifier Language: Data from Korean 2-4 Year Olds. *Language Learning and Development*, *18*(1), 1–15. https://doi.org/10.1080/15475441.2021.1922280

Akhtar, N. and Tomasello, M. (2000). The social nature of words and word learning. In R. Golinkoff & K. Hirsh-Pasek (Eds.), *Becoming a word learner: A debate on lexical acquisition* (pp. 115–135). Oxford University Press.

Albirini, A. (2015). Factors affecting the acquisition of plural morphology in Jordanian Arabic. *Journal of Child Language, 42*(4), 734–762. https://doi.org/10.1017/S0305000914000270

Albirini, A., & Benmamoun, E. (2014). Concatenative and nonconcatenative plural formation in L1, L2, and heritage speakers of Arabic. *The Modern Language Journal, 98*(3), 854–871. https://doi.org/10.1111/modl.12126

Albright, A., & Hayes, B. (2003). Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition, 90*(2), 119–161.

Allen, S. (2024). *Is acquisition any different in a polysynthetic language? The case of Inuktitut* [Keynote presentation]. Conference of the International Max Planck Research School for Language Sciences, Nijmegen, Netherlands.

Allen, S. E., & Crago, M. B. (1992). First language acquisition of Inuktitut. In M. J. Dufour & F. Therien (Eds.), *Looking to the future: Papers from the Seventh Inuit Studies Conference* [Inuit Studies Occasional Papers 4] (p. 273-281). Quebec, QC: Association Inuksiutiit Katimajiit.

Allotey, J., Zamora, J., Cheong-See, F., Kalidindi, M., Arroyo-Manzano, D., Asztalos, E., van der Post, J. A. M., Mol, B. W., Moore, D., Birtles, D., Khan, K. S., & Thangaratinam, S. (2018). Cognitive, motor, behavioural, and academic performances of children born preterm: A meta-analysis and systematic review involving 64 061 children. *BJOG: An International Journal of Obstetrics and Gynaecology, 125*(1), 16–25. https://doi.org/10.1111/1471-0528.14832

Allott, N., Lohndal, T., & Rey, G. (2021). *A companion to Chomsky*. Wiley Blackwell. https://10.1002/9781119598732

Almoammer, A., Sullivan, J., Donlan, C., Marušič, F., Žaucer, R., O'Donnell, T., & Barner, D. (2013). Grammatical morphology as a source of early number word meanings. *Proceedings of the National Academy of Sciences - PNAS*, *110*(46), 18448–18453. https://doi.org/10.1073/pnas.1313652110

Ambridge, B. (2020). Against stored abstractions: A radical exemplar model of language acquisition. *First Language*, *40*(5–6), 509–559. https://doi.org/10.1177/0142723719869731

Ambridge, B., & Lieven, E. (2011). *Child language acquisition*. Cambridge: Cambridge University Press. https://10.1017/CBO9780511975073

Ambridge, B., & Rowland, C. F. (2013). Experimental methods in studying child language acquisition. *Wiley Interdisciplinary Reviews. Cognitive Science*, *4*(2), 149–168. https://doi.org/10.1002/wcs.1215

Ambridge, B., & Wagner, L. (2021). Testable theories of core first language acquisition. *Journal of Child Language, 48*(5), 859–861. https://10.1017/S0305000921000581

Ambridge, B., Pine, J. M., Rowland, C. F., & Young, C. R. (2008). The effect of verb semantic class and verb frequency (entrenchment) on children's and adults' graded judgements of argument-structure overgeneralization errors. *Cognition, 106*(1), 87–129. https://doi.org/10.1016/j.cognition.2006.12.015

Ambridge, B., Pine, J. M., Rowland, C. F., Freudenthal, D., & Chang, F. (2012). Avoiding dative overgeneralisation errors: Semantics, statistics, or both? *Language, Cognition and Neuroscience, 29*(2), 218–243. https://doi.org/10.1080/01690965.2012.738300

Ambridge, B., Rowland, C. F., Theakston, A. L. & Kidd, E. J. (2015). The ubiquity of frequency effects in first language acquisition. Journal of Child Language 42(2), 239–273.

Andersen, E. S. (1990). Speaking with style: The sociolinguistic skills of children. London: Routledge.

Anderson, N. J., Graham, S. A., Prime, H., Jenkins, J. M., & Madigan, S. (2021). Linking quality and quantity of parental linguistic input to child language skills: A

Meta-Analysis. *Child Development, 92*(2), 484–501. https://10.1111/cdev.13508

Anderson, S. R. (2015). The morpheme: Its nature and use. In M. Baerman (Ed.), *The Oxford handbook of inflection* (pp. 11–33). Oxford University Press.

Applebee, A. (1978). The child's concept of a story: Ages 2 to 17. Chicago, IL: University of Chicago Press.

Aquilina, J., & Isserlin, B. S. J. (Eds.). (1981). *A survey of contemporary dialectal Maltese: Volume I – Gozo.* Leeds University Press.

Aquilina, Joseph. 1959. *The Structure of Maltese*. Msida: Malta University Press.

Arunachalam, S., Leddon, E. M., Song, H., Lee, Y., & Waxman, S. R. (2013). Doing more with less: Verb learning in Korean-acquiring 24-month-olds. *Language Acquisition, 20*(4), 292–304. https://10.1080/10489223.2013.828059

Attard, D. (2016). *Comprehension and production of nouns and verbs in Maltese five-year olds* [Bachelor's dissertation, University of Malta]. University of Malta.

Ayoun, D. (2022). Indeterminacy in L1 French grammars: the case of gender and number agreement. *Journal of French Language Studies*, *32*(3), 327–359. doi:10.1017/S0959269521000247

Azzopardi-Alexander, M. (2016). Language profiling: The weaving of Maltese and English in Maltese children's conversations. In G. Puech & B. Saade (Eds.), *Shifts and patterns in Maltese* (Vol. 19, pp. 331–385). De Gruyter Mouton.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390–412.

Baayen, R. H., Milin, P., Đurđević, D. F., Hendrix, P., & Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review, 118*(3), 438–481. https://doi.org/10.1037/a0023851

Babatsouli, E. (2024). Input, Universals, and Transfer in Developing Rhotics: A Sketch in Bilingualism. *Languages (Basel)*, *9*(10), 328-. https://doi.org/10.3390/languages9100328

Baetens Beardsmore, H. (1986). *Bilingualism: Basic principles* (2nd ed.). Multilingual Matters.

Bagheri, M. S., Rahimi, F., & Riasati, M. J. (2012). Communicative Interaction in Language Learning Tasks among EFL Learners. Journal of Language Teaching and Research, 3(5), 948-. https://doi.org/10.4304/jltr.3.5.948-952

Bail, A., Morini, G., & Newman, R. S. (2015). Look at the gato! Code-switching in speech to toddlers. Journal of Child Language, 42, 1073–1101. https://doi.org/10. 1017/S0305000914000695

Baird, P. (2024). Language choice and code-mixing in a longitudinal study of an Estonian–English bilingual child. *Taikomoji Kalbotyra*, *21*, 35-51. https://doi.org/10.15388/Taikalbot.2024.21.3

Bajada, M. A. (2008). *Investigating the relationship between expressive vocabulary and MLU* [Bachelor's dissertation, University of Malta]. University of Malta.

Balason, D. V., & Dollaghan, C. A. (2002). Grammatical Morpheme Production in 4-Year-Old Children. *Journal of Speech, Language, and Hearing Research*, *45*(5), 961–969. https://doi.org/10.1044/1092-4388(2002/078)

Baldacchino, R. (2020). *The effects of primary caregiver language input on the lexical skills of 3-year-old children from different socioeconomic backgrounds* [Master's dissertation, University of Malta]. University of Malta.

Baldwin, D. A. & Moses, L. J. (2001). Links between social understanding and early word learning: Challenges to current accounts. *Social Development, 10*(3), 309–329.

Barner, D., Thalwitz, D., Wood, J., Yang, S.-J., & Carey, S. (2007). On the relation between the acquisition of singular-plural morpho-syntax and the conceptual distinction between one and more than one. *Developmental Science*, *10*(3), 365–373. https://doi.org/10.1111/j.1467-7687.2007.00591.x

Baron, A., Wagley, N., Hu, X., & Kovelman, I. (2023). Neural Correlates of Morphosyntactic Processing in Spanish–English Bilingual Children: A Functional Near-Infrared Spectroscopy Study. *Journal of Speech, Language, and Hearing Research*, *66*(9), 3500–3514. https://doi.org/10.1044/2023_JSLHR-22-00598

Bartolo, D. (2014). *Il-Lingwaġġ tat-Tfal: Studju tat-Taħdit ta' Tifla ta' Sitt Snin [The Language of Children: A Study of the Speech of a Six-Year-Old Girl]* [Bachelor's dissertation, University of Malta]. University of Malta.

Bartolo, K. (2021). *Il-parteċipju passiv fil-Malti [The passive participle in Maltese]*. [Bachelor's dissertation, University of Malta]. University of Malta.

Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P., Reznick, J. S., Reilly, J., & Hartung, J. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language, 21*(1), 85–123. https://doi.org/10.1017/S0305000900008680

Bauer, L. (2016). Classical morphemics: Assumptions, extensions, and alternatives. In A. Hippisley & G. Stump (Eds.), *The Cambridge handbook of morphology* (pp. 331-355). Cambridge University Press. https://doi.org/10.1017/9781139814720.013

Bauer, L. (2019). *Rethinking morphology*. Edinburgh University Press.

Bavin, E. L., & Naigles, L. R. (Eds.). (2015). *The Cambridge handbook of child language* (2nd ed.). Cambridge University Press.

Bavin, E. L., & Naigles, L. R. (2015). Introduction: Perspectives on child language. In *The Cambridge Handbook of Child language* (2nd ed) (pp. 1–12). Cambridge: Cambridge University Press.

Behrens, H. (2006). The input-output relationship in first language acquisition. *Language and Cognitive Processes*, *21*(1–3), 2–24. https://doi.org/10.1080/01690960400001721

Behrens, H. (2015). The acquisition of grammatical categories. In E. L. Bavin & L. R. Naigles (Eds.), *The Cambridge handbook of child language* (2nd ed., pp. 250–270). Cambridge University Press.

Behrens, H. (2021). Constructivist approaches to first language acquisition. *Journal of Child Language, 48*(5), 959–983. https://10.1017/S0305000921000556

Bergelson, E., Soderstrom, M., Schwarz, I.-C., Rowland, C. F., Ramírez-Esparza, N., R Hamrick, L., Marklund, E., Kalashnikova, M., Guez, A., Casillas, M., Benetti, L., Alphen, P. van, & Cristia, A. (2023). Everyday language input and production in 1,001 children from six continents. *Proceedings of the National Academy of Sciences - PNAS*, *120*(52), e2300671120-. https://doi.org/10.1073/pnas.2300671120

Berglund, E., Eriksson, M., & Westerlund, M. (2005). Communicative skills in relation to gender, birth order, childcare and socioeconomic status in 18-month-old children. *Scandinavian Journal of Psychology, 46*(6), 485-491. https://10.1111/j.1467-9450.2005.00480.x

Berk, S., & Lillo-Martin, D. (2012). The two-word stage: Motivated by linguistic or cognitive constraints? *Cognitive Psychology, 65*(1), 118–140.

Berko Gleason, J., & Ratner, N. B. (2016). *The Development of language*. Pearson Australia Pty Ltd.

Berko, J. (1958). The child's learning of English morphology. *Word, 14*(2-3), 150–177. https://doi.org/10.1080/00437956.1958.11659661

Berko, J. (1985). Studying language development. In J. Berko Gleason, (Ed.), *The development of language* (4th ed.). (pp. 1-35). Colombus, Ohio: Charles E. Merrill Publishing Company.

Berman, R. (2003). 13. Children's lexical innovations: Developmental perspectives on Hebrew verb structure. In J. Shimron (Ed.), *Language Processing and Acquisition in Languages of Semitic, Root-Based, Morphology* (pp. 243-291). John Benjamins Publishing Company. https://doi.org/10.1075/lald.28.13ber

Berman, R. (2016). *Acquisition and development of Hebrew: From infancy to adolescence*. Amsterdam, The Netherlands: John Benjamins.

Berman, R. A. (1985). The acquisition of Hebrew. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition, Vol. 1: The data* (pp. 255-371). Lawrence Erlbaum Associates.

Berman, R., & Slobin, D., I. (1994). *Relating Events in Narrative: A crosslinguistic Developmental Study.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Bernardini, P., & Schlyter, S. (2004). Growing syntactic structure and code-mixing in the weaker language: The ivy hypothesis. Bilingualism: Language and Cognition, 7, 49–69. doi:10.1017/S1366728904001270

Berwick, R. C. (1985). *The acquisition of syntactic knowledge*. Cambridge, MA: MIT Press. https://10.7551/mitpress/1074.001.0001

Berwick, R., & Chomsky, N. (2016). *Why Only US? Language and Evolution*. Cambridge MA: MIT press.

Bezzina, R. (2020). *L-allomorfija fil-Malti: B'referenza speċjali għall-imtenni u d-diminuttiv [Allomorphy in Maltese: With special reference to the dual and the diminutive]* [Bachelor's dissertation, University of Malta]. University of Malta.

Bidgood, A., Pine, J., Rowland, C., Sala, G., Freudenthal, D., & Ambridge, B. (2021). Verb argument structure overgeneralisations for the English intransitive and transitive constructions: Grammaticality judgments and production priming. *Language and Cognition, 13*(3), 397–437. https://doi.org/10.1017/langcog.2021.8

Bigi, B., & Priego-Valverde, B. (2022). The automatic search for sounding segments of SPPAS: Application to Cheese! corpus. In *Human language technology: Challenges for computer science and linguistics* (Vol. 13212, pp. 16–27). *Lecture Notes in Computer Science* [978-3-031-05328-3]. https://doi.org/10.1007/978-3-031-05328-3_2

Bittner, D., Dressler, W. U., & Kilani-Schoch, M. (2003). Introduction. In D. Bittner, W. U. Dressler, & M. Kilani-Schoch (Eds.), *Development of verb inflection in first language acquisition: A cross-linguistic perspective* (vii–xl). Studies on Language Acquisition 21, De Gruyter Mouton.

Björnsdóttir, S. (2021). Productivity and the acquisition of gender. *Journal of Child Language, 48*(6), 1209-1234. doi:10.1017/S0305000920000732

Björnsdóttir, S. M., (2023). Predicting ineffability: Grammatical gender and noun pluralization in Icelandic. *Glossa: a journal of general linguistics* 8(1). doi: https://doi.org/10.16995/glossa.5823

Blom, E., & Unsworth, S. (2010). Experimental Methods in Language Acquisition Research. John Benjamins Publishing Company.

Blom, E., Bosma, E., & Heeringa, W. (2021). Regular and Irregular Inflection in Different Groups of Bilingual Children and the Role of Verbal Short-Term and Verbal Working Memory. *Languages*, *6*(1), 56. https://doi.org/10.3390/languages6010056

Blom, E., Jong, J., Orgassa, A., Baker, A., & Weerman, F. (2013). Verb inflection in monolingual Dutch and sequential bilingual Turkish–Dutch children with and without SLI. *International Journal of Language & Communication Disorders*, *48*(4), 382–393. https://doi.org/10.1111/1460-6984.12013

Bloom, L. (1970). Language development: Form and function in emerging grammars. Cambridge, MA: MIT Press.

Bloom, L., Lightbown, P., & Hood, L. (1975). Structure and variation in child language. *Monographs of the Society for Research in Child Development*, *40*(2), 1–97. https://doi.org/10.2307/1165986

Bloom, P. (2004). Myths of Word Learning. In D. G. Hall & S. R. Waxman (Eds.), *Weaving a lexicon* (pp. 205–224). Boston Review.

Bloomfield, L. (1933). *Language*. New York: Holt, Rinehart and Winston.

Blount, B. G. (1988). Cognition and phonology in acquisition of plurals and possessives by Luo children. *Language Sciences*, 10(1), 225–240.

Blum-Kulka, S. (2017). Language socialization and family dinnertime discourse. In P. Duff & S. May (Eds.), *Language socialization* (3rd ed., pp. 3-16). Springer.

Boersma, P., & Weenik, D. (2009). *Praat: Doing Phonetics by Computer* (Version 5.1.07).

Borg, A. (1988). The maintenance of Maltese as a language: What chances? *Council of Europe 1988*, 89-106.

Borg, A., & Azzopardi-Alexander, M. (1997). *Maltese* (Descriptive grammars/ series editor, B. Comrie). London: Routledge.

Borg, C., & Gatt, A. (2017). Morphological Analysis for the Maltese Language: The challenges of a hybrid system. In Proceedings of the Third Arabic Natural Language Processing Workshop, Valencia. (25–34).

Bornstein, M. H., Cote, L. R., Maital, S., Painter, K., Park, S. Y., Pascual, L., Pêcheux, M. G., Ruel, J., Venuti, P., & Vyt, A. (2004). Cross-linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. *Child Development*, *75*(4), 1115–1139. https://doi.org/10.1111/j.1467-8624.2004.00729.x

Bosch, S., Veríssimo, J., & Clahsen, H. (2019). Inflectional morphology in bilingual language processing: An age-of-acquisition study. *Language Acquisition*, *26*(3), 339–360. https://doi.org/10.1080/10489223.2019.1570204

Bosse, M. L., Brissaud, C., & Le Levier, H. (2021). French Pupils' Lexical and Grammatical Spelling from Sixth to Ninth Grade: A Longitudinal Study. *Language and speech*, *64*(1), 224–249. https://doi.org/10.1177/0023830920935558

Bovingdon, R. (2004). From Language to Ethnolect: Maltese to Maltraljan. *Current Issues in Language Planning*, *5*(2), 166–175. https://doi.org/10.1080/13683500408668256

Bowerman, M. (1973). *Early syntactic development: A cross-linguistic study with special reference to Finnish.* Cambridge: Cambridge University Press.

Bowerman, M. (2011). Linguistic typology and first language acquisition. In J. J. Song (Ed.), *The Oxford handbook of linguistic typology* (pp. 591–617). Oxford University Press.

Braginsky, M., Yurovsky, D., Marchman, V. A., & Frank, M. C. (2019). Consistency and variability in children's word learning across languages. *Open Mind: Discoveries in Cognitive Science, 3*, 52–67. https://doi.org/10.1162/opmi_a_00026

Braine, M. D. S. (1963). The ontogeny of English phrase structure: The 1st phase. *Language, 39*(1), 1–13.

Braine, M. D. S., & Brooks, P. J. (1995). Verb argument structure and the problem of avoiding an overgeneral grammar. In M. Tomasello, & W. E. Merriman (Eds.), *Beyond names for things: Young children's acquisition of verbs* (pp. 353–376). Lawrence Erlbaum Associates, Inc.

Brennan-Jones, C. G., Whitehouse, A. J. O., Calder, S. D., Da Costa, C., Eikelboom, R. H., Swanepoel, D. W., & Jamieson, S. E. (2020). Does otitis media affect later language ability? A prospective birth cohort study. *Journal of Speech, Language, and Hearing Research*, *63*(7), 2441–2452. https://doi.org/10.1044/2020_JSLHR-19-00005

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

Brown, P. (2012). The cultural organization of attention. In A. Duranti, E. Ochs, & B. B. Schieffelin (Eds.), *The handbook of language socialization* (pp. 29–55). Wiley-Blackwell.

Brown, R. (1973). *A first language: The early stages*. Harvard University Press.

Bruner, J. S. (1981). The social context of language acquisition. *Language & Communication, 1*(2-3), 155–178. https://doi.org/10.1016/0271-5309(81)90010-0

Bruner, J. S. (1983). *Child's talk: Learning to use language*. Oxford University Press.

Bybee, J. (1985). Morphology: A study of the relation between meaning and form. Amsterdam, The Netherlands: John Benjamins.

Bybee, J. (2001). *Phonology and language use*. Cambridge: Cambridge University Press.

Bybee, J. (2007). *Frequency of use and the organization of language*. In M. Tomasello (Ed.), *The new psychology of language: Cognitive and functional approaches to language structure* (pp. 229–254). Erlbaum.

Bybee, J. L. (2010). *Language, usage and cognition*. Cambridge: Cambridge University Press.

Byers-Heinlein, K., Burns, T. C., & Werker, J. F. (2010). The Roots of Bilingualism in Newborns. *Psychological Science*, *21*(3), 343–348. https://doi.org/10.1177/0956797609360758

Byers-Heinlein, K. (2013). Parental language mixing: Its measurement and the relation of mixed input to young bilingual children's vocabulary size. *Bilingualism: Language and Cognition, 16*(1), 32–48. https://doi.org/10.1017/S1366728912000120

Cabiddu, F., Bott, L., Jones, G., & Gambi, C. (2023). CLASSIC utterance boundary: A chunking-based model of early naturalistic word segmentation. *Language Learning*, *73*(3), 942–975. https://doi.org/10.1111/lang.12559

Cachia, P. (1982). *Aspects of the language of Maltese five year olds* [Bachelor's dissertation, University of Malta]. University of Malta.

Camilleri Grima, A. (2016). Young children living bilingually in Malta. *Lingwistyka Stosowana, 17*(2), 1–13.

Camilleri Grima, A. (2020). The vitality of Maltese dialects in Gozo. The Gozo Observer, 41, 8-18.

Camilleri, A. (1995). *Bilingualism in education: The Maltese experience*. Groos.

Camilleri, P. M. (1987). *Christopher: a child in perspective*. [Bachelor's dissertation, University of Malta]. University of Malta.

Campbell, R., & Wales, R. (1970). The study of language acquisition. In J. Lyons (Ed.), *New horizons in linguistics* (pp. 242–260). Penguin Books.

Cantone, K. N., Kupisch, T., Müller, K., & Schmitz, K. (2008). Rethinking language dominance. *Linguistische Berichte, 215*, 129–160.

Caruana Lia, D. (2016). *L-użu tal-Malti u l-Ingliż f'xi klassijiet fil-livell tal-edukazzjoni bikrija [The use of Maltese and English in some early education classes]* [Bachelor's dissertation, University of Malta]. University of Malta.

Caruana, S. (2025). Maltese English. In K. Bolton (Ed.), *The Wiley Blackwell Encyclopedia of World Englishes*. Wiley Blackwell. https://doi.org/10.1002/9781119518297.eowe00357

Caruana, S., Vassallo Gauci, P., & Scaglione, S. (2019). Multilingualism and the inclusion of migrant learners in Maltese schools. In C. Bezzina & S. Caruana (Eds.), *Teacher education matters: Transforming lives ... transforming schools* (pp. 330–343). University of Malta.

Casillas, M., Brown, P., & Levinson, S. C. (2020). Early language experience in a Tseltal Mayan village. *Child Development, 91*, 1819–1835. https://doi.org/10.1111/cdev.13349

Cassar, M. (1986). *Maria, the first five years* [Bachelor's dissertation, University of Malta]. University of Malta.

Castilla-Earls, A., Ronderos, J., & Francis, D. J. (2023). Longitudinal examination of morphosyntactic skills in bilingual children: Spanish and English standardized scores. *Journal of Speech, Language, and Hearing Research, 66*(8), 2671–2687. https://doi.org/10.1044/2023_JSLHR-22-00495

Cattell, R. (2000). *Children's language: Consensus and controversy.* Cassell.

Cazden, C. B. (1968). The acquisition of noun and verb inflections. *Child Development, 39*(2), 433–448. https://doi.org/10.2307/1126956

Chafe, W. (Ed). 1980. *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production.* Ablex. Video URL: https://www.youtube.com/watch?v=bRNSTxTpG7U

Chamam, S., Forcella, A., Musio, N., Quinodoz, F., & Dimitrova, N. (2024). Effects of digital and non-digital parental distraction on parent–child interaction and communication. *Frontiers in Child and Adolescent Psychiatry, 3*, 1330331. https://doi.org/10.3389/frcha.2024.1330331

Chaney, C. (1992). Language development, metalinguistic skills, and print awareness in 3-year-old children. *Applied Psycholinguistics*, *13*(4), 485–514. doi:10.1017/S0142716400005774

Checa, E., Galeote, M., & Soto, P. (2016). The composition of early vocabulary in Spanish children with Down syndrome and their peers with typical development. *American Journal of Speech-Language Pathology*, *25*(4), 605–619. https://doi.org/10.1044/2016_AJSLP-15-0095

Cheng, Y., Cunnings, I., Miller, D., & Rothman, J. (2022). Double-number marking matters for both L1 and L2 processing of nonlocal agreement similarly: An ERP investigation. *Studies in Second Language Acquisition, 44*(5), 1309–1329. https://doi.org/10.1017/S0272263121000772

Choi, S., & Gopnik, A. (1995). Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of Child Language*, *22*(3), 497–529. https://doi.org/10.1017/s0305000900009776

Chomsky, N. (1957). *Syntactic structures*. Mouton.

Chomsky, N. (1959). A review of B. F. Skinner's *Verbal Behavior*. *Language, 35*(1), 26–58.

Chomsky, N. (1965). *Aspects of the theory of syntax.* MIT Press.

Chomsky, N. (1975). *Reflections on Language*. Pantheon Books.

Chomsky, N. (1981a). On the representation of form and function. *The Linguistics Review, 1*(1), 3–40.

Chomsky, N. (1981b). *Lectures on government and binding*. Foris Publications.

Chomsky, N. (1988). *Language and problems of knowledge.* MIT Press.

Chomsky, N. (1995). *The minimalist program*. MIT Press.

Chomsky, N. (2001). Derivation by phase. In M. Kentowicz (Ed.), *A life in language* (pp. 1–52). MIT Press.

Chomsky, N. (2015). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (2017). The language capacity: architecture and evolution. *Psychonomic Bulletin & Review*, *24*(1), 200–203. https://doi.org/10.3758/s13423-016-1078-6

Clark, E. (2016). *First language acquisition* (3rd ed.). Cambridge University Press.

Clark, E. (2024). *First language acquisition* (4th ed.). Cambridge University Press.

Clark, E. V. (1985). The acquisition of Romance, with special reference to French. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition* (Vol. 1, pp. 687–782). Hillsdale, NJ: Lawrence Erlbaum Associates.

Clark, E. V. (1993). *The lexicon in acquisition.* Cambridge University Press.

Clark, E. V. (2009). *First language acquisition* (2nd ed.). Cambridge University Press.

Clark, E., & Nikitina, T. (2009). One vs. more than one: Antecedents to plural marking in early language acquisition. *Linguistics, 47*(1), 103–139. https://doi.org/10.1515/LING.2009.004

Comeau, L., Genesee, F., & Lapaquette, L. (2003). The modeling hypothesis and child bilingual codemixing. *International Journal of Bilingualism, 7*(2), 113–126.

Comrie, B., Fabri, R., Hume, E., Mifsud, M., Stolz, T., & Vanhove, M. (Eds.). (2009). *Introducing Maltese linguistics: Proceedings of the First Meeting of the International Association of Maltese Linguistics.* John Benjamins.

Comrie, B. (2009). Malta and the World Atlas of Language Structures. In B. Comrie, R. Fabri, E. Hume, M. Mifsud, T. Stolz, & M. Vanhove (Eds.), *Introducing Maltese linguistics* (pp. 3–11). John Benjamins.

Comrie, B., & Spagnol, M. (2016). *Maltese loanword typology.* De Gruyter Mouton.

Conboy, B. T., & Thal, D. J. (2006). Ties between the lexicon and grammar: Cross-sectional and longitudinal studies of bilingual toddlers. *Child Development, 77*(3), 712–735. https://doi.org/10.1111/j.1467-8624.2006.00899.x

Conboy, B. T., Rivera-Gaxiola, M., Silva-Pereyra, J., & Kuhl, P. K. (2008). Event-related potential studies of early language processing at the phoneme, word, and sentence levels. In A. D. Friederici & G. Thierry (Eds.), *Early language development: Bridging brain and behaviour* (pp. 23–64). John Benjamins Publishing Company. https://doi.org/10.1075/tilar.5.04con

Coppini, A. (2002). *The acquisition of number and gender agreement between nouns and adjectives in 3- to 5-year-old Maltese-speaking children* [Bachelor's dissertation, University of Malta]. University of Malta.

Croft, W., & Cruse, D. A. (2004). *Cognitive linguistics.* Cambridge University Press. https://doi.org/10.1017/CBO9780511803864

Crystal, D. (1974). [Review of the book *A first language: The early stages*, by R. Brown]. *Journal of Child Language*, *1*(2), 289–307. https://doi.org/10.1017/S030500090000074X

Crystal, D. (1998). *Language play*. University of Chicago Press.

Culbertson, J., Jarvinen, H., Haggarty, F., & Smith, K. (2019). Children's sensitivity to phonological and semantic cues during noun class learning: Evidence for a phonological bias. *Language, 95*(2), 268–293. https://doi.org/10.1353/lan.0.0234

Cyr, M., & Shi, R. (2013). Development of abstract grammatical categorization in infants. *Child Development, 84*(2), 617–629.

Dąbrowska, E. (2001). Learning a morphological system without a default: The Polish genitive. Journal of Child Language, 28(3), 545–574.

Dąbrowska E. (2015). What exactly is Universal Grammar, and has anyone seen it? *Frontiers in psychology*, *6*, 852. https://doi.org/10.3389/fpsyg.2015.00852

Dąbrowska, E. (2020). Language as a phenomenon of the third kind. *Cognitive Linguistics, 31*(2), 213–229. https://doi.org/10.1515/cog-2019-0029

Daelemans, W. (2002). A comparison of analogical modeling to memory-based language processing. In R. Skousen, D. Lonsdale, & D. B. Parkinson (Eds.), *Analogical modeling* (pp. 157–179). Benjamins.

Dandria, R. (2002). *Lexical access and prototypicality in the English of Maltese speakers of English* [Bachelor's dissertation, University of Malta]. University of Malta.

Darwin, C. (1877). A biographical sketch of an infant. *Mind: A Quarterly Review of Psychology and Philosophy, 2*, 285–294.

Davies, B., Xu Rattanasone, N., & Demuth, K. (2017). Two-year-olds' sensitivity to inflectional plural morphology: Allomorphic effects. *Language Learning and Development, 13*, 38–53.

Davies, B., Xu Rattanasone, N., & Demuth, K. (2020). Acquiring the Last Plural: Morphophonological Effects on the Comprehension of /-əz. *Language Learning and Development*, *16*(2), 161–179. https://doi.org/10.1080/15475441.2020.1717956

Davies, B., Xu Rattanasone, N., Schembri, T., & Demuth, K. (2019). Preschoolers' developing comprehension of the plural: The effects of number and allomorphic variation. *Journal of Experimental Child Psychology*, *185*, 95–108. https://doi.org/10.1016/j.jecp.2019.04.015

Davison, M. D., & Hammer, C. S. (2012). Development of 14 English grammatical morphemes in Spanish-English preschoolers. *Clinical Linguistics & Phonetics*, *26*(8), 728–742. https://doi.org/10.3109/02699206.2012.700679

Dawdy-Hesterberg, L. G., & Pierrehumbert, J. B. (2014). Learnability and generalisation of Arabic broken plural nouns. *Language, Cognition and Neuroscience, 29*(10), 1268–1282. https://doi.org/10.1080/23273798.2014.899377

De Houwer, A. (2007). Parental language input patterns and children's bilingual use. *Applied Psycholinguistics, 28*(3), 411–424. https://doi.org/10.1017/S0142716407070221

De Houwer, A. (2009). *Bilingual first language acquisition*. Multilingual Matters.

De Houwer, A. (2018). The role of language input environments for language outcomes and language acquisition in young bilingual children. In D. Miller, F. Bayram, J. Rothman, & L. Serratrice (Eds.), *Bilingual cognition and language: The state of the science across its subfields* (pp. 127–153). John Benjamins Publishing Company. https://doi.org/10.1075/sibil.54.07hou

De Houwer, A. (2023). The danger of bilingual–monolingual comparisons in applied psycholinguistic research. *Applied Psycholinguistics, 44*(3), 343–357. https://doi.org/10.1017/S014271642200042X

De Houwer, A. (2024a). Language Development Milestones for Bilingual and Monolingual Children. *Babylonia Journal of Language Education*, *3*, 14–17. https://doi.org/10.55393/babylonia.v3i.505

De Houwer, A. (2024b). Early Bilingualism. In The Encyclopedia of Applied Linguistics (Bilingualism and Multilingualism), C.A. Chapelle & D. Gabryś-Barker (Eds.). https://doi.org/10.1002/9781405198431.!wbeal0351.pub3.

de León, L. (2012). Language socialization and multi-party participation frameworks. In A. Duranti, E. Ochs, & B. B. Schieffelin (Eds.), *The handbook of language socialization* (pp. 81–111). Wiley-Blackwell.de León, L. (2012). Language socialization and multi-party participation frameworks. In A. Duranti, E. Ochs, & B. B. Schieffelin (Eds.), *The handbook of language socialization* (pp. 81–111). Wiley-Blackwell.

de Villiers, J. (1985). Learning how to use verbs: Lexical coding and the influence of the input. *Journal of Child Language*, 12(3), 587–595.

de Villiers, J. G., & de Villiers, P. A. (1978). *Language acquisition.* Harvard University Press.

de Villiers, J. G., & de Villiers, P. (1973). A cross-sectional study of the acquisition of grammatical morphemes. *Journal of Psycholinguistic Research, 2*, 267–278.

Debattista, M. (2024). *Does childcare attendance influence lexical skills of young Maltese children?* [Bachelor's dissertation, University of Malta]. University of Malta.

Debono, G. (2006). *The proportion of nouns and verbs in two and three year old Maltese-speaking children* [Bachelor's dissertation, University of Malta]. University of Malta.

Dediu, D., & Graham, S. (2014). Genetics and language. *Linguistics.* https://doi.org/10.1093/obo/9780199772810-0184

Demuth, K., & Weschler, S. (2012). The acquisition of Sesotho nominal agreement. *Morphology (Dordrecht)*, *22*(1), 67–88. https://doi.org/10.1007/s11525-011-9192-7

Desira, G. (2002). *Inflection and agreement in child language: A case study* [Bachelor's dissertation, University of Malta]. University of Malta.

Diehm, E. A., Wood, C., Puhlman, J., & Callendar, M. (2020). Young children's narrative retell in response to static and animated stories. *International Journal of Language & Communication Disorders*, *55*(3), 359–372. https://doi.org/10.1111/1460-6984.12523

Diepeveen, F., Dommelen, P., Oudesluys-Murphy, A., & Verkerk, P. (2018). Children with specific language impairment are more likely to reach motor milestones late. *Child: Care, Health & Development, 44*(6), 857-862.

Diessel, H., & Tomasello, M. (2005). A new look at the acquisition of relative clauses. *Language*, 81(4), 882–906.

Długosz, K., & Brown-Bousfield, M. (2025). The acquisition of grammatical gender in German as an additional language: A comparison of L1 English and L1 Polish learners with and without knowledge of other gendered languages. *Lingua, 324*, 103989. https://doi.org/10.1016/j.lingua.2025.103989

Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative and mixed methodologies.* Oxford University Press.

Dressler, W. U. (2005). Morphological typology and first language acquisition: Some mutual challenges. In G. Booij et al. (Eds.), *Morphology and linguistic typology (MMM4)* (pp. 7–20). University of Bologna.

Dressler, W. U. (2011). The rise of complexity in inflectional morphology. *Papers and Studies in Contemporary Linguistics, 47*, 159–176.

Dressler, W. U., Sommer-Lolei, S., Korecky-Kröll, K., et al. (2019). First-language acquisition of synthetic compounds in Estonian, Finnish, German, Greek, Lithuanian, Russian and Saami. *Morphology, 29*, 409–429. https://doi.org/10.1007/s11525-019-09339-0

Dressler, W. U., & Karpf, A. (1995). The theoretical relevance of pre- and protomorphology in language acquisition. *Yearbook of Morphology, 1994*, 99–122.

Dromi, E. (1987). *Early lexical development*. Cambridge: Cambridge University Press.

Dromi, E., & Berman, R. A. (1982). A morphemic measure of early language development: Data from modern Hebrew. *Journal of Child Language, 9*(2), 403–424. https://doi.org/10.1017/S0305000900004785

Eisenbeiß, S. (2006). Documenting child language. In P. K. Austin (Ed.), *Language documentation and description* (Vol. 3, pp. 106–140). SOAS, The Hans Rausing Endangered Languages Project.

Eisenbeiß, S. (2009). Contrast is the Name of the Game: Contrast-Based Semi-Structured Elicitation Techniques for Studies on Children's Language Acquisition. *Essex Research Reports in Linguistics, 57*(7), 1-29. Dept. of Language and Linguistics, University of Essex, Colchester, UK.

Eisenbeiß, S. (2010). Production methods in language acquisition research. In E. Blom & S. Unsworth (Eds.), *Experimental methods in language acquisition research* (pp. 11–34). John Benjamins. https://doi.org/10.1075/lllt.27.03eis

Eisenbeiß, S. (2015). Syntax and language acquisition. In T. Kiss & A. Alexiadou (Eds.), *Syntax – theory and analysis* (2nd ed., Vol. 3, pp. 1792–1833). De Gruyter Mouton.

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.

Engelmann, F., Granlund, S., Kolak, J., Szreder, M., Ambridge, B., Pine, J., Theakston, A., & Lieven, E. (2019). How the input shapes the acquisition of verb morphology: Elicited production and computational modelling in two highly inflected languages. *Cognitive Psychology, 110*, 30–69. https://doi.org/10.1016/j.cogpsych.2019.02.001

Ensminger, M. E., & Fothergill, K. (2014). A decade of measuring SES: What it tells us and where to go from here. In M. H. Bornstein & R. H. Bradley (Eds.), *Socioeconomic status, parenting, and child development* (pp. 13–27). Erlbaum.

Eriksson, M., Marschik, P. B., Tulviste, T., Almgren, M., Pérez Pereira, M., Wehberg, S., Marjanovič-Umek, L., Križnar, Z. G., & Gallego, C. (2012). Differences between girls and boys in emerging language skills: Evidence from 10 language communities. *British Journal of Developmental Psychology*, *30*(2), 326–343. https://doi.org/10.1111/j.2044-835X.2011.02042.x

Ervin, S. M., & Miller, W. R. (1963). Language development. In H. W. Stevenson (Ed.), *Child psychology: The sixty-second yearbook of the National Society for the Study of Education* (pp. 108–143). University of Chicago Press.

Estigarribia, B., & Clark, E. V. (2007). Getting and maintaining attention in talk to young children. *Journal of Child Language, 34*(4), 799–814.

European Commission. (2024). *Europeans and their languages: Special Eurobarometer SP540*. Retrieved from https://europa.eu/eurobarometer/surveys/detail/2979

Eurostat. (2019). *People on the move: Statistics on mobility in Europe.* Directorate-General of the European Commission. https://ec.europa.eu/eurostat/cache/digpub/eumove_2019

Eurostat. (2024). *Children aged less than 3 years in formal childcare by duration* [Data set]. Retrieved from https://ec.europa.eu/eurostat/databrowser/view/tepsr_sp210/default/table?lang=en

Fabri, R. (2009). Stem allomorphy in the Maltese verb. *Ilsienna/Our Language, 1*, 1–20. International Association of Maltese Linguistics (GHILM).

Fabri, R. (2010). Maltese. In C. Delcourt & P. van Sterkenburg (Eds.), *The languages of the new EU member states. Revue Belge de Philologie et d'Histoire, 88*(3), 791–816.

Fabri, R. (2014). Maltese. In I. Zitouni (Ed.), *Natural language processing of Semitic* (pp. 26–31). Springer.

Fabri, R. (2015). The island adventures of Robinson Crusoe and Man Friday: The life and strange surprising island adventures of the Maltese language. In R. Heimrath, & K. Arndt (Eds.), *Insularity: Small Worlds in Linguistic, Social and Cultural Perspectives* (pp. 45–61). Würzburg: Königshausen & Neumann.

Fabri, R. (2001). *Definiteness marking and the structure of the NP in Maltese*. Verbum, 23(2), 153–172.

Fabri, R. (2023). Grammatical agreement in Maltese. Journal of Maltese Studies, 30, 77-110.

Farrugia, G. (2015). Il-ġens grammatikali fil-Malti f'kuntest tipoloġiku u psikolingwistiku [Grammatical gender in Maltese in a typological and psycholinguistic context]. *Il-Malti, LXXXVIII*, 24–44.

Farrugia, G. (2018). *Grammatical gender in Maltese*. De Gruyter Mouton.

Farrugia, J. (2024). *The effect of the home language learning environment on lexical and narrative skills of 5-year-old children* [Undergraduate dissertation, University of Malta]. University of Malta.

Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J. P., Pethick, S. J., & Reilly, J. S. (1993). *The MacArthur Communicative Development Inventories: User's guide and technical manual*. Singular.

Fenson, L., Marchman, V. A., Thal, D., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *The MacArthur Communicative Development Inventories: User's guide and technical manual* (2nd ed.). Brookes.

Ferguson, C. A. (1959). Diglossia. *Word, 15*(2), 325–340. https://doi.org/10.1080/00437956.1959.11659702

Ferguson, C. A. (1964). Baby talk in six languages. *American Anthropologist, New Series, 66*, Part 2: The ethnography of communication, 103–114.

Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science, 16*(2), 234–248. https://doi.org/10.1111/desc.12019

Ferry, A. L., Corcoran, M. G., Williams, E., Curtis, S. M., Gale, C. J., & Twomey, K. E. (2025). Bigger versus smaller: Children's understanding of size comparison words becomes more precise with age. *Child Development, 96*(2), 492–507. https://doi.org/10.1111/cdev.14182

Fletcher, P., & MacWhinney, B. (1995). *The Handbook of Child Language*. Blackwell.

Flewitt, R., El Gemayel, S., Arnott, L., Gillen, J., Goodall, J., Winter, K., Dalziell, A., Liu, M., Savadova, S., & Timmins, S. (2024). Toddlers, tech and talk: Very young children's language and literacy learning at home in a post-digital age. Summary report. Manchester Metropolitan University.

Fodor, J. (2001). Doing Without What Within: Fiona Cowie's Critique of nativism. *Mind. A Quarterly Review of Psychology and Philosophy*, *110*, 94–148.

Forshaw, W. (2021). The acquisition of complex morphology: Insights from Murrinhpatha. John Benjamins.

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2021). *Variability and consistency in early language learning: The Wordbank project*. MIT Press.

Frank, M. C., Goodman, N. D., & Tenenbaum, J. B. (2009). Using speakers' referential intentions to model early cross-situational word learning. *Psychological Science, 20*(5), 578–585. https://doi.org/10.1111/j.1467-9280.2009.02335.x

French, R. M., Addyman, C., & Mareschal, D. (2011). TRACX: A recognition-based connectionist framework for sequence segmentation and chunk extraction. *Psychological Review*, *118*(4), 614–636. https://doi.org/10.1037/a0025255

Frendo Wirth, L. (2020). *Investigating the use of a novel multimodal device in lexical assessment and intervention for 3- to 6-year-old bilingual Maltese children* [Master's dissertation, University of Malta]. University of Malta.

Friederici, A. D. (2005). Neurophysiological markers of early language acquisition: From syllables to sentences. *Trends in Cognitive Sciences, 9*(10), 481–488. https://doi.org/10.1016/j.tics.2005.08.008

Friederici A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in cognitive sciences*, *6*(2), 78–84. https://doi.org/10.1016/s1364-6613(00)01839-8

Friedrich, M., & Friederici, A. D. (2005). Phonotactic knowledge and lexical-semantic processing in one-year-olds: Brain responses to words and nonsense words in picture contexts. *Journal of Cognitive Neuroscience, 17*(11), 1785–1802.

Frizelle, P., Thompson, P. A., McDonald, D., & Bishop, D. V. M. (2018). Growth in syntactic complexity between four years and adulthood: Evidence from a narrative task. *Journal of Child Language, 45*(5), 1174–1197. https://doi.org/10.1017/S0305000918000144

Fromkin, V., & Rodman, R. (2011). *An introduction to language* (9th ed.). Harcourt Brace College.

Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Balčiūnienė, I., Bohnacker, U., & Walters, J. (2012). MAIN: Multilingual Assessment Instrument for Narratives. *ZAS Papers in Linguistics, 56*, 1–140.

Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Bohnacker, U., & Walters, J. (2020). MAIN: Multilingual Assessment Instrument for Narratives. Revised version – Materials for use. *ZAS Papers in Linguistics, 63*. [Maltese version translated and adapted by D. Gatt]

Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Bohnacker, U., & Walters, J. (2019). MAIN: Multilingual Assessment Instrument for Narratives. Revised version. *ZAS Papers in Linguistics, 63*, 1–36.

Gallaway, C., & Richards, B. (Eds.). (1994). *Input and interaction in language acquisition.* Cambridge University Press.

Gardner-Neblett, N. (2024). Becoming fictional storytellers: African American children's oral narrative development in early elementary school. *Child Development, 95*(4), 1218–1236. https://doi.org/10.1111/cdev.14075

Gaskins, D., Quick, A. E., Verschik, A., & Backus, A. (2022). Usage-based approaches to child code-switching: State of the art and ways forward. *Cognitive Development, 64*, 101269. https://doi.org/10.1016/j.cogdev.2022.101269

Gaskins, D., Frick, M., Palola, E., & Endesfelder Quick, A. (2019). Towards a usage-based model of early code-switching: Evidence from three language pairs. *Applied Linguistics Review, 12*(2), 179–206.

Gatt, A., & Čéplö, S. (2013). Digital corpora and other electronic resources for Maltese. In *Proceedings of the International Conference on Corpus Linguistics* (pp. 96–97). UCREL.

Gatt, D. (2010). *Early expressive lexical development: Evidence from children brought up in Maltese-speaking families* [Doctoral dissertation, University of Malta]. University of Malta.

Gatt, D. (2017). Bilingual vocabulary production in young children receiving Maltese-dominant exposure: Individual differences and the influence of demographic and language exposure factors. *International Journal of Bilingual Education and Bilingualism, 20*(2), 163–182. https://doi.org/10.1080/13670050.2016.1179255

Gatt, D., & Borg Cutajar, R. (2023). The adaptation of MAIN to Maltese. *ZAS Papers in Linguistics, 65*, 61–71.

Gatt, D., Camilleri, L., & Grech, C. (2023). Word usage as measured by parental checklists and language samples: Trends, comparisons, and implications. *Frontiers in Psychology, 14*, 1214518. https://doi.org/10.3389/fpsyg.2023.1214518

Gatt, D., & Dodd, B. (2020). Preschoolers' lexical skills in two majority languages: Is there evidence for the onset of sequential bilingualism? *International Journal of Bilingualism, 24*(2), 222–245. https://doi.org/10.1177/1367006919826408

Genesee, F. (2015). Myths about early childhood bilingualism. *Canadian Psychology/Psychologie canadienne, 56*(1), 6–15. https://doi.org/10.1037/a0038599

Genesee, F. (2022). The monolingual bias: A critical analysis. *Journal of Immersion and Content-Based Language Education, 10(2), 153–181.* https://doi.org/10.1075/jicb.21016.gen

Genesee, F., & Nicoladis, E. (2007). Bilingual first language acquisition. In E. Hoff & M. Shatz (Eds.), *Handbook of language development* (pp. 324–342). Blackwell.

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. A. Kuczaj (Ed.), *Language development: Vol. 2. Language, thought, and culture* (pp. 301–334). Lawrence Erlbaum Associates.

Gentner, D., & Boroditsky, L. (2001). Individuation, relativity, and early word learning. In M. Bowerman & S. C. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 215–256). Cambridge University Press.

Gervain, J. (2020). Typical language development. *Handbook of Clinical Neurology, 173*, 171–183. https://doi.org/10.1016/B978-0-444-64150-2.00016-2

Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.

Goldin-Meadow, S. (2015). From gesture to word. In E. L. Bavin & L. R. Naigles (Eds.), *The Cambridge handbook of child language* (pp. 183–204). Cambridge University Press.

Goldin-Meadow, S., Levine, S. C., Hedges, L. V., Huttenlocher, J., Raudenbush, S. W., & Small, S. L. (2014). New evidence about language and cognitive development based on a longitudinal study: Hypotheses for intervention. *American Psychologist, 69*(6), 588–599. https://doi.org/10.1037/a0036886

Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language, 14*(1), 23–45. https://doi.org/10.1017/S030500090001271X

Golinkoff, R. M., Mervis, C. B., & Hirsh-Pasek, K. (1994). Early object labels: The case for a developmental lexical principles framework. *Journal of Child Language, 21*(1), 125–155. https://doi.org/10.1017/S0305000900008692

Gómez, R. L., Bootzin, R. R., & Nadel, L. (2006). Naps promote abstraction in language-learning infants. *Psychological Science, 17*(8), 670–674. https://doi.org/10.1111/j.1467-9280.2006.01764.x

Goodluck, H. (2020). *Language acquisition by children: A linguistic introduction*. Edinburgh University Press.

Graf Estes, K., Edwards, J., & Saffran, J. R. (2011). Phonotactic Constraints on Infant Word Learning. *Infancy*, *16*(2), 180–197. https://doi.org/10.1111/j.1532-7078.2010.00046.x

Granlund, S., Kolak, J., Vihman, V., Engelmann, F., Lieven, E. V. M., Pine, J. M., Theakston, A. L., & Ambridge, B. (2019). Language-general and language-specific phenomena in the acquisition of inflectional noun morphology: A cross-linguistic elicited-production study of Polish, Finnish and Estonian. *Journal of Memory and Language, 107*, 169–194. https://doi.org/10.1016/j.jml.2019.04.004

Grech, H., & Dodd, B. (2008). Phonological acquisition in Malta: A bilingual language learning context. *International Journal of Bilingualism, 12*(3), 155–171. https://doi.org/10.1177/1367006908098564

Grech, S. (2015) *Variation in English: Perception and patterns in the identification of Maltese English*. Unpublished doctoral dissertation, Institute of Linguistics. Malta: University of Malta.

Grech, S. (2024). The development of Maltese English. *World Englishes, 43*, 270–283. https://doi.org/10.1111/weng.12645

Grosjean, F. (1997). The bilingual individual. *Interpreting*, *2*(1–2), 163–187.

Grosjean, F. (2010). *Bilingual: Life and reality*. Harvard University Press.

Grosjean, F. (2012). An attempt to isolate, and then differentiate, transfer and interference. *International Journal of Bilingualism*, *16*(1), 11–21.

Guasti, M. T. (2002). *Language acquisition: The growth of grammar*. MIT Press.

Gutova, E. (2016). Baby talk in the Maghreb. In G. Grigore & G. Bițună (Eds.), *Arabic Varieties: Far and Wide. Proceedings of the 11th International Conference of AIDA – Bucharest 2015* (pp. 282–290). Bucharest: Editura Universității din București.

Hadley, P. A., Rispoli, M., Fitzgerald, C., & Bahnsen, A. (2011). Predictors of morphosyntactic growth in typically developing toddlers: Contributions of parent input and child sex. *Journal of Speech, Language, and Hearing Research, 54*(2), 549–566. https://doi.org/10.1044/1092-4388(2010/09-0216)

Hamre, B. K., & Pianta, R. C. (2007). Learning opportunities in preschool and early elementary classrooms. In R. C. Pianta, M. J. Cox, & K. Snow (Eds.), *School readiness and the transition to school* (pp. 49–84). Baltimore, MD: Brookes.

Hashoul-Essa, L., & Armon-Lotem, S. (2025). Gender effects in lexical and morphosyntactic acquisition of Arabic: A CDI study. *First Language*, *45*(3), 231–259. https://doi.org/10.1177/01427237251329971

Haspelmath, M. (2020). The morph as a minimal linguistic form. *Morphology, 30*(2), 117–134. https://doi.org/10.1007/s11525-020-09355-5

Haspelmath, M., & Sims, A. (2010). *Understanding morphology* (2nd ed.). Hodder Education.

He, A. X. (2022). Optimal input for language development: Tailor nurture to nature. *Infant and Child Development, 31*(1), e2269. https://doi.org/10.1002/icd.2269

Hellwig, B., Allen, S. E. M., Davidson, L., Defina, R., Kelly, B. F., & Kidd, E. (Eds.). (2023). *The acquisition sketch project* (Language Documentation & Conservation Special Publication No. 28). University of Hawaiʻi Press.

Henry, G. T., & Rickman, D. K. (2007). Do peers influence children's skill development in preschool? *Economics of Education Review, 26*, 100–112. https://doi.org/10.1016/j.econedurev.2005.09.006

Hockett, C. F. (1958). *A course in modern linguistics*. Macmillan.

Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review, 26*(1), 55–88. https://doi.org/10.1016/j.dr.2005.11.002

Hoff, E. (2013). *Language development* (5th ed.). Wadsworth/Cengage Learning.

Hoff, E. (2015). Language development in bilingual children. In E. L. Bavin & L. R. Naigles (Eds.), *The Cambridge Handbook of Child Language* (pp. 483–503). chapter, Cambridge: Cambridge University Press.

Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language, 39*(1), 1–27. https://doi.org/10.1017/S0305000910000759

Höhle, B., Bijeljac-Babic, R., & Nazzi, T. (2020). Variability and stability in early language acquisition: Comparing monolingual and bilingual infants' speech perception and word recognition. *Bilingualism: Language and Cognition, 23*, 56–71. https://doi.org/10.1017/S1366728919000348

Höhle, B. (2009). Bootstrapping mechanisms in first language acquisition. *Linguistics*, *47*(2), 359-382. https://doi.org/10.1515/LING.2009.013

Horváth, K., Liu, S., & Plunkett, K. (2016). A daytime nap facilitates generalization of word meanings in young toddlers. *Sleep (New York, N.Y.), 39*(1), 203–207. https://10.5665/sleep.5348

Howe, N., Della Porta, S., Recchia, H., & Ross, H. (2016). "Because if you don't put the top on, it will spill": A longitudinal study of sibling teaching in early childhood. *Developmental Psychology, 52*(11), 1832–1842. https://doi.org/10.1037/dev0000193

Howe, N., Paine, A. L., Recchia, H., & Ross, H. (2022). Sibling relations in early childhood. In C. Hart & P. K. Smith (Eds.), *Wiley-Blackwell handbook of childhood social development* (3rd ed., pp. 443–458). Wiley.

Hržica, G. (2012). Daj mi to napisaj: preopćavanja glagolske osnove u usvajanju hrvatskog jezika (Overgeneralisations of the verb stem in the acquisition of Croatian language). *Suvremena Lingvistika*, *38*(74), 189–208.

Hsu, N., Hadley, P. A., & Rispoli, M. (2017). Diversity matters: Parent input predicts toddler verb production. *Journal of Child Language, 44*(1), 63–86.

Hutton, J. S., Dudley, J., Horowitz-Kraus, T., DeWitt, T., & Holland, S. K. (2019). Associations between screen-based media use and brain white matter integrity in preschool-aged children. *JAMA Pediatrics, 174*(1), e193869. https://doi.org/10.1001/jamapediatrics.2019.3869

Ibbotson, P. (2020). *What it takes to talk*. De Gruyter Mouton.

Ibbotson, P., & Browne, W. J. (2024). The effects of family, culture and sex on linguistic development across 20 languages. *Developmental Science*, *27*(6), e13547-n/a. https://doi.org/10.1111/desc.13547

Jachimek, A., Dimroth, C., & Köpcke, K.-M. (2022). The acquisition of the NP in a German–Polish bilingual child: Evidence for cross-linguistic influence. *Discours, 29*. https://doi.org/10.4000/discours.11800

Jakobson, R. (1968). *Child language, aphasia, and phonological universals.* Mouton. (Original work published 1941)

Joanisse, M. F., & McClelland, J. L. (2015). Connectionist perspectives on language learning, representation, and processing. *Wiley Interdisciplinary Reviews: Cognitive Science, 6*(3), 235–247. https://doi.org/10.1002/wcs.1340

Johnston J. R. (2001). An alternate MLU calculation: magnitude and variability of effects. *Journal of speech, language, and hearing research: JSLHR*, *44*(1), 156–164. https://doi.org/10.1044/1092-4388(2001/014)

Kalashnikova, M., Carignan, C., & Burnham, D. (2017). The origins of babytalk: Smiling, teaching or social convergence? *Royal Society Open Science, 4*(8), 170306. https://doi.org/10.1098/rsos.170306

Kapantzoglou, M., Brown, J. E., Cycyk, L. M., & Fergadiotis, G. (2021). Code-switching and language proficiency in bilingual children with and without developmental language disorder. *Journal of Speech, Language, and Hearing Research, 64*(5), 1605–1620. https://doi.org/10.1044/2020_JSLHR-20-00182

Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences, 2*(10), 389–398. https://10.1016/S1364-6613(98)01230-3

Katsos, N., Cummins, C., Ezeizabarrena, M.-J., Gavarró, A., Kuvač Kraljević, J., Hrzica, G., Grohmann, K. K., Skordi, A., Jensen de López, K., Sundahl, L., van Hout, A., Hollebrandse, B., Overweg, J., Faber, M., van Koert, M., Smith, N., Vija, M., Zupping, S., Kunnari, S., … Noveck, I. (2016). Cross-linguistic patterns in the acquisition of quantifiers. *Proceedings of the National Academy of Sciences of the United States of America, 113*(33), 9244–9249. https://doi.org/10.1073/pnas.1601341113

Kauschke, C., Lawatsch, K., Tenhagen, A., & Dörfler, T. (2024). Exploring grammatical development in children aged 2;6 to 7: A novel approach using elicited production. *Clinical Linguistics & Phonetics*, Advance online publication, 1–24. https://doi.org/10.1080/02699206.2024.2328200

Kauschke, C., Kurth, A., & Domahs, U. (2011). Acquisition of German noun plurals in typically developing children and children with specific language impairment. *International Journal of Speech-Language Pathology, 13*(3), 255–265. https://doi.org/10.3109/17549507.2011.564438

Kaushanskaya, M., & Crespo, K. (2019). Does Exposure to Code-Switching Influence Language Performance in Bilingual Children? *Child Development*, *90*(3), 708–718. https://doi.org/10.1111/cdev.13235

Kedar, Y. & College, B. B. (2019). Innateness of language. In *Oxford Research Encyclopedia of Linguistics*. Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.86

Kellerman, E. (1983). *Now you see it, now you don't*. In S. Gass & L. Selinker (Eds.), Language transfer in language learning (pp. 112–134). Newbury House.

Kelly, B. F., Forshaw, W., Nordlinger, R., & Wigglesworth, G. (2015). Linguistic diversity in first language acquisition research: Moving beyond the challenges. *First Language, 35*(4–5), 286–304. https://doi.org/10.1177/0142723715602350

Kenanidis, P., Chondrogianni, V., Legendre, G., & Culbertson, J. (2021). Cue reliability, salience and early comprehension of agreement: Evidence from Greek. *Journal of Child Language*, *48*(4), 815–833. https://doi.org/10.1017/S0305000920000628

Kern, S., Gayraud, F., & Chenu, F. (2014). The role of input in early first language morphosyntactic development. *Language, Interaction and Acquisition, 5*(1), 1–18.

Keydeniers, D., Aalberse, S. P., Andringa, S., & Kuiken, F. (2023). Acquiring the Dutch plural in a bilingual setting: Investigating the effects of language dominance, overlap, and variability. *Languages, 8*(3), 156. https://doi.org/10.3390/languages8030156

Khan, K. S., Gugiu, M. R., Justice, L. M., Bowles, R. P., Skibbe, L. E., & Piasta, S. B. (2016). Age-related progressions in story structure in young children's narratives. *Journal of Speech, Language, and Hearing Research*, *59*(6), 1395–1408. https://doi.org/10.1044/2016_JSLHR-L-15-0275

Khattab, G., & Keren-Portnoy, T. (2025). How templatic is Arabic input to children? The role of child-directed speech in the acquisition of Semitic morphophonology. *Language and Speech*, Advance online publication. https://doi.org/10.1177/00238309241311230

Kidd, E., & Arciuli, J. (2016). Individual Differences in Statistical Learning Predict Children's Comprehension of Syntax. *Child development*, *87*(1), 184–193. https://doi.org/10.1111/cdev.12461

Kidd, E., & Garcia, R. (2022). How diverse is first language acquisition research? *First Language*, Advance online publication, 1–33. https://doi.org/10.1177/01427237211066405

Kidd, E., & Donnelly, S. (2020). Individual differences in first language acquisition. *Annual Review of Linguistics, 6*(1), 319–340. https://doi.org/10.1146/annurev-linguistics-011619-030326

Kids Count Data Center. (2023). *Children who speak a language other than English at home in United States.* Annie E. Casey Foundation. https://datacenter.aecf.org/data/tables/81-children-who-speak-a-language-other-than-english-at-home

Kim, S. (2014). The effects of using storytelling with a Charlie Chaplin silent movie clip on oral proficiency development. *STEM Journal, 15*(1), 1–20.

Kim, Y. J., & Sundara, M. (2021). 6–month–olds are sensitive to English morphology. *Developmental Science*, *24*(4), e13089-n/a. https://doi.org/10.1111/desc.13089

Klinzing, J. G., Niethard, N., & Born, J. (2019). Mechanisms of systems memory consolidation during sleep. *Nature Neuroscience, 22*(10), 1598–1610. https://doi.org/10.1038/s41593-019-0467-

Kovačević, M., Palmović, M., & Hržica, G. (2009). The acquisition of case, number, and gender in Croatian. In U. Stephany & M. D. Voeikova (Eds.), *Development of nominal inflection in first language acquisition: A cross-linguistic perspective* (pp. 153–177). De Gruyter Mouton.

Kover, S. T., Edmunds, S. R., & Ellis Weismer, S. (2016). Ages of Language Milestones as Predictors of Developmental Trajectories in Young Children with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *46*(7), 2501–2507. https://doi.org/10.1007/s10803-016-2756-y

Krashen, S. (1982). *Second language acquisition and second language learning.* Pergamon Press.

Kremin, L. V., Alves, J., Orena, A. J., Polka, L., & Byers-Heinlein, K. (2022). Code-switching in parents' everyday speech to bilingual infants. *Journal of child language*, *49*(4), 714–740. https://doi.org/10.1017/S0305000921000118

Kuhl, P. K. (2010). Brain mechanisms in early language acquisition. *Neuron, 67*(5), 713-727.

Kupisch, T. (2007). Determiners in bilingual German–Italian children: What they tell us about the relation between language influence and language dominance. *Bilingualism: Language and Cognition, 10*(1), 57–78. https://doi.org/10.1017/S1366728906002823

Kuvač-Kraljević, J., Blaži, A., Schults, A., Tulviste, T., & Stolt, S. (2021). Influence of internal and external factors on early language skills: A cross-linguistic study. *Infant Behavior & Development, 63*, 101552. https://doi.org/10.1016/j.infbeh.2021.101552

Kwon, J., & Martínez-Álvarez, P. (2021). A young linguistic and cultural mediator: a case of trilingual siblings' interaction. *International Multilingual Research Journal*, *16*(1), 47–64. https://doi.org/10.1080/19313152.2021.1930647

Laaha, S., Ravid, D., Korecky-Kröll, K., Laaha, G., & Dressler, W. U. (2006). Early noun plurals in German: regularity, productivity or default? *Journal of Child Language*, *33*(2), 271–302. https://doi.org/10.1017/S0305000906007379

Laalo, K., & Argus, R. (2020). Linguistic recycling in language acquisition: Child-directed speech and child speech in the study of language acquisition. *AILA Review, 33*(1), 86–103. https://doi.org/10.1075/aila.00031.laa

L-Aġenzija Nazzjonali tal-Litteriżmu. (2021). *Promoting Maltese at home | Bilingual children* [Video]. YouTube. https://www.youtube.com/watch?v=dqGwLzgEeC4

Langacker, R. W. (1987). *Foundations of cognitive grammar. Vol. 1: Theoretical prerequisites.* Stanford University Press.

Langacker, R. W. (2000). A dynamic usage-based model. In M. Barlow & S. Kemmer (Eds.), *Usage-based models of language* (pp. 1–63). CSLI Publications.

Lange, B. P., Euler, H. A., & Zaretsky, E. (2016). Sex differences in language competence of three- to six-year-old children. *Applied Psycholinguistics*, *37*(6), 1417–1438. https://doi.org/10.1017/S0142716415000624

Lanza, E., & Wei, L. (2023). Code-switching and children. In J. L. Jackson Jr. (Ed.), *Oxford bibliographies in childhood studies*. Oxford University Press. https://doi.org/10.1093/obo/9780199791231-0271

Lenneberg, E. H. (1967). *Biological foundations of language.* Wiley.

Leonard, L. B., Salameh, E.-K., & Hansson, K. (2001). Noun phrase morphology in Swedish-speaking children with specific language impairment. *Applied Psycholinguistics*, *22*(4), 619–639. https://doi.org/10.1017/S0142716401004076

Lester, N. A., Moran, S., Küntay, A. C., Allen, S. E. M., Pfeiler, B., & Stoll, S. (2022). Detecting structured repetition in child-surrounding speech: Evidence from maximally diverse languages. *Cognition, 221*, 104986.

Levelt, W. J. M. (2013). *A history of psycholinguistics: The pre-Chomskyan era*. Oxford: Oxford University Press.

Levy, Y. (1983). The acquisition of Hebrew plurals: The case of the missing gender category. *Journal of Child Language, 10*(1), 107–121. https://doi.org/10.1017/S0305000900005171

Li, P. (2009). Lexical organization and competition in first and second languages: Computational and neural mechanisms. *Cognitive Science, 33*(4), 629–664. https://doi.org/10.1111/j.1551-6709.2009.01028.x

Lieber, R. (Ed.). (2020). *The Oxford encyclopedia of morphology.* Oxford University Press.

Lieven, E. (1994). Variation in crosslinguistic context. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition* (Vol. 5, pp. 219–263). Lawrence Erlbaum.

Lieven, E. (2010). Input and first language acquisition: Evaluating the role of frequency. *Lingua*, *120*(11), 2546–2556. https://doi.org/10.1016/j.lingua.2010.06.005

Lieven, E., & Brandt, S. (2011). The constructivist approach. *Infancia y Aprendizaje, 34*(3), 281–296. https://doi.org/10.1174/021037011797238586

Lipsky, M. G. (2013). Head Start teachers' vocabulary instruction and language complexity during storybook reading: Predicting vocabulary outcomes of students in linguistically diverse classrooms. *Early Education and Development, 24*, 640–667.

Liu, H., & MacWhinney, B. (2024). Morphosyntactic analysis for CHILDES. *Language Development Research, 2*(1).

Lohndal, T., & Westergaard, M. (2021). Grammatical Gender: Acquisition, Attrition, and Change. *Journal of Germanic Linguistics*, *33*(1), 95–121. https://doi.org/10.1017/S1470542720000057

Long, H. L., Ramsay, G., Griebel, U., Bene, E. R., Bowman, D. D., Burkhardt-Reed, M. M., & Oller, D. K. (2022). Perspectives on the origin of language: Infants vocalize most during independent vocal play but produce their most speech-like vocalizations during turn taking. *PloS One*, *17*(12), e0279395–e0279395. https://doi.org/10.1371/journal.pone.0279395

Lorang, E., Venker, C. E., & Sterling, A. (2020). An investigation into maternal use of telegraphic input to children with down syndrome. *Journal of Child Language*, *47*(1), 225–249. https://doi.org/10.1017/S0305000919000503

Lukács, A., Leonard, L. B., & Kas, B. (2010). Use of noun morphology by children with language impairment: the case of Hungarian. *International journal of language & communication disorders*, *45*(2), 145–161. https://doi.org/10.3109/13682820902781060

Lust, B. (2006). *Child language: Acquisition and growth.* Cambridge: Cambridge University Press.

Lytle, S. R., & Kuhl, P. K. (2017). Social interaction and language acquisition: Toward a neurobiological view. In E. M. Fernández, & H. S. Cairns (Eds.). The Handbook of Psycholinguistics (pp. 615-634). Wiley. https://doi.org/10.1002/9781118829516.ch27

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Psychology Press.

MacWhinney, B. (2005). *A unified model of language acquisition*. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 49–67). Oxford University Press.

MacWhinney, B. (2008). A unified model. In P. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 341–371). Routledge/Taylor & Francis Group.

MacWhinney, B., & Bates, E. (1989). *The crosslinguistic study of sentence processing.* Cambridge University Press.

Majorano, M., Vihman, M. M., & DePaolis, R. A. (2014). The relationship between infants' production experience and their processing of speech. *Language Learning and Development, 10*(2), 179–204. https://doi.org/10.1080/15475441.2013.829740

Malvern, D. D., Richards, B. J., Chipere, N., & Durán, P. (2004). *Lexical diversity and language development: Quantification and assessment*. Palgrave Macmillan.

Maneva, B., & Genesee, F. (2001). *Bilingual babbling: Evidence for language differentiation in bilingual first language acquisition* [Conference presentation]. Boston University Conference on Language Development.

Marchman, V. A. (1997). Children's productivity in the English past tense: The role of frequency, phonology and neighborhood structure. *Cognitive Science, 21*, 283–304.

Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J., & Xu, F. (1992). Overregularization in language acquisition. *Monographs of the Society for Research in Child Development*, *57*(4), 1–182.

Marinaccio, M. (2024). *Getting to the Root of the Maltese Broken Plural.* [Master's dissertation, University of Malta]. University of Malta.

Mariscal, S. (2009). Early acquisition of gender agreement in the Spanish noun phrase: starting small. *Journal of Child Language*, *36*(1), 143–171. https://doi.org/10.1017/S0305000908008908

Mashaqba, B., Al-Khawaldeh, N., AlGweirien, H. & Al-Edwan, Y. (2020). Acquisition of broken plural patterns by Jordanian children. *Linguistics*, *58*(4), 1009-1022. https://doi.org/10.1515/ling-2020-0024

Maslen, R. J. C., Theakston, A. L., Lieven, E. V. M., & Tomasello, M. (2004). A dense corpus study of past tense and plural overregularization in English. *Journal of Speech, Language, and Hearing Research, 47*(6), 1319–1333. https://doi.org/10.1044/1092-4388(2004/099)

Masrai, A. M. (2016). The influence of morphological knowledge on lexical processing and acquisition: The case of Arab EFL learners. *Ampersand (Oxford, UK)*, *3*, 52–60. https://doi.org/10.1016/j.amper.2016.04.001

Massaroni, V., Delle Donne, V., Marra, C., Arcangeli, V., & Chieffo, D. P. R. (2024). The relationship between language and technology: How screen time affects language development in early life—A systematic review. *Brain Sciences, 14*(1), 27. https://doi.org/10.3390/brainsci14010027

Matsuo, A., Kita, S., Shinya, Y., Wood, G. C., & Naigles, L. (2012). Japanese two-year-olds use morphosyntax to learn novel verb meanings. *Journal of Child Language, 39*, 637–663.

Mayer, M. (1969). *Frog, where are you?* Dial Books for Young Readers.

Mazara, J., Loporcaro, M., & Stoll, S. (2024). Acquisition of gender agreement depends on frequency distributions in specific contexts. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (pp. 2496–2502). Cognitive Science Society, UC Merced.

McCarthy, J. J., & Prince, A. S. (1990). Foot and word in prosodic morphology: The Arabic broken plural. *Natural Language & Linguistic Theory*, *8*(2), 209–283. https://doi.org/10.1007/BF00208524

McClelland, J. L., & Patterson, K. (2002). Rules or connections in past-tense inflections: What does the evidence rule out? *Trends in Cognitive Sciences, 6*, 465–472.

McDaniel, B. T., & Radesky, J. S. (2018). Technoference: Parent distraction with technology and associations with child behavior problems. *Child Development, 89*(1), 100–109. https://doi.org/10.1111/cdev.12822

McLaughlin, B. (1985). *Second-language acquisition in childhood.* Lawrence Erlbaum Associates.

McLeod, S., & Crowe, K. (2018). Children's Consonant Acquisition in 27 Languages: A Cross-Linguistic Review. *American Journal of Speech-Language Pathology, 27*(4), 1546–1571. https://doi.org/10.1044/2018_AJSLP-17-0100

Meir, N. (2023). Individual differences in bilingual child language acquisition: A plunge into a complex and dynamic network. *Journal of Child Language, 50*, 827–831. https://doi.org/10.1017/S0305000923000016

Meisel, J. M. (1994). Code-switching in young bilingual children: The acquisition of grammatical constraints. *Studies in Second Language Acquisition, 16*(4), 413–439. https://doi.org/10.1017/S0272263100013449

Meisel, J. M. (2011). *First and second language acquisition: Parallels and differences.* Cambridge University Press.

Meyer, C. (2019). *Rule and order: Acquiring ordinals in Dutch and English* [Doctoral dissertation, University of Amsterdam]. LOT Dissertations, 530.

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

Mifsud, M. (1996). The collective in Maltese. *Italian Journal of Linguistics*, *8*(1), 29–52.

Miller, G. A. (2003). The cognitive revolution: a historical perspective. *Trends in Cognitive Sciences 7* (3), 141–144.

Miller, J. F., & Chapman, R. S. (1981). The relation between age and mean length of utterance in morphemes. *Journal of speech and hearing research*, *24*(2), 154–161. https://doi.org/10.1044/jshr.2402.154

Mills, A. (1986). *The acquisition of gender: A study of English and German.* Springer.

Ministry for Education and Employment. (2016). *A language policy for the early years in Malta and Gozo.* Ministry for Education and Employment.

Mishina-Mori, S. (2011). A longitudinal analysis of language choice in bilingual children: The role of parental input and interaction. *Journal of Pragmatics 43*(13), 3122–3138.

Mitterer, H., Cassar, R., Pantelmann, J., & Reinisch, E. (2025). Can a second language trigger sound change in the first? An apparent-time study on rhotic production in Maltese [Preprint]. https://osf.io/nmq57/overview

Moawad, R. (2006). *The acquisition of the Arabic gender and number systems* [Doctoral dissertation, Bangor University]. Bangor University Research Repository.

Moerk, E. L. (1980). Relationships between parental input frequencies and children's language acquisition: A reanalysis of Brown's data. *Journal of Child Language, 7*(1), 105–118. https://doi.org/10.1017/S0305000900007054

Molnar, M., Alemán Bañón, J., Mancini, S., & Caffarra, S. (2021). The processing of Spanish article–noun gender agreement by monolingual and bilingual toddlers. *Language and Speech, 64*(4), 980–990. https://doi.org/10.1177/0023830920977050

Morgan, J. L. (1986). *From simple input to complex grammar.* Cambridge, MA: MIT Press.

Mulford, R. (1985). Comprehension of Icelandic pronoun gender: semantic versus formal factors. *Journal of Child Language*, *12*(2), 443–453. doi:10.1017/S030500090000653X

Murchison, C., & Langer, S. (1927). Tiedemann's observations on the development of the mental faculties of children. *The Pedagogical Seminary and Journal of Genetic Psychology, 34,* 205–230. https://doi.org/10.1080/08856559.1927.10532379

Murray Curtis, J. (1997). *The emergence of the plural in 3- to 5-year-old Maltese-speaking children* [Bachelor's dissertation, University of Malta]. University of Malta.

Muszyńska, K., Krajewski, G., Dynak, A., Garmann, N.G., Romøren, A.S.H., Łuniewska, M., Alcock, K., Katsos, N., Kołak,J., Simonsen, H.G., Hansen,P., Krysztofiak, M., Sobota,K., & Haman,E. (2025). Bilingual children reach early language milestones at the same age as monolingual peers. *Journal of Child Language* 1–24, https://doi.org/10.1017/S0305000924000655

Muscat, D., Mifsud, G., & Coleiro, D. (2017). *Hekk aħjar! [That's better!]*. Midsea Books.

Myers-Scotton, C. (2002). *Contact linguistics: Bilingual encounters and grammatical outcomes*. Oxford University Press.

Nag, S., John, S., & Agrawal, A. (2024). NSP-SCD: A corpus construction protocol for child-directed print in understudied languages. *Behavior Research Methods*, *56*(4), 2751–2764. https://doi.org/10.3758/s13428-024-02339-x

Nakano, N., Park, H., & Schmitt, C. (2009). A comparative study of the acquisition of plural morphemes in Japanese and Korean. Proceedings of the 11th Tokyo conference on psycholinguistics. Tokyo: Hituzi Shobo.

Nakisa, R., Plunkett, K., & Hahn, U. (2000). Single- and dual-route models of inflectional morphology. In P. Broeder & J. Murre (Eds.), *Models of language acquisition: Inductive and deductive approaches* (pp. 201–222). Oxford University Press. https://doi.org/10.1093/oso/9780198299899.003.0010

Nappa, J. Y. (2024). *Speech and language development in children born during the COVID-19 pandemic* [Bachelor's dissertation, University of Malta]. University of Malta.

Narafshan, M. H., Sadighi, F., Bagheri, M. S., & Shokrpour, N. (2014). The role of input in first language acquisition. *International Journal of Applied*

Linguistics & English Literature, 3(1), 86–91. https://doi.org/10.7575/aiac.ijalel.v.3n.1p.86

National Statistics Office. (2023a). *Census of population and housing 2021: Final report.* NSO.

National Statistics Office. (2023b). *Malta Skills Survey 2022 (Preliminary report).* NSO.

Nazzi, T., Barrière, I., Goyet, L., Kresh, S. and Legendre, G. (2011). Tracking irregular morphophonological dependencies in natural language: evidence from the acquisition of subject-verb agreement in French. Cognition 120:119–135.

Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development, 38*(1–2, Serial No. 149), 1–135. https://doi.org/10.2307/1165788

Newman, R. S., Rowe, M. L., & Bernstein Ratner, N. (2016). Input and uptake at 7 months predicts toddler vocabulary: The role of child-directed speech and infant processing skills in language development. *Journal of Child Language, 43*(5), 1158–1173. https://doi.org/10.1017/S0305000915000446

Newsbook. (2025). *Jesprimu tħassib fuq il-Malti li jintuża fiċ-ċentri taċ-childcare [Expressing concerns on the Maltese used in childcare centres]*. Retrieved from https://newsbook.com.mt/jesprimu-thassib-fuq-il-malti-li-jintuza-fic-centri-tac-childcare/

Nguyen, T. A., Kharitonov, E., Copet, J., Adi, Y., Hsu, W.-N., Elkahky, A., Tomasello, P., Algayres, R., Sagot, B., Mohamed, A., & Dupoux, E. (2023). Generative Spoken Dialogue Language Modeling. *Transactions of the Association for Computational Linguistics, 11*, 250–266. https://doi.org/10.1162/tacl_a_00545

Nic Fhlannchadha, S., & Hickey, T. M. (2017). Acquiring an opaque gender system in Irish, an endangered indigenous language. *First Language*, *37*(5), 475–499. https://doi.org/10.1177/0142723717702942

Nicoladis, E., & Marchak, K. (2011). Le Carte Blanc or la Carte Blanche? Bilingual Children's Acquisition of French Adjective Agreement. *Language Learning*, *61*(3), 734–758. https://doi.org/10.1111/j.1467-9922.2011.00660.x

Nicoladis, E., & Secco, G. (2000). The role of a child's productive vocabulary in the language choice of a bilingual family. *First Language 20*(58), 003–028.

Nieder, J., van de Vijver, R., & Mitterer, H. (2020). Knowledge of Maltese singular-plural mappings. *Morphology, 31*(2), 147–170. https://doi.org/10.1007/s11525-020-09353-7

Nikolaus, M., Agrawal, A., Kaklamanis, P., Warstadt, A., & Fourtassi, A. (2024). Automatic annotation of grammaticality in child-caregiver conversations. In N. Calzolari, M.-Y. Kan, V. Hoste, A. Lenci, S. Sakti, & N. Xue (Eds.), *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)* (pp. 1832–1844). ELRA and ICCL. https://aclanthology.org/2024.lrec-main.164/

Ninio, A. (2011). *Syntactic development, its input and output.* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199565962.001.0001

Nittrouer, S., & Lowenstein, J. H. (2024). Early otitis media puts children at risk for later auditory and language deficits. *International Journal of Pediatric Otorhinolaryngology, 176*, 111801. https://doi.org/10.1016/j.ijporl.2023.111801

Ntelitheos, D., & Idrissi, A. (2017). Language growth in child Emirati Arabic. In H. Ouali (Ed.), *Perspectives on Arabic linguistics XXIX: Papers from the annual symposium on Arabic linguistics, Milwaukee, Wisconsin, 2015* (pp. 229–248). John Benjamins Publishing Company. https://doi.org/10.1075/sal.5.10nte

Nwokah, E. E., & Graves, K. N. (2009). Word play: The creation and function of novel words in the pretend play of two siblings. *American Journal of Play, 1*(3), 327–351.

Ochs, E. (1987). Introduction. In B. B. Schieffelin & E. Ochs (Eds.), *Language socialization across cultures* (pp. 1–14). Cambridge University Press.

Ochs, E. & Schieffelin, B. (2017). Language socialization: A historical overview. In P. Duff & S. May (Eds.), *Language socialization* (3rd ed., pp. 3-16). Springer.

Ochs, E., & Schieffelin, B. (1995). The impact of language socialization on grammatical development. In P. Fletcher & B. MacWhinney (Eds.), *The handbook of child language* (pp. 73–94). Blackwell.

Ochs, E., & Schieffelin, B. (2012). The theory of language socialization. In A. Duranti, E. Ochs, & B. B. Schieffelin (Eds.), *The handbook of language socialization* (pp. 1–21). Wiley-Blackwell.

Oliver, B. R., & Plomin, R. (2007). Twins' Early Development Study (TEDS): A multivariate, longitudinal genetic investigation of language, cognition and behavior problems from childhood through adolescence. *Twin Research and Human Genetics, 10*(1), 96–105. https://doi.org/10.1375/twin.10.1.96

Oller, D. K. (2000). *The emergence of the speech capacity.* Lawrence Erlbaum Associates.

Omar, A. S. (2007). Acquisition of gender and number agreement in Egyptian Arabic. *SKASE Journal of Theoretical Linguistics*, 4(2), 42–53.

Oosthuizen, H., & Southwood, F. (2009). Methodological issues in the calculation of mean length of utterance (MLU). *South African Journal of Communication Disorders, 56*(1), 76–87. https://doi.org/10.4102/sajcd.v56i1.194

Otwinowska, A., Opacki, M., Mieszkowska, K., Białecka-Pikul, M., Wodniecka, Z., & Haman, E. (2022). Polish–English bilingual children overuse referential markers: MLU inflation in Polish-language narratives. *First Language*, *42*(2), 191–215. https://doi.org/10.1177/0142723720933769

Özçalışkan, Ş. (2007). Metaphors we move by: Children's developing understanding of metaphorical motion in typologically distinct languages. *Metaphor and Symbol, 22*(2), 147–168. https://doi.org/10.1080/10926480701235429

Ozturk, S., Pinar, E., Ketrez, F. N., & Özçalişkan, Ş. (2021). Effect of sex and dyad composition on speech and gesture development of singleton and twin children. *Journal of Child Language*, *48*(5), 1048–1066. https://doi.org/10.1017/S0305000920000744

Pace, T., & Borg, A. (2017). The status of Maltese in national language-related legislation and implications for its use. *Revista de Llengua i Dret / Journal of Language and Law, 67*, 70–85. https://doi.org/10.2436/rld.i67.2017.2935

Padilla, A. M. (1978). The acquisition of fourteen English grammatical morphemes in the speech of bilingual children. *The Bilingual Review, 5*(3), 163–168.

Paradis J. (2023). Sources of individual differences in the dual language development of heritage bilinguals. *Journal of Child Language 50*, 793–817, https://doi.org/10.1017/S0305000922000708

Paradis, J. (2010). The interface between bilingual development and specific language impairment. *Applied Psycholinguistics, 31*(2), 227–252. https://doi.org/10.1017/S0142716409990373

Paradis, J., & Genesee, F. (1996). Syntactic Acquisition in Bilingual Children: Autonomous or Interdependent? *Studies in Second Language Acquisition*, *18*(1), 1–25. doi:10.1017/S0272263100014662

Paradis, J., Genesee, F., & Crago, M. (2011). *Dual language development and disorders: A handbook on bilingualism and second language learning* (2nd ed.). Brookes Publishing.

Paradis, J., Nicoladis, E., & Genesee, F. (2000). Early emergence of structural constraints on code-mixing: Evidence from French–English bilingual children. *Bilingualism: Language and Cognition 3*, 245–61.

Parisse, C. (2019). How large should a dense corpus be for reliable studies in early language acquisition? *Cognitextes, 19*. https://doi.org/10.4000/cognitextes.1483

Parisse, C. (2023). Acquisition of Inflection in Romance Languages. In *Oxford Research Encyclopedia of Linguistics*. Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.728

Parisse, C., & Le Normand, M.-T. (2000). How children build their morphosyntax: the case of French. *Journal of Child Language*, *27*(2), 267–292. doi:10.1017/S0305000900004116

Pearson, B. Z., Fernández, S. C., & Oller, D. K. (1993). Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language Learning, 43*(1), 93–120. https://doi.org/10.1111/j.1467-1770.1993.tb00174.x

Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009). Statistical learning in a natural language by 8-month-old infants. Child Development, 80, 674–685. https://doi.org/10.1111/j.1467-8624.2009.01290.x

Pérez-Pereira, M. (1991). The acquisition of gender: What Spanish children tell us. *Journal of Child Language, 18*(3), 571–590.

Pérez-Pereira, M. (2011). The debate on language acquisition. *Infancia y Aprendizaje, 34*(3), 259–261. https://doi.org/10.1174/021037011797238522

Perkins, K., & Zhang, L. J. (2024). The Effect of First Language Transfer on Second Language Acquisition and Learning: From Contrastive Analysis to Contemporary Neuroimaging. *RELC Journal*, *55*(1), 162-178. https://doi.org/10.1177/00336882221081894

Perpiñán, S. (2015). The locative paradigm in the L2 Spanish of Catalan native speakers. In T. Judy & S. Perpiñán (Eds.), *Acquisition of Spanish in Understudied Language Pairings* (Vol. 3). John Benjamins Publishing Company.

Peters, A. M. (1983). The units of language acquisition. Cambridge University Press.

Petitto, L. A., & Marentette, P. F. (1991). Babbling in the Manual Mode: Evidence for the Ontogeny of Language. *Science (American Association for the Advancement of Science)*, *251*(5000), 1493–1496. https://doi.org/10.1126/science.2006424

Piasta, S. B., Groom, L. J., Khan, K. S., Skibbe, L. E., & Bowles, R. P. (2018). Young children's narrative skill: Concurrent and predictive associations with emergent literacy and early word reading skills. *Reading and Writing, 31*(7), 1479–1498. https://doi.org/10.1007/s11145-018-9844-7

Piattelli-Palmarini, M. (1979). *Language and learning: The debate between Jean Piaget and Noam Chomsky.* Harvard University Press.

Piattelli-Palmarini, M. (1994). Ever since language and learning: Afterthoughts on the Piaget-Chomsky debate. *Cognition, 50*(1), 315–346. https://10.1016/0010-0277(94)90034-5

Piccin, T. B., & Waxman, S. R. (2007). Why Nouns Trump Verbs in Word Learning: New Evidence from Children and Adults in the Human Simulation Paradigm. *Language Learning and Development*, *3*(4), 295–323. https://doi.org/10.1080/15475440701377535

Pine, J. M. (1995). Variation in vocabulary development as a function of birth order. *Child Development, 66*(1), 272–281. https://doi.org/10.2307/1131205

Pine, J. M., Freudenthal, D., Krajewski, G., & Gobet, F. (2013). Do young children have adult-like syntactic categories? Zipf's law and the case of the determiner. *Cognition*, *127*(3), 345–360. https://doi.org/10.1016/j.cognition.2013.02.006

Pine, J. M., Lieven, E. V. M., & Rowland, C. (1996). Observational and checklist measures of vocabulary composition: What do they mean? *Journal of Child Language*, *23*(3), 573–589. https://doi.org/10.1017/S0305000900008953

Pinker, S. (1984). *Language learnability and language development*. Harvard Univ. Press.

Pinker, S. (1991). Rules of language. *Science, 253*, 530–535.

Place, S., & Hoff, E. (2016). Effects and noneffects of input in bilingual environments on dual language skills in 2½-year-olds. *Bilingualism: Language and Cognition, 19*, 1023–1041. https://doi.org/10.1017/S1366728915000322

Ponvannan, P., Chinnasami, B., Vaishali, S., Pinnamaneni, M., & Subash, S. (2024). Gender differences in the performance of Language Evaluation Scale Trivandrum (LEST) in children aged between zero and three years. *Cureus, 16*(7), e63786. https://doi.org/10.7759/cureus.63786

Portelli, J. (1997). *Mean length of utterance in three to four year old Maltese speaking children* [Bachelor's dissertation, University of Malta]. University of Malta.

Portelli, J. (2005). *The Mean Length of Utterance in Maltese Speaking Children*. [Master's dissertation, University of Malta]. University of Malta.

Potratz, J. R., Gildersleeve-Neumann, C., & Redford, M. A. (2022). Measurement Properties of Mean Length of Utterance in School-Age Children. *Language, Speech & Hearing Services in Schools*, *53*(4), 1088–1100. https://doi.org/10.1044/2022_LSHSS-21-00115

Preyer, W. (1882). *Die Seele des Kindes: Beobachtungen über die geistige Entwicklung des Menschen in den ersten Lebensjahren [The Soul of the Child:*

Observations on the Mental Development of Man in the First Years of Life]. Th. Grieben's Verlag.

Prince, A., & Smolensky, P. (2004). *Optimality theory: Constraint interaction in generative grammar.* Blackwell.

Pulis, L. (2021). *Exploring relationships between parental input and children's expressive vocabulary skills* [Bachelor's dissertation, University of Malta]. University of Malta.

Pye, C. (1992). The acquisition of K'iche' (Maya). In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition* (pp. 221–308). Lawrence Erlbaum Associates.

Quick, A., Hartmann, S., Backus, A., & Lieven, E. (2019). Entrenchment and productivity: The role of input in the code-mixing of a German–English bilingual child. *Applied Linguistics Review, 12*(2), 225–247. https://doi.org/10.1515/applirev-2019-0027

Quick, A. E., Lieven, E., Backus, A., & Tomasello, M. (2018). Constructively combining languages: The use of code-mixing in German–English bilingual child language acquisition. *Linguistic Approaches to Bilingualism, 8*(3), 393–409.

Racz, P., Pierrehumbert, J. B., Hay, J. B., & Papp, V. (2015). Morphological emergence. In B. MacWhinney & W. O'Grady (Eds.), *The handbook of language emergence* (pp. 123–146). Wiley-Blackwell.

Radford, A. (1990). *Syntactic theory and the acquisition of English syntax: The nature of early child grammars of English*. Blackwell.

Ramscar, M., Dye, M., & McCauley, S. (2013). Error and expectation in language learning: The curious absence of 'mouses' in adult speech. Language, 89(4), 760–793.

Ravid, D. (2019). First-language acquisition of morphology. In *Oxford research encyclopedia of linguistics*. Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.603

Ravid, D., & Avidor, A. (1998). Acquisition of derived nominals in Hebrew: Developmental and linguistic principles. *Journal of Child Language*, *25*, 229–266.

Ravid, D., & Farah, R. (2009). Noun plurals in early Palestinian Arabic: A longitudinal case study. In U. Stephany & M. D. Voeikova (Eds.), *Development of nominal inflection in first language acquisition* (pp. 411–432). De Gruyter Mouton. https://doi.org/10.1515/9783110217117.411

Ravid, D. & Schiff, R. (2009). Morpho-phonological categories of noun plurals in Hebrew: A developmental study. *Linguistics,* 47 (1), 45-63.

Ravid, D., & Schiff, R. (2012). From dichotomy to divergence: Number/gender marking on Hebrew nouns and adjectives across school ages. *Language Learning, 62*(S1), 66–106. https://doi.org/10.1111/j.1467-9922.2011.00675.x

Ravid, D. & Farah, R. (1999). Learning about noun plurals in early Palestinian Arabic. *First Language 19*, 187–206. https://doi.org/10.1177/014272379901905603.

Redlinger, W. E., & Park, T.-Z. (1980). Language mixing in young bilinguals. *Journal of Child Language*, *7*(2), 337–352. doi:10.1017/S030500090000266X

Rescorla, L. A. (1980). Overextension in early language development. Journal of Child Language 7, 321-335.

Rice, M. L., Zubrick, S. R., Taylor, C. L., Gayán, J., & Bontempo, D. E. (2014). Late language emergence in 24-month-old twins: Heritable and increased risk for late language emergence in twins. *Journal of Speech, Language, and Hearing Research*, *57*(3), 917–928. https://doi.org/10.1044/1092-4388(2013/12-0350)

Rinaldi, P., Pasqualetti, P., Volterra, V., & Caselli, M. C. (2021). Gender differences in early stages of language development. Some evidence and possible explanations. *Journal of Neuroscience Research*, *101*(5), 643–653. https://doi.org/10.1002/jnr.24914

Rispens, J., & Krikhaar, E. (2010). Using event-related potentials in language acquisition research. In E. Blom & S. Unsworth (Eds.), *Experimental methods in language acquisition research* (pp. 95–114). John Benjamins Publishing Company. https://doi.org/10.1075/lllt.27.07ris

Roberson, D., Davidoff, J., Davies, I. R. L., & Shapiro, L. R. (2004). The Development of Color Categories in Two Languages: A Longitudinal Study. *Journal of Experimental Psychology: General, 133*(4), 554–571. https://doi.org/10.1037/0096-3445.133.4.554

Romaine, S. (1995). *Bilingualism*. Blackwell Publishers.

Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews. Cognitive Science*, *1*(6), 906–914. https://doi.org/10.1002/wcs.78

Ronjat, J. (1913). *Le développement du langage observé chez un enfant bilingue*. (The development of language observed in a bilingual child). Librairie Ancienne H. Champion.

Roseberry, S., Hirsh-Pasek, K., & Golinkoff, R. M. (2014). Skype Me! Socially Contingent Interactions Help Toddlers Learn Language. *Child Development, 85*(3), 956–970. https://doi.org/10.1111/cdev.12166

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development, 83*(5), 1762–1774. https://doi.org/10.1111/j.1467-8624.2012.01805.x

Rowe M. L. (2015). Input versus intake - a commentary on Ambridge, Kidd, Rowland, and Theakson's 'the ubiquity of frequency effects in first language acquisition'. *Journal of child language*, *42*(2), 301–322. https://doi.org/10.1017/S030500091400066X

Rowe, M. L., & Snow, C. E. (2020). Analyzing input quality along three dimensions: interactive, linguistic, and conceptual. *Journal of Child Language, 47*(1), 5–21. https://doi.org/10.1017/S0305000919000655

Rowland, C. (2014). *Understanding Child Language Acquisition*. Routledge.

Roy, J., Oetting, J. B., & Moland, C. W. (2013). Linguistic constraints on children's overt marking of BE by dialect and age. *Journal of Speech, Language, and Hearing Research*, *56*(3), 933–944. https://doi.org/10.1044/1092-4388(2012/12-0099)

Royle, P., & Valois, D. (2010). Acquisition of adjectives in Quebec French as revealed by elicitation data. *Journal of French Language Studies, 20*(3), 313–338. https://doi.org/10.1017/S0959269510000013

Saffran, J. R. (2020). Statistical language learning in infancy. *Child Development Perspectives, 14*(1), 49–54.

Saffran, J., Aslin, R., & Newport, E. (1996). Statistical Learning by 8-Month-Old Infants. *Science (American Association for the Advancement of Science), 274*(5294), 1926–1928.

Saiegh-Haddad, E., Hadieh, A., & Ravid, D. (2012). Acquiring noun plurals in Palestinian Arabic: Morphology, familiarity, and pattern frequency. *Language Learning, 62*(4), 1079–1109. https://doi.org/10.1111/j.1467-9922.2012.00727.x

Saji, N., Imai, M., & Asano, M. (2020). Acquisition of the Meaning of the Word Orange Requires Understanding of the Meanings of Red, Pink, and Purple: Constructing a Lexicon as a Connected System. *Cognitive Science*, *44*(1), e12813-n/a. https://doi.org/10.1111/cogs.12813

Sanjeevan, T., Rosenbaum, D. A., Miller, C. A., Hell, J. G., Weiss, D. J., & Mainela-Arnold, E. (2015). Motor issues in specific language impairment: A window into the underlying impairment. *Current Developmental Disorders Reports, 2*(3), 228–236. https://doi.org/10.1007/s40474-015-0051-9

Sansavini, A., Guarini, A., Savini, S., Broccoli, S., Justice, L., Alessandroni, R., & Faldella, G. (2011). Longitudinal trajectories of gestural and linguistic abilities in very preterm infants in the second year of life. *Neuropsychologia, 49*(13), 3677–3688. https://doi.org/10.1016/j.neuropsychologia.2011.09.023

Sauerland, U., Grohmann, K. K., Guasti, M. T., Anđelković, D., Argus, R., Armon-Lotem, S., Arosio, F., Avram, L., Costa, J., Dabašinskienė, I., Jensen de López, K., Gatt, D., Grech, H., Haman, E., van Hout, A., Hrzica, G., Kainhofer, J., Kamandulytė-Merfeldienė, L., Kunnari, S., … Yatsushiro, K. (2016). How do 5-year-olds understand questions? Differences in languages across Europe. *First Language, 36*(3), 169–202. https://doi.org/10.1177/0142723716640236

Savičiūtė, E., Ambridge, B., & Pine, J. M. (2018). The roles of word-form frequency and phonological neighbourhood density in the acquisition of Lithuanian noun morphology. *Journal of Child Language, 45*(3), 641–672.

Scarborough, H. S. (1990). Index of Productive Syntax. *Applied Psycholinguistics*, *11*, 1–22.

Scerri, T. (2015). *Bilingual Interaction by Young Children in Public Play Areas* [Bachelor’s dissertation, University of Malta]. University of Malta.

Schembri, T. (2006). *The Broken Plural in Maltese: An Analysis* [Bachelor’s dissertation, University of Malta]. University of Malta.

Schembri, T. (2012). *The Broken Plural in Maltese: A Description*. Bochum: Universitätsverlag Brockmeyer.

Schmid, H.-J. (2020). *The dynamics of the linguistic system: Usage, conventionalization, and entrenchment.* Oxford University Press.

Schmid, M. S. (2011). *Language attrition* (1st ed.). Cambridge University Press.

Schmitt, M. B., Logan, J. A. R., Tambyraja, S. R., Farquharson, K., & Justice, L. M. (2017). Establishing language benchmarks for children with typically developing language and children with language impairment. *Journal of Speech, Language, and Hearing Research*, *60*(2), 364–378. https://doi.org/10.1044/2016_JSLHR-L-15-0273

Schnell de Acedo, B. (1994). Early morphological development: The acquisition of articles in Spanish. In Handbook of research in language development using CHILDES, eds J. L. Sokolov and C. E. Snow, 210–53. Hillsdale, NJ: Lawrence Erlbaum.

Scott, M. E., Kanero, J., Saji, N., Chen, Y., Imai, M., Golinkoff, R. M., & Hirsh-Pasek, K. (2023). From green to turquoise: Exploring age and socioeconomic status in the acquisition of color terms. First Language, 43(1), 3–21. https://doi.org/10.1177/01427237221112499

Scovel, T. (2000). A Critical Review of the Critical Period Research. *Annual Review of Applied Linguistics, 20*, 213-223. https://ejournals.um.edu.mt/login?url=https://www.proquest.com/scholarly-journals/critical-review-period-research/docview/85517612/se-2

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*(4), 523–568.

Serratrice, L. (2013). Cross-linguistic influence in bilingual development: Determinants and mechanisms. Linguistic Approaches to Bilingualism, 3(1), 3–25. https://doi.org/10.1075/lab.3.1.01ser

Serratrice, L. (2019). Becoming bilingual in early childhood. In A. De Houwer & L. Ortega (Eds.), *The Cambridge handbook of bilingualism* (pp. 15–35). Cambridge University Press.

Skidelsky, L. (2013). Faculty of language, functional models, and mechanisms. *Journal of Cognitive Science, 14*(2), 111–149. https://doi.org/10.17791/jcs.2013.14.2.111

Skidelsky, L. (2016). The poverty of the stimulus argument once again. *Análisis Filosófico, 36*(2), 143–170. http://www.scielo.org.ar/scielo.php?script=sci_arttext&pid=S1851-96362016000200001&lng=en&tlng=en

Skinner, B. F. (1957). *Verbal behavior.* Appleton-Century-Crofts. https://doi.org/10.1037/11256-000

Slobin, D. I. (1985). Crosslinguistic evidence for the language-making capacity. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition: Vol. 1. The data; Vol. 2. Theoretical issues* (pp. 1157–1256). Lawrence Erlbaum Associates.

Smith-Hefner, N. J. (1988). The linguistic socialization of Javanese children in two communities. *Anthropological Linguistics, 30*(2), 166–198. https://www.jstor.org/stable/30027978

Smoczyńska, M. (1985). The acquisition of Polish. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition: Vol. 1. The data; Vol. 2. Theoretical issues* (pp. 595–686). Lawrence Erlbaum Associates.

Snyder, W. (2007). *Child language: The parametric approach.* Oxford: Oxford University Press.

Snyder, W. (2021). A parametric approach to the acquisition of syntax. *Journal of Child Language, 48*(5), 862–887. https://doi.org/10.1017/S0305000921000465

Socarrás, G. M. (2011). *First language acquisition in Spanish: A minimalist approach to nominal agreement*. London: Continuum International Publishing Group.

Soderstrom, M. (2007). Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants. *Developmental Review*, *27*(4), 501–532. https://doi.org/10.1016/j.dr.2007.06.002

Sorenson Duncan, T., & Paradis, J. (2020). Home language environment and children's second language acquisition: The special status of input from older siblings. *Journal of Child Language*, Advance online publication. https://doi.org/10.1017/S0305000919000977

Spagnol, M. (2011). *A tale of two morphologies* [Doctoral dissertation, University of Konstanz]. University of Konstanz.

Spagnol, M., & Fabri, R. (2023). *Journal of Maltese Studies: On Maltese syntax: volume: 30*. University of Malta. Dipartiment tal-Malti.

Squires, J., & Bricker, D. (2009). *Ages & Stages Questionnaires®, Third Edition (ASQ-3™)* [Database record].

Stankiewicz, E. (Ed.). (1972). *A Baudouin de Courtenay anthology*. Indiana University Press.

Stark, R. E. (1980). Stages of speech development in the first year of life. In G. Yeni-Komshian, J. Kavanagh, & C. Ferguson (Eds.), *Child phonology* (Vol. 1, pp. 73–90). Academic Press.

Stengelin, R., Haun, D. B. M., & Kanngiesser, P. (2023). Simulating peers: Can puppets simulate peer interactions in studies on children's socio-cognitive development? *Child Development*, *94*(5), 1117–1135. https://doi.org/10.1111/cdev.13913

Stephany, U., & Voeikova, M. D. (2009). *Development of nominal inflection in first language acquisition: A cross-linguistic perspective.* De Gruyter Mouton. https://doi.org/10.1515/9783110217117

Stern, C., & Stern, W. (1928). *Die Kindersprache: Eine psychologische und sprachtheoretische Untersuchung* (4th rev. ed.; 1st ed. 1907). Barth. (Reprinted *Darmstadt: Wissenschaftliche Buchgesellschaft*, 1965.).

Stoll, S. (2015). Inflectional morphology in language acquisition. In M. Baerman (Ed.), *The Oxford handbook of inflection* (pp. 351–374). Oxford University Press.

Stoll, S., & Bickel, B. (2013). Capturing diversity in language acquisition research. In B. Bickel, L. A. Grenoble, D. A. Peterson, & A. Timberlake (Eds.), *Language typology and historical contingency: In honor of Johanna Nichols* (pp. 195–216). John Benjamins. https://doi.org/10.1075/tsl.104.08slo

Stolz, C. (2009). A different kind of gender problem: Maltese loanword gender from a typological perspective. In B. Comrie, R. Fabri, E. Hume, M. Mifsud, T. Stolz, & M. Vanhove (Eds.), *Introducing Maltese linguistics* (pp. 321–353). John Benjamins.

Stump, G. T. (2001). *Inflectional morphology: A theory of paradigm structure*. Cambridge University Press.

Sultana, M. (2014). *Bilingual exposure at home and at school in Early Childhood.* [Bachelor's dissertation, University of Malta]. University of Malta.

Sundara, M. (2022). Development of speech perception and production in bilingual preverbal infants. In B. Goldstein (Ed.), *Bilingual language development and disorders in Spanish- English speakers* (3rd ed.) (pp. 67–94). Brookes Publishing.

Svenonius, P. (2019). Syntactic features. In *Oxford research encyclopedia of linguistics*. Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.179

Swearengin, P., Mihalicz, P., Marsh, L., Baker, Z., Perry, F., Graham, S. A., & Childers, J. B. (2025). Examining the influence of encoding and retrieval contexts on 2- to 4-year-olds' acquisition of nouns and verbs. *British Journal of Developmental Psychology*, *43*(3), 730–741. https://doi.org/10.1111/bjdp.12547

Syrnyk, C., & Meints, K. (2017). Bye-bye mummy – Word comprehension in 9-month-old infants. *British Journal of Developmental Psychology, 35*(2), 202–217. https://doi.org/10.1111/bjdp.12157

Szagun, G., Stumper, B., Sondag, N., & Franik, M. (2007). The acquisition of gender marking by young German-speaking children: Evidence for learning guided by phonological regularities. *Journal of Child Language, 34*(3), 445–472. https://doi.org/10.1017/S0305000906007951

Taine, H. (1876). De l'acquisition du langage chez les enfants et dans l'espèce humaine. *Revue Philosophique, 1*, 23–32.

Takada, A. (2012). Preverbal infant–caregiver interaction. In A. Duranti, E. Ochs, & B. B. Schieffelin (Eds.), *The handbook of language socialization* (pp. 56–80). Wiley-Blackwell.

Tallas-Mahajna, N., & Dromi, E. (2023). Arabic morpheme per utterance: A morphological measure of child language development in spoken Arabic. *Language Acquisition*, *31*(2), 130–144. https://doi.org/10.1080/10489223.2023.2195386

Tamis-LeMonda, C. S., Kuchirko, Y., & Song, L. (2014). Why is infant language learning facilitated by parental responsiveness? *Current Directions in Psychological Science, 23*(2), 121–126. https://doi.org/10.1177/0963721414522813

Tardif, T. (1996). Nouns are not always learned before verbs: Evidence from Mandarin speakers' early vocabularies. *Developmental Psychology*, *32*(3), 492–504. https://doi.org/10.1037/0012-1649.32.3.492

Tavassoli, N., Recchia, H., & Ross, H. (2019). Preschool children's prosocial responsiveness to their siblings' needs in naturalistic interactions: A longitudinal study. *Early Education and Development, 30*(6), 724–742. https://doi.org/10.1080/10409289.2019.1599095

Taylor, K., & Kan, P. (2018). The impact of older siblings on vocabulary learning in bilingual children. *International Journal of Bilingual Education and Bilingualism*, Advance online publication. https://doi.org/10.1080/13670050.2018.1518969

Taylor, T. J., & van den Herik, J. C. (2021). Metalinguistic exchanges in child language development. *Language Sciences, 88*, Article 101434. https://doi.org/10.1016/j.langsci.2021.101434

Tchintcharauli, T., Tsintsadze, N., Damenia, T., Kalkhitashvili, T., Doborjginidze, N., & Uziel-Karl, S. (2024). Establishing guidelines for MLU measurement in an agglutinating language: An illustration of Georgian. *First Language*, *44*(4), 345–371. https://doi.org/10.1177/01427237241247930

Tek, S., Mesite, L., Fein, D., & Naigles, L. (2014). Longitudinal analyses of expressive language development reveal two distinct language profiles among young children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 44*(1), 75–89. https://doi.org/10.1007/s10803-013-1853-4

Theakston, A. L. (2004). The role of entrenchment in children's and adults' performance on grammaticality judgment tasks. *Cognitive Development, 19*(1), 15–34. https://doi.org/10.1016/j.cogdev.2003.08.001

Theobald, M., Bateman, A., Busch, G., Laraghy, M., & Danby, S. (2017). "I'm your best friend": Peer interaction and friendship in a multilingual preschool. In Theobald, M. (Ed.), *Friendship and peer culture in multilingual settings* [Sociological Studies of Children and Youth, Volume 21] (pp. 171–196). Emerald Group Publishing Limited.

Thordardottir, E. (2015). The relationship between bilingual exposure and morphosyntactic development. *International Journal of Speech-Language Pathology, 17*(2), 97–114. https://doi.org/10.3109/17549507.2014.923509

Thornton, R. (2016). Children's acquisition of syntactic knowledge. In *Oxford research encyclopedia of linguistics*. Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.72

Thothathiri, M., Kidd, E., & Rowland, C. (2025). The role of executive function in the processing and acquisition of syntax. *Royal Society Open Science, 12*(3), 201497. https://doi.org/10.1098/rsos.201497

Tiedemann, D. (1787). Beobachtungen über die Entwickelung der Seelenfähigkeiten bei Kindern [Observations on the development of the mental faculties of children]. *Hessische Beiträge zur Gelehrsamkeit und Kunst, 2*, 313–315; 3, 486–488.

Tippenhauer, N., Fourakis, E. R., Watson, D. G., & Lew-Williams, C. (2020). The scope of audience design in child-directed speech: Parents' tailoring of word lengths for adult versus child listeners. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*(11), 2163–2178.

Tomas, E., & Dorofeeva, S. (2019). Mean length of utterance and other quantitative measures of spontaneous speech in Russian-speaking children. *Journal of Speech, Language, and Hearing Research*, *62*(12), 4483–4496. https://doi.org/10.1044/2019_JSLHR-L-18-0339

Tomasello, M. (2000). First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics, 11*, 61–82.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Harvard University Press.

Tomasello, M. (2015). The usage-based theory of language acquisition. In E. L. Bavin & L. R. Naigles (Eds.), *The Cambridge handbook of child language* (pp. 89–106). Cambridge University Press. https://doi.org/10.1017/CBO9781316095829.005

Tomasello, M. (2022). *The evolution of agency: Behavioral organization from lizards to humans.* MIT Press.

Tomblin, J. B., Harrison, M., Ambrose, S. E., Walker, E. A., Oleson, J. J., & Moeller, M. P. (2015). Language outcomes in young children with mild to severe hearing loss. *Ear and Hearing, 36*(Suppl. 1), 76S–91S. https://doi.org/10.1097/AUD.0000000000000219

Tomlin, R. S. (1997). Mapping conceptual representations into linguistic representations: The role of attention in grammar. In J. Nuyts & E. Pederson (Eds.), *Language and conceptualization* (pp. 162–189). Cambridge University Press. https://doi.org/10.1017/CBO9781139086677.007

Torrens, V. (2020). *Typical and impaired processing in morphosyntax.* John Benjamins Publishing Company.

Tribushinina, E., & Mak, P. (2022). Production of pronoun gender by children acquiring Russian as a minority language: Comparison with the effects of developmental language disorder. *Stem-, Spraak- en Taalpathologie, 27*, 223–249. https://doi.org/10.21827/32.8310/2022-SG-223

Tsimpli, I. M., & Hulk, A. (2013). Grammatical gender and the notion of default: Insights from language acquisition. *Lingua, 137*, 128–144. https://doi.org/10.1016/j.lingua.2013.09.001

Tsinivits, D., & Unsworth, S. (2021). The impact of older siblings on the language environment and language development of bilingual toddlers. *Applied Psycholinguistics*, *42*(2), 325–344. https://doi.org/10.1017/S0142716420000570

Tulloch, M. K., & Hoff, E. (2023). Early language development: Major milestones and individual differences. In R. J. Tierney, F. Rizvi & K. Ercikan (Eds.), *International encyclopedia of education (fourth edition)* (pp. 467–476). Elsevier. https://10.1016/B978-0-12-818630-5.14063-1

Ullman, M. T. (2002). Contributions of memory circuits to language: The declarative/procedural model. *Cognition, 92*, 231–270.

Unsworth, S. (2013). Assessing the role of current and cumulative exposure in simultaneous bilingual acquisition: The case of Dutch gender. *Bilingualism: Language and Cognition, 16*(1), 86–110. https://doi.org/10.1017/S1366728912000284

Vainio, S., Pajunen, A., & Häikiö, T. (2019). Acquisition of Finnish derivational morphology: School-age children and young adults. First Language, 39(2), 139–157.

Vaks, A., & Vihman, V. A. (2025). Bilingual acquisition of morphology: Norwegian and Russian influence on children's sentence repetition in Estonian. *Language and Speech*, Advance online publication. https://doi.org/10.1177/00238309251327209

Vaks, A., Vihman, V.-A., & de Jong, J. (2025). Language exposure and competence in Estonian-Norwegian bilingual children. *Eesti Ja Soome-Ugri Keeleteaduse Ajakiri. Journal of Estonian and Finno-Ugric Linguistics*, *16*(2), 177–216. https://doi.org/10.12697/jeful.2025.16.2.06

Valian, V. (1986). Syntactic categories in the speech of young children. *Developmental Psychology, 22*(4), 562–579. https://doi.org/10.1037/0012-1649.22.4.562

Valian, V. (2014). Arguing about innateness. *Journal of Child Language, 41*(S1), 78–92. https://doi.org/10.1017/S0305000914000336

Valian, V., & Casey, L. (2003). Young children's acquisition of wh-questions: the role of structured input. *Journal of Child Language*, *30*(1), 117–143. https://doi.org/10.1017/S0305000902005457

Valleau, M. J., Konishi, H., Golinkoff, R. M., Hirsh-Pasek, K., & Arunachalam, S. (2018). An Eye-Tracking Study of Receptive Verb Knowledge in Toddlers. *Journal of Speech, Language, and Hearing Research*, *61*(12), 2917–2933. https://doi.org/10.1044/2018_JSLHR-L-17-0363

van de Vijver, R., & Baer-Henney, D. (2014). Developing biases. *Frontiers in Psychology, 5*, 634. https://doi.org/10.3389/fpsyg.2014.00634

van Dijk, C., van Wonderen, E., Koutamanis, E., Kootstra, G. J., Dijkstra, T., & Unsworth, S. (2022). Cross-linguistic influence in simultaneous and early sequential bilingual children: A meta-analysis. *Journal of Child Language, 49*, 897–929. https://doi.org/10.1017/S0305000921000337

Vannest, J., Karunanayaka, P. R., Schmithorst, V. J., Szaflarski, J. P., & Holland, S. K. (2009). Language networks in children: Evidence from functional MRI studies. *AJR: American Journal of Roentgenology, 192*(5), 1190–1196. https://doi.org/10.2214/AJR.08.2246

Vasilyeva, M., Waterfall, H., & Huttenlocher, J. (2008). Emergence of syntax: Commonalities and differences across children. *Developmental Science, 11*, 84–97.

Vassallo, M. (2006). *The language development of a Maltese toddler: A case study* [Bachelor's dissertation, University of Malta]. University of Malta.

Vella, A. (2012). Languages and language varieties in Malta. *International Journal of Bilingual Education and Bilingualism, 16*(5), 532–552. https://doi.org/10.1080/13670050.2012.716812

Vella, A., Spagnol, M., Grech, S., & Chetcuti, F. (2011). A preliminary investigation of filled pauses in Maltese. In S. Caruana, R. Fabri, & T. Stolz (Eds.), *Variation and change: The dynamics of Maltese in space, time and society* (pp. 255–275). Akademie Verlag.

Vella, L. A., Mifsud, C. L., & Muscat, D. (2022). *A language policy for the junior years in Malta and Gozo.* Ministry for Education, Sport, Youth, Research and Innovation.

Vella, M. (2020). *Morpho-syntactic errors in Maltese, performed by children whose first language (L1) is English and whose second language (L2) is Maltese* [Bachelor's dissertation, University of Malta]. University of Malta.

Vella, M. (2025). Maltese first language acquisition of the plural. In M. Vorholt, R. Bezzina, & M. Vella (Eds.), *The next century of Maltese linguistics* (pp. 279–300). Studia Typologica, De Gruyter Mouton.

Vescovi, G., Schilling, L., & Rowe, M. L. (2024). Digital media and language development: The role of child-directed speech (pp. 109–123). In G. B. Frizzo (Ed.), *Digital media and early child development*. Springer, Cham. https://doi.org/10.1007/978-3-031-69224-6_8

Vihman, M. M. (1982). The acquisition of morphology by a bilingual child: A whole-word approach. *Applied Psycholinguistics, 3*(2), 141–160. https://doi.org/10.1017/S0142716400006676

Vihman, M. M., Thierry, G., Lum, J., Keren-Portnoy, T., & Martin, P. (2007). Onset of word form recognition in English, Welsh, and English-Welsh bilingual infants. *Applied Psycholinguistics, 28*, 475–493.

Vihman, V.-A. (2018). Language Interaction in Emergent Grammars: Morphology and Word Order in Bilingual Children's Code-Switching. *Languages (Basel), 3*(4), 40-. https://doi.org/10.3390/languages3040040

Vihman, V.-A., Engelmann, F., Lieven, E. V. M., & Theakston, A. L. (2021). Many ways to decline a noun: Elicitation of children's novel noun inflection in Estonian. *Language and Cognition, 13*(4), 693–733. https://doi.org/10.1017/langcog.2021.19

Volling, B. L., Gonzalez, R., Oh, W., Song, J.-H., Yu, T., Rosenberg, L., Kuo, P. X., Thomason, E., Beyers-Carlson, E., Safyer, P., Stevenson, M. M., & Howe, N. (2017). *Developmental trajectories of children's adjustment across the transition to siblinghood: Pre-birth predictors and sibling outcomes at one year* (*Monographs of the Society for Research in Child Development, 82*(3)). https://www.jstor.org/stable/45106959

Volterra, Virginia & Traute Taeschner. (1978). The acquisition and development of language by bilingual children. *Journal of Child Language 5*, 311–326.

Von Stutterheim, C., Bouhaous, A., Carroll, M. & Sahonenko, N. (2012). Grammaticalized temporal categories, language specificity, and macroplanning in expository texts. *Linguistics*, *50*(2), 341-371. https://doi.org/10.1515/ling-2012-0012

Vorobyeva, T., Bel, A., & Voeikova, M. (2024). Grammatical gender agreement in production: The case of heritage Russian. *International Journal of Bilingualism, 28*(2), 234–254. https://doi.org/10.1177/13670069231155333

Vuksanovic, J. (2015). Relationship between social interaction bids and language in late talking children. *International Journal of Speech-Language Pathology, 17*(6), 527–536.

Wagner, K., Dobkins, K., & Barner, D. (2013). Slow mapping: Color word learning as a gradual inductive process. *Cognition*, *127*(3), 307–317. https://doi.org/10.1016/j.cognition.2013.01.010

Wallentin, M. (2009). Putative sex differences in verbal abilities and language cortex: A critical review. *Brain and Language, 108*(3), 175–183. https://doi.org/10.1016/j.bandl.2008.07.001

Wallentin, M., & Trecca, F. (2023). Cross-Cultural Sex/Gender Differences in Produced Word Content Before the Age of 3 Years. *Psychological Science*, *34*(4), 411–423. https://doi.org/10.1177/09567976221146537

Wasik, B. A., Bond, M. A., & Hindman, A. (2006). The effects of a language and literacy intervention on Head Start children and teachers. *Journal of Educational Psychology, 98*(1), 63–74. https://doi.org/10.1037/0022-0663.98.1.63

Wei, L. (2007). *The bilingualism reader* (2nd ed.). Routledge.

Weisleder, A., & Fernald, A. (2013). Talking to Children Matters: Early Language Experience Strengthens Processing and Builds Vocabulary. *Psychological Science, 24*(11), 2143–2152. https://doi.org/10.1177/0956797613488145

Werfel, K. L. (2018). Morphosyntax production of preschool children with hearing loss: An evaluation of the extended optional infinitive and surface accounts. *Journal of Speech, Language, and Hearing Research*, *61*(9), 2313–2324. https://doi.org/10.1044/2018_JSLHR-L-17-0406

White, L. J., Alexander, A., & Greenfield, D. B. (2017). The relationship between executive functioning and language: Examining vocabulary, syntax, and language learning in preschoolers attending Head Start. *Journal of Experimental Child Psychology, 164*, 16–31. https://doi.org/10.1016/j.jecp.2017.06.01

Williams, S. E., & Horst, J. S. (2014). Goodnight book: Sleep consolidation improves word learning via storybooks. *Frontiers in Psychology, 5*, 184. https://doi.org/10.3389/fpsyg.2014.00184

Winskel, H. (2006). The effects of an early history of otitis media on children's language and literacy skill development. *British Journal of Educational Psychology*, *76*(4), 727–744. https://doi.org/10.1348/000709905X68312

Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology, 24*, 220–251.

Xu, Z. (2011). Optimality Theory and Morphology. *Language and Linguistics Compass*, *5*(7), 466–484. https://doi.org/10.1111/j.1749-818X.2011.00290.x

Xuan, L., & Dollaghan, C. (2013). Language-specific noun bias: Evidence from bilingual children. *Journal of Child Language*, *40*(5), 1057–1075. https://doi.org/10.1017/S0305000912000529

Yip, V., & Matthews, S. (2016). Code-mixing and mixed verbs in Cantonese-English bilingual children: Input and innovation. *Languages, 1*(1), 4. https://doi.org/10.3390/languages1010004

Yip, V., & Matthews, S. (2022). Language diversity and bilingual first language acquisition: A commentary on Kidd and Garcia (2022). *First Language*, *42*(6), 832-836. https://doi.org/10.1177/01427237221097581 (Original work published 2022)

Zambrana, I. M., Ystrom, E., & Pons, F. (2012). Impact of gender, maternal education, and birth order on the development of language comprehension: A longitudinal study from 18 to 36 months of age. *Journal of Developmental & Behavioral Pediatrics, 33*(2), 146–155. https://doi.org/10.1097/DBP.0b013e31823d4f83

Zammit, D. (2020). *Il-ħakma tal-plural mit-tfal tar-raba' sena primarja [The mastery of the plural by children in the fourth year of primary school]* [Bachelor's dissertation, University of Malta]. University of Malta.

Zammit, R. (2008). *The use of number and gender markers in Maltese in 3- to 3;6-year-old children* [Bachelor's dissertation, University of Malta]. University of Malta.

Zapf, J. A., & Smith, L. B. (2007). When do children generalize the plural to novel nouns? *First Language*, *27*(1), 53–73. https://doi.org/10.1177/0142723707070286

Zheng, Z., Degotardi, S., & Djonov, E. (2021). Supporting multilingual development in early childhood education: A scoping review. *International Journal of Educational Research*, *110*, 101894-. https://doi.org/10.1016/j.ijer.2021.101894

Zimmerman, F. J., Christakis, Dimitri A., & Meltzoff, A. N. (2007). Associations between Media Viewing and Language Development in Children Under Age 2 Years. *The Journal of Pediatrics, 151*(4), 364–368. https://doi.org/10.1016/j.jpeds.2007.04.071

Zuniga-Montanez, C., Davies, C., Ligoxygakis, L., Kašćelan, D., & Gonzalez-Gomez, N. (2025). Annual research review: How did COVID-19 affect young children's language environment and language development? A scoping review. *Journal of Child Psychology and Psychiatry, 66*(4), 569–587. https://doi.org/10.1111/jcpp.14102

Appendices

Appendix A.1: Information sheet – Maltese

Karta ta' Informazzjoni

Dan l-istudju jifforma parti mir-riċerka tad-dottorat li qed nagħmel taħt is-superviżjoni tal-Professur Ray Fabri u l-Professur Alexandra Vella. L-għan tar-riċerka huwa li tippreżenta deskrizzjoni dettaljata tas-sekwenzi u l-istadji tal-iżvilupp li jikkaratterizzaw l-Akkwist tal-Malti bħala l-Ewwel Lingwa, b'enfasi speċjali fuq il-mod li bih jiġu akkwistati l-karatteristiċi tal-morfoloġija u s-sintassi tal-Malti, bħall-qbil bejn il-ġens grammatikali. Jeħtieġ li waqt il-proċess tal-ġbir tad-dejta, inżur lil bintek/ibnek għal madwar 45 minuta, f'intervalli ta' bejn 4 u 6 xhur, fuq medda ta' bejn 20 u 24 xahar (sentejn) sabiex nistudja l-iżvilupp tal-lingwa tagħha/tiegħu.

Se tingħata wkoll ittra ta' reklutaġġ, formola ta' kunsens, kwestjonarju dwar l-użu tiegħek tal-lingwa u kwestjonarju dwar l-iżvilupp ta' bintek/ibnek. Jekk taċċetta li bintek/ibnek t/jipparteċipa f'dan l-istudju, hi/u se t/jintalab t/jilgħab, t/jitkellem u t/jipparteċipa f'logħob tal-lingwa. Is-sessjonijiet, li se jiġu rrekordjati bl-użu kemm ta' awdjo kif ukoll ta' video recorder, se jseħħu f'ambjent informali. Kull sessjoni se ddum madwar 45 minuta u ġenitur/kustodju għandu jkun preżenti mill-bidu sal-aħħar. Ir-recordings se jintużaw biss għal skopijiet ta' traskrizzjoni u qatt mhuma se jkunu ppubblikati.

Jekk tiddeċiedi li tħalli lil bintek/ibnek t/jipparteċipa f'dan l-istudju, inti se tintalab tiffirma formola ta' kunsens u ittra ta' reklutaġġ. Tintalab ukoll biex timla 2 kwestjonarji. L-informazzjoni personali miġbura mill-kwestjonarji, li tista' twassal għall-identifikazzjoni tiegħek/ ta' bintek/ibnek, se titħassar u/jew tinbidel b'psewdonimi. Jekk f'kwalunkwe punt tiddeċiedi li ma tridx tipparteċipa aktar, tkun tista' tagħmel dan mingħajr ma jkollok bżonn tipprovdi raġuni. L-ebda dejta miġbura mingħandek/bintek/ibnek matul dan l-istudju mhi se tiġi diskussa ma' persuni oħra.

Is-superviżur ewlieni tiegħi huwa l-Professur Ray Fabri. Jekk ikollok bżonn tikkuntattjah, jekk jogħġbok iktiblu fuq ray.fabri@um.edu.mt.

Dawn li ġejjin huma d-dettalji ta' kuntatt tiegħi:

Isem: Michela Vella

Indirizz elettroniku: michela.vella.17@um.edu.mt

Firma: ______________________________ Data: ________________

Isem: _________________________________

Appendix A.2: Information sheet – English

Information Sheet

This study forms part of the doctoral research which I am carrying out under the supervision of Professor Ray Fabri and Professor Alexandra Vella. The aim of the research is to present a detailed description of the developmental patterns and stages that characterise Maltese First Language Acquisition (FLA), with a special focus on the way in which features of Maltese morphology and syntax, such as gender agreement, are acquired. The data collection process will require me to visit your child for around 45 minutes, at 4-to-6-month intervals, over a span of 20 to 24 months (2 years) so as to study her/his language development.

You will also be given a recruitment letter, a consent form, a language background questionnaire and a questionnaire about your child's development. If you allow your child to participate in this study, s/he will be requested to engage in free play and conversation and participate in language games. The sessions, which will be recorded using both an audio and a video recorder, will take place in an informal setting. Each session will last around 45 minutes and a parent/guardian must be present throughout. The recordings will be used solely for transcription purposes and will never be made public.

Should you decide to allow your child to participate in this study, you will be asked to sign a consent form and a recruitment letter. You will also be required to complete 2 questionnaires. Personal information collected from the questionnaires, which could lead to your/ your child's identification, will be deleted and/or replaced by pseudonyms. If you decide to withdraw your participation at any time, you will be able to do so without needing to provide a reason. None of the data collected from you/ your child in the course of this study will be discussed with third parties.

My main supervisor is Professor Ray Fabri. Should you need to contact him, please write to him at ray.fabri@um.edu.mt.

The following are my contact details:

Name: Michela Vella

Email: michela.vella.17@um.edu.mt

Signature: ______________________________ Date: ________________

Name: _________________________________

Appendix B.1: Recruitment letter – Maltese

Ittra ta' Reklutaġġ

Għażiż ġenitur/kustodju,

Jiena jisimni Michela Vella. Jiena studenta fl-Istitut tal-Lingwistika u t-Teknoloġija tal-Lingwa fl-Università ta' Malta. Bħala parti mill-Ph.D. tiegħi fil-Lingwistika, biħsiebni nirrekluta numru ta' tfal Maltin biex jipparteċipaw fi studju lonġitudinali li fih se nosservahom, ninteraġixxi magħhom u nirrekordjahom fuq bażi regolari għal total ta' 20 sa 24 xahar (sentejn). Qed nagħmel din ir-riċerka taħt is-superviżjoni prinċipali tal-Prof. Ray Fabri u l-kosuperviżjoni tal-Prof. Alexandra Vella.

L-istudju tiegħi jirrikjedi li t-tfal jilagħbu u jitkellmu b'mod ħieles u jipparteċipaw f'logħob u eżerċizzji marbutin mal-lingwa. Fil-bidu tal-istudju t-tfal se jkollhom bejn 18 u 36 xahar. Imbagħad inkompli insegwihom fuq bażi regolari, jiġifieri kull 4 sa 6 xhur, sabiex nosserva l-iżvilupp tal-lingwa tagħhom fuq medda ta' bejn 20 u 24 xahar (sentejn). Kull sessjoni se ddum madwar 30 minuta u se sseħħ f'ambjent informali li t-tifla/tifel tkun/ikun familjari miegħu. Ġenitur/kustodju għandu jkun preżenti mill-bidu sal-aħħar ta' kull sessjoni. Għall-finijiet ta' dan l-istudju, it-tfal se jkunu rrekordjati bl-użu ta' kemm awdjo kif ukoll video recorder biex niżgura li nkun nista' nittraskrivi dak kollu li jgħidu. Ir-recordings se jintużaw biss għal skopijiet ta' riċerka. Barra minn hekk, l-informazzjoni personali tiegħek se tinżamm f'post sigur li jien biss se jkolli aċċess għalih u kwalunkwe dettalji li jistgħu jwasslu għall-identifikazzjoni tiegħek jew ta' bintek/ibnek se jiġu anonimizzati qabel il-pubblikazzjoni ta' dan l-istudju. Jekk inti lest/a tagħti 45 minuta minn ħinek u ħin bintek/ibnek bosta drabi fuq medda ta' sentejn, napprezza ħafna jekk tipparteċipa f'dan il-proġett.

Inti tista' tibdel fehmtek dwar il-parteċipazzjoni tiegħek f'dan l-istudju f'kwalunkwe punt mingħajr ma tipprovdi raġuni. Jekk f'xi mument it-tifla/tifel tkun/ikun qed tħossha/iħossu skomdu jew taħt pressjoni, is-sessjoni tintemm immedjatament. Ladarba l-istudju jkun tlesta, inti tkun tista' taqra t-teżi tiegħi. Jekk tixtieq titlob għal aktar dettalji dwar l-istudju, toqgħodx lura milli tikkuntattjani billi tibgħat email fuq michela.vella.17@um.edu.mt jew tibgħat email lis-superviżur tiegħi, il-Professur Ray Fabri, fuq ray.fabri@um.edu.mt.

Grazzi bil-quddiem,

Michela Vella

Appendix B.2: Recruitment letter – English

Recruitment Letter

Dear parent/guardian,

My name is Michela Vella. I am a student at The Institute of Linguistics and Language Technology at the University of Malta. As part of my Ph.D. research in Linguistics, I intend to recruit a number of Maltese children to participate in a longitudinal study in which I will observe, interact with and record the children at regular intervals for a total of 20 to 24 months (2 years). I am conducting this research under the main supervision of Prof. Ray Fabri and the cosupervision of Prof. Alexandra Vella.

My study requires the children to engage in free play and conversation and to participate in language games and tasks. At the start of the study the children will be between 18 and 36 months old. I will follow them up at intervals of around 4 to 6 months so as to monitor their language development over a span of 20 to 24 months (2 years). Each session will last around 45 minutes and will take place in an informal setting which the child is familiar with. A parent/guardian must be present throughout. For the purposes of this study, the children will be recorded using both an audio and video recorder to ensure that I can transcribe all of their utterances. The recordings will be used solely for research purposes. Additionally, your personal information will be kept under lock and key and any details which might lead to you or your child's identification will be anonymised prior to publication. If you are willing and able to give up around 45 minutes of your and your child's time on a number of occasions over a span of 2 years, I would very much appreciate your participation in this project.

It will be possible for you to change your mind about participating in this study at any time without providing a reason. If at any point the child is feeling uncomfortable or pressured, the session will be ended immediately. Once the study has been completed, you will be more than welcome to read the thesis. If you wish to ask for more details about the study, do not hesitate to contact me by sending an email to michela.vella.17@um.edu.mt or emailing my supervisor, Professor Ray Fabri, at ray.fabri@um.edu.mt.

Thanks in advance,

Michela Vella

Appendix C.1: Consent form – Maltese

Formola ta' Kunsens

Jien, ____, ġenitur/kustodju ta'

____, nagħti l-permess tiegħi lil Michela Vella, studenta tal-Università ta' Malta, biex tosserva, tinteraġixxi ma' u tirrekordja lil binti/ibni diversi drabi għall-istudju tagħha dwar l-iżvilupp tal-Malti fit-tfal. Jien konxja/u li hi se tkun qed tuża awdjo recorder kif ukoll video recorder biex imbagħad tkun tista' tittraskrivi d-dejta. Jiena nagħtiha l-permess biex tuża d-dejta għal skopijiet ta' riċerka biss. Jiena konxja/u li d-dettalji personali li jistgħu jwasslu għall-identifikazzjoni tiegħi jew ta' binti/ibni, bħall-ismijiet tagħna, mhux se jintużaw fir-rappurtar tar-riżultati tar-riċerka u mhux se jsiru pubbliċi. Jiena konxja wkoll li d-dejta tista' tiġi ppublikata u użata għal aktar minn studju wieħed waqt, u anke wara, li Michela Vella tkun lestiet id-dottorat. Naqbel li d-dejta tiġi ppublikata (eż., bħala parti minn database, fuq sit elettroniku marbut mar-riċerka eċċ.) u użata biss wara li din tkun kompletament anonimizzata. Naf li nista' nwaqqaf il-parteċipazzjoni tiegħi f'kwalunkwe punt matul l-istudju mingħajr ma nagħti raġuni.

L-Università tirrikonoxxi d-dritt tiegħi, taħt l-Att dwar il-Protezzjoni tad-Dejta, li naċċessa d-dejta li tappartjeni għalija jew għal binti/ibni matul dan l-istudju, kif ukoll li nitlob li d-dejta tinqered jekk nixtieq.

Firma: ________________________________

Data: ____________________________________

Jien, Michela Vella, studenta tal-Università ta' Malta, inwiegħed li nuża d-dejta rrekodjata għal skopijiet ta' ricerka biss. Dettalji personali li jistgħu jwasslu għall-identifikazzjoni tal-parteċipanti mhux se jintużaw fir-rappurtaġġ tar-riżultati tar-riċerka tiegħi jew ippublikata.

Firma: ________________________________

Data: ____________________________________

Dettalji biex tikkunttatja r-riċerkatriċi

Isem: Michela Vella

Indirizz elettroniku: michela.vella.17@um.edu.mt

Appendix C.2: Consent form – English

Consent Form

I, ____, parent/guardian of ____, give my permission to Michela Vella, a student at the University of Malta, to observe, interact with and record my daughter/son on a number of occasions for her study about language development in Maltese children. I am aware that she will be using an audio recorder as well as a video recorder to later transcribe the data. I give my permission for the data to be used solely for research purposes. I am aware that personal details which might lead to the identification of myself or my child, such as our names, will not be used in reporting the results of the research and will not be made public. I am also aware that the data may be published and used for more than one study during, and even after, Michela Vella has completed her doctoral studies. I agree that the data will be published (e.g., as part of a database, on a website linked to research etc.) and used only after it has been completely anonymised. I know that I can stop my participation at any time without giving a reason.

The University recognises my right, under the Data Protection Act, to seek access to the data pertaining to myself/my child in the course of this study, as well as to request that the data be destroyed should I so wish.

Signature: ____________________________________

Date: ____

I, Michela Vella, a student at the University of Malta, promise to use the recorded data solely for the purposes of this study. Personal details which may lead to the identification of the participants will not be used in reporting the results of the research carried out.

Signature: ____________________________________

Date: ____

Contact information of researcher
Name: Michela Vella
Email address: michela.vella.17@um.edu.mt

Appendix D: Language background questionnaire

Kwestjonarju dwar l-Użu tal-Lingwa

Language Background Questionnaire

L-informazzjoni li se tingħabar permezz ta' dan il-Kwestjonarju dwar l-Użu tal-Lingwa hija maħsuba biex tgħinni, Michela Vella, nagħżel dawk il-parteċipanti li huma eliġibbli għall-istudju tiegħi.

Id-dettalji kollha mogħtija f'dan il-kwestjonarju, inkluż kwalunkwe informazzjoni personali, se jinżammu strettament kunfidenzjali.

Grazzi bil-quddiem tal-parteċipazzjoni tiegħek.

The information which will be collected through this Language Background Questionnaire is intended to help me, Michela Vella, select participants for my study with a view to selecting those who fit the eligibility criteria best.

All the details given in this questionnaire, including any personal information, will be treated in full confidentiality.

Thank you in advance for your participation.

A. Dettalji Personali

Personal details

1. Ismek u kunjomok: ______________________________

   Your name and surname: ______________________________

2. Kif tiġi mit-tifla/tifel (e.ż., ommha/u)? ________________

   How are you related to the child (e.g., mother)? ________________

3. Isem u kunjom tat-tifla/tifel: ________________________

   Name and surname of the child: ________________________

4. Età u data tat-twelid tat-tifla/tifel: ____________ ________________

   Child's age and date of birth: ____________ ________________

B. Esperjenza tal-Lingwi

Language experience

1. Agħti numru minn 1 (l-inqas livell) sa 4 (l-ogħla livell) biex tindika kemm tħossok komdu bil-Malti f'dawn l-erba' ħiliet tal-lingwa.

Tick a number from 1 (lowest level) to 4 (highest level) to indicate how comfortable you feel in Maltese in each of the four language skills (listening, speaking, reading, writing).

	1	2	3	4
Is-smigħ	○	○	○	○
It-taħdit	○	○	○	○
Il-qari	○	○	○	○
Il-kitba	○	○	○	○

2. Agħti numru minn 1 (l-inqas livell) sa 4 (l-ogħla livell) biex tindika kemm tħossok komdu bl-Ingliż f'dawn l-erba' ħiliet tal-lingwa.

Tick a number from 1 (lowest level) to 4 (highest level) to indicate how comfortable you feel in Maltese in each of the four language skills (listening, speaking, reading, writing).

	1	2	3	4
Is-smigħ	○	○	○	○
It-taħdit	○	○	○	○
Il-qari	○	○	○	○
Il-kitba	○	○	○	○

3. Bejn wieħed u ieħor, ta' liema età bdejt tuża l-lingwi attwali tiegħek?

Malti: ______________

Ingliż: ______________

At what approximate age did you start using your current language/s?

Maltese: ______________

English: ______________

C. L-Użu tal-Lingwa

Language Use

1. Jekk jogħġbok immarka liema minn dawn jiddeskrivi l-aħjar l-użu tal-lingwa tiegħek ma':

Kindly mark the description which best fits your language use with:

	Familtek Family	Sħabek Friends	Bintek/Ibnek Your child
Malti biss Maltese only			
Malti u ftit Ingliż Maltese and some English			
Bejn wieħed u ieħor Malti u Ingliż indaqs Approximately equal use of Maltese and English			
Ingliż u ftit Malti English and some Maltese			
Ingliż biss English only			

2. Liema lingwa tippreferi tuża meta:

   What is your preferred language when:

	Malti Maltese	Ingliż English	Lingwa oħra Other language
Tisma' programmi fuq ir-radju Listening to radio programmes			
Tara t-televiżjoni Watching television			
Tuża l-mezzi soċjali Using social media			
Tikteb nota/messaġġ informali Writing an informal note/message			

3. Tuża djalett? Jekk iva, liema? ______________________

Do you speak a dialect? If yes, which one? ______________________

4. Minn 1 (l-inqas) sa 5 (l-iktar), kemm taħseb li t-tifla/tifel hija/huwa esposta/espost għall-**Malti** matul il-ġurnata? __________

On a scale of 1 (lowest) to 5 (highest), to what extent do you think the child is exposed to **Maltese** on a daily basis? __________

5. Minn 1 (l-inqas) sa 5 (l-iktar), kemm taħseb li t-tifla/tifel hija/huwa esposta/espost għall-**Ingliż** matul il-ġurnata? __________

On a scale of 1 (lowest) to 5 (highest), to what extent do you think the child is exposed to **English** on a daily basis? __________

6. Jekk jogħġbok, imla din it-tabella għal dawk il-persuni li t-tifla/tifel t/jqatta' magħhom mill-inqas **2 sigħat** fil-ġimgħa, billi tindika liema lingwa jintuża l-aktar meta jinteraġixxu mat-tifla/tifel.

Kindly complete this table for people the child spends a minimum of **2 hours** a week with, indicating the language usually used when they interact with the child.

	Hin flimkien matul il-ġimgħa Time spent together per week	Persentaġġ (bejn wieħed u ieħor) Rough estimate percentage
Omm Mother	☐ Inqas minn 5 sigħat / Less than 5 hrs ☐ 5–15 ☐ 16–30 ☐ Aktar minn 30 / More than 30	Malti: _____ Ingliż: _____ Maltese:____English: _____
Missier	☐ <5 ☐ 5–15 ☐ 16–30 ☐ >30	Malti: _____ Ingliż: _____

Father		Maltese: ____ English: ______
Nanna min-naħa tal-omm Maternal grandmother	☐ <5 ☐ 5–15 ☐ 16–30 ☐ >30	Malti: ______ Ingliż: ______ Maltese: ____ English: ______
Nannu min-naħa tal-omm Maternal grandfather	☐ <5 ☐ 5–15 ☐ 16–30 ☐ >30	Malti: ______ Ingliż: ______ Maltese:____ English: ______
Nanna min-naħa tal-missier Paternal grandmother	☐ <5 ☐ 5–15 ☐ 16–30 ☐ >30	Malti: ______ Ingliż: ______ Maltese: ____ English: ______
Nannu min-naħa tal-missier Paternal grandfather	☐ <5 ☐ 5–15 ☐ 16–30 ☐ >30	Malti: ______ Ingliż: ______ Maltese: ____ English: ______
Aħwa Siblings	☐ <5 ☐ 5–15 ☐ 16–30 ☐ >30	Malti: ______ Ingliż: ______ Maltese: ____ English: ______
Nanny Nanny	☐ <5 ☐ 5–15 ☐ 16–30 ☐ >30	Malti: ______ Ingliż: ______ Maltese: ____ English: ______
Nies li jaħdmu fid-day care Day care staff	☐ <5 ☐ 5–15 ☐ 16–30 ☐ >30	Malti: ______ Ingliż: ______ Maltese: ____ English: ______
Persuni oħra Others	☐ <5 ☐ 5–15 ☐ 16–30 ☐ >30	Malti: ______ Ingliż: ______ Maltese: ____ English: ______

7. Jekk jogħġbok, imla t-tabella skont is-sitwazzjoni.

Kindly complete this table according to the situation.

Sitwazzjoni Situation	Malti biss Maltese only	Ingliż biss English only	It-tnejn indaqs Both equally	L-aktar bil-Malti Mostly Maltese	L-aktar bl-Ingliż Mostly English
Meta nitkellem mat-tifla/tifel When I speak to the child					
Meta nitkellem ma' membri oħra tal-familja When I speak to other family members					
Meta t-tifla/tifel t/jara t-televiżjoni When the child watches TV					

Meta t- tifla/tifel t/juża apparat elettroniku (eż. tablet, mowbajl, kompjuter) When the child uses electronic devices (e.g., tablet, phone, computer)					
Meta jinqraw il-kotba lit-tifla/tifel When books are read to the child					

Appendix E: Child development questionnaire

Kwestjonarju dwar l-Iżvilupp tat-Tifla/Tifel

Child's Development Questionnaire

L-informazzjoni li se tinġabar permezz ta' dan il-Kwestjonarju dwar l-Iżvilupp tat-Tifla/Tifel hija maħsuba biex tgħinni, Michela Vella, nagħżel dawk il-parteċipanti li huma eliġibbli għall-istudju tiegħi.

Id-dettalji kollha mogħtija f'dan il-kwestjonarju, inkluż kwalunkwe informazzjoni personali, se jinżammu strettament kunfidenzjali.

Grazzi bil-quddiem tal-parteċipazzjoni tiegħek.

The information which will be collected through this Language Background Questionnaire is intended to help me, Michela Vella, select participants for my study with a view to selecting those who fit the eligibility criteria best.

All the details given in this questionnaire, including any personal information, will be treated in full confidentiality.

Thank you in advance for your participation.

A. Informazzjoni dwar it-tifla/tifel u familtha/u

A. Child's background information

Isem u kunjom tat-tifla/tifel :______________________________

Child's name and surname: ______________________________

Data tat-twelid tat-tifla/tifel: ______________________________

Child's date of birth: ______________________________

Indirizz residenzjali tat-tifla/tifel: ____________________________________

Child's residential address: ____

Sess tat-tifla/tifel: ___________________________

Child's sex: ___________________________

L-età tat-tifla/tifel: ___________________________

Child's age: ___________________________

Isem u kunjom tal-omm: ______________________

Mother's name and surname: _____________________________

Età tal-omm: ___________________________

Mother's age: _____________________________

Raħal/belt twelidha: ___________________________

Mother's hometown: _____________________________

Livell ta' edukazzjoni tal-omm: ___________________________

Mother's level of education: _____________________________

Xogħol tal-omm: ________________________

Mother's occupation: _____________________________

Omm it-tifla/tifel qatt għexet barra minn Malta għal sena jew iktar? Iva Le

Has the child's mother lived abroad for a period of one year or more? Yes No

Jekk 'Iva', jekk jogħġbok semmi l-pajjiż. _________________________

If 'Yes', kindly specify the country. _____________________________

Isem u kunjom tal-missier: ______________________

Father's name and surname: _____________________________

Età tal-missier: ___________________________

Father's age: _____________________________

Raħal/belt twelidu: ___________________________

Father's hometown: _____________________________

Livell ta' edukazzjoni tal-missier: ___________________________

Father's level of education: _____________________________

Xogħol tal-missier: ________________________

Father's occupation: _____________________________

Missier it-tifla/tifel qatt għex barra minn Malta għal sena jew iktar? Iva Le

Has the child's father lived abroad for a period of one year or more? Yes No

Jekk 'Iva', jekk jogħġbok semmi l-pajjiż. _________________________

If 'Yes', kindly specify the country. _____________________________

It-tifla/tifel għandha/u aħwa? Iva Le

Does the child have siblings? Yes No

Jekk 'Iva', jekk jogħġbok niżżel l-età u s-sess tal-aħwa kollha.

____

If 'Yes', kindly specify the age and sex of each sibling.

____

Jekk 'Iva', ħutha/u jmorru skola? Iva Le

If 'Yes', does the sibling/s attend school? Yes No

Jekk 'Iva', jekk jogħġbok semmi l-isem/ismijiet tal-iskola/skejjel

____

If 'Yes', kindly specify the name/s of the school/s.

____

Hemm persuni oħra li qed jgħixu mat-tifla/tifel? Iva Le

Jekk 'Iva', semmi min. ______________________________

Are there any other persons residing with the child? Yes No

If 'Yes', kindly specify. ______________________________

Hemm xi ħadd fil-familja li għandha/u problemi biex t/jitkellem/ t/jifhem/ t/jinftiehem meta t/jikkomunika? Iva Le

Jekk 'Iva', elabora. ______________________________

Are there any family members who have difficulty speaking, understanding or being understood when communicating? Yes No

If 'Yes', kindly specify. ______________________________

Hemm xi ħadd fil-familja li huwa trux jew għandu problemi ta' smigħ? Iva Le

Jekk 'Iva', elabora. ______________________________

Are there any family members who suffer/ed from childhood deafness/hearing impairment? Yes No

If 'Yes', kindly specify. ______________________________

B. Il-funzjoni komunikattiva tat-tifla/tifel bħalissa

B. Child's current communicative functioning

Bħalissa t-tifla/tifel jikkomunika:

bil-ħsejjes biss | bil-mossi biss | bil-ħsejjes u l-mossi

bil-kliem biss | bil-kliem u l-mossi

The child currently communicates:

solely by babbling | solely by gesturing | by babbling and gesturing

through verbal communication | both verbal communication and gesturing

Bħalissa t-tifla/tifel qiegħed fi stadju fejn juża:

kelma waħda biss f'sekwenza | żewġ kelmiet f'sekwenza

tliet kelmiet jew aktar f'sekwenza

The child is currently at the:

single-word stage | two-word stage | multi-word stage

C. Stadji ta' żvilupp

C. Developmental milestones

Kemm kellha/u żmien it-tifla/tifel meta bdiet/beda t/jimxi? ____________

At what age did the child start to walk? ____________________

Kemm kellha/u żmien it-tifla/tifel meta qalet/qal l-ewwel kelma? ____________

At what age did the child utter his/her first words? ___________________

X'kienu l-ewwel kelmiet tat-tifla/tifel?

____

What were the child's first words? _______________________________________

It-tifla/tifel t/jixrob minn flixkun jew tazza? ________________

Does the child drink from a bottle or a cup? ________________

It-tifla/tifel t/jiekol ikel solidu? Iva Le

Does the child eat solid food? Yes No

It-tifla/tifel t/juża gażaża? Iva Le

Does the child use a pacifier? Yes No

It-tifla/tifel t/jitma lilu nnifisha/u b'mgħarfa minkejja li forsi t/jaqleb ftit mill-ikel?

Iva Le

Does the child feed herself/himself with a spoon even though s/he may spill some food?

Yes No

It-tifla/tifel potty trained? Iva Le

Is the child potty trained? Yes No

It-tifla/tifel t/jista' t/jirrepeti sentenza magħmula minn żewġ kelmiet? Iva Le

Does the child imitate a two-word sentence? Yes No

D. Informazzjoni prenatali u perinatali

D. Prenatal and perinatal history

Kemm damet it-tqala? ____________________

What was the duration of the gestation period? ______________________

Kien hemm kumplikazzjonijiet waqt it-tqala/twelid? Iva Le

Jekk 'Iva', jekk jogħġbok elabora. _____________________________________

Were there any complications during the pregnancy/at birth? Yes No

If 'Yes', kindly specify. ____

It-tifla/tifel twieldet b'mod normali? Iva Le

Was the child born through normal delivery? Yes No

E. Informazzjoni medika

E. Medical history

It-tifla/tifel tbati minn kundizzjonijiet mediċi? Iva Le

Jekk 'Iva', jekk jogħġbok niżżilhom. _____________________________

Does the child suffer from any medical conditions? Yes No

If 'Yes', kindly specify. ____

It-tifla/tifel qatt kellha/u jinżamm l-isptar? Iva Le

Jekk 'Iva', jekk jogħġbok elabora. _______________________________

Has the child ever been hospitalised? Yes No

If 'Yes', kindly specify. ____

It-tifla/tifel t/jbati minn xi allerġiji? Iva Le

Jekk 'Iva', jekk jogħġbok niżżilhom. ______________________

Does the child suffer from any allergies? Yes No

If 'Yes', kindly specify. ______________________________

It-tifla/tifel qatt kellha/u infezzjoni f'widnejha/widnejh/glue ear? Iva Le

Jekk 'Iva', jekk jogħġbok elabora. ______________________

Did the child ever have any episodes of an ear infection/glue ear? Yes No

If 'Yes', kindly specify. ______________________________

F. Mistoqsijiet marbutin mad-drawwiet u l-imġiba tat-tifla/tifel

F. Questions related to habits and behaviour

It-tifla/tifel t/jirrispondi meta tgħajtilha/tgħajjatlu minn kamra oħra? Iva Le

Does the child respond when called from another room? Yes No

It-tifla/tifel għandha/u tendenza li ma t/joqgħodx bi kwietha/u jew t/iqum spiss matul il-lejl? Iva Le

Does the child tend to be restless or wake up often during the night? Yes No

It-tifla/tifel t/jmur day care centre? Iva Le

Jekk 'Iva', bejn wieħed u ieħor, kemm t/jqatta' sigħat id-day care centre. _____

Does your child attend a day care centre? Yes No

If 'Yes', approximately how many hours a week does your child spend at the day care centre. _____

Bejn wieħed u ieħor, kemm-il siegħa t/jqatta' t-tifla/tifel quddiem skrin f'ġurnata. _____

Approximately how many hours a day does the child spend in front of a screen. _____

It-tifla/tifel tiegħek t/jieħu gost t/jinteraġixxi ma' tfal oħra? Iva Le

Does your child enjoy interacting with other children? Yes No

It-tifla/tifel tiegħek t/jieħu gost t/jinteraġixxi ma' adulti? Iva Le

Does the child enjoy interacting with adults? Yes No

Jirnexxilek tifhem il-maġġorparti ta' dak li t/jgħid it-tifla/tifel? Iva Le

Can you understand most of what the child says? Yes No

Taħseb li t-tifla/tifel t/jimxi/ t/jiġri/ t/jitla' bħal tfal żgħar oħra tal-età tagħha/tiegħu? Iva Le

Do you think the child walks/runs/climbs like other toddlers her/his age? Yes No

Tinsab imħassba/imħasseb dwar l-imġiba tat-tifla/tifel? Iva Le

Do you have any concerns about the child's behaviour? Yes No

X'inhuma l-ġugarelli favoriti tat-tifla/tifel? ______________________________

What are your child's favourite toys? ______________________________

It-tifla/tifel tiegħek x't/jieħu gost t/jagħmel l-aktar? ______________________________

What does your child enjoy doing most? ______________________________

Appendix F: Child assent form

<u>Trid tipparteċipa fl-istudju tiegħi?</u>

Jien se nkun qed nirrekordjak b'awdjo recorder u video recorder waqt li nilagħbu u nitkellmu flimkien. Imbagħad se mmur id-dar u se nikteb dak kollu li nkunu għidna.

Tifla/tifel:

☐ **Jiena fhimt x'se nkun qed nagħmel.**

Jiena nixtieq nipparteċipa fl-istudju ta' Michela Vella.

Jiena ma nixtieqx nipparteċipa fl-istudju ta' Michela Vella.

Isem tat-tifla/tifel: ______________________________

Data: ________________________

Appendix G: Questions for semi-structured interview with a parent

Mistoqsijiet li se ssaqsi r-riċerkatriċi

Questions to be asked by the researcher

- Min taħseb li hi/hu l-Mudell Primarju tal-Lingwa tat-tifla/tifel bħalissa?
  Who do you think is currently the child's Primary Language Model?
  (Primary Language Model = the person the child spends most time with in conversation)

- Taħseb li l-Mudell Primarju tal-Lingwa tat-tifla/tifel inbidel matul dawn l-aħħar sentejn?
  Do you think the child's Primary Language Model changed over the past two years?

- Tuża djalett? Jekk iva, liema? Jekk iva, mat-tifla/tifel titkellem bid-djalett?
  Do you use a dialect? If yes, which? If yes, do you speak in dialect with your child?

- Hemm persuni oħra fil-familja li jużaw djalett? Jekk iva, min? Jekk iva, mat-tifla/tifel jitkellmu bid-djalett?
  Do any other family members use a dialect? If yes, which? If yes, do they speak in dialect with the child?

- Qatt osservajt lit-tifla/tifel tuża/juża djalett? Jekk iva, elabora.
  Have you ever observed your child using a dialect? If so, elaborate.

- X'inhi l-opinjoni tiegħek dwar l-użu tad-djalett f'Malta u Għawdex?
  What is your opinion on the use of dialects in Malta and Gozo?

- Meta titkellem mat-tifla/tifel taqleb ta' spiss bejn il-Malti u l-Ingliż?
  Do you frequently switch between Maltese and English when speaking to your child?

- Taħseb li d-dominanza u l-użu tat-tifla/tifel tal-Malti u l-Ingliż inbidlu matul dawn l-aħħar sentejn? Jekk iva, elabora.
  Do you think the child's dominance and use of Maltese and English changed over the course of the past two years? If so, elaborate.

- Taħseb li ċ-childcare/skola għandha effett fuq id-dominanza u l-użu tal-lingwa tat-tifla/tifel? Jekk iva, elabora.
  Do you think childcare/school has an effect on your child's language dominance and use? If so, elaborate.

- Taħseb li l-preżenza u/jew l-influwenza ta' tfal oħra f'ħajjet it-tifla/tifel (aħwa/kuġini/ħbieb) għandhom effett fuq id-dominanza u l-użu tal-lingwa tat-tifla/tifel?
  Do you think the presence and/or influence of other children in the child's life (siblings/cousins/friends) have an effect on your child's language dominance and use? If so, elaborate.

- Tinsab imħassba/imħasseb dwar l-iżvilupp tal-lingwa tat-tifla/tifel?
  Are you concerned about any aspects of your child's language development?

- Tinsab imħassba/imħasseb dwar l-iżvilupp tat-tifla/tifel b'mod ġenerali?
  Are you concerned about any aspects of your child's general development?

- It-tifla/tifel tiegħek qatt ġiet/ġie referuta/referut għand patoloġista tad-diskors u tal-lingwa (speech and language pathologist/speech therapist)? Jekk iva, spjega.
  Was your child ever referred to a speech and language pathologist? If yes, specify.

- It-tifla/tifel tiegħek qatt attendiet/attenda sessjonijiet tal-patoloġija tad-diskors u tal-lingwa (speech and language pathology/speech therapy sessions)? Jekk iva, spjega.
  Did your child ever attend speech and language pathology sessions? If yes, specify.

Appendix H: Representative sample of toys presented

The images in this appendix illustrate examples of the toys and materials used in the study, or items as closely resembling them as possible. They are reproduced under fair use for academic and illustrative purposes. A numbered list of corresponding image sources is provided at the end of this appendix.

1

2

3

4

6

5

7

9

8

10

12

11

List of Image Sources:

Dog hand-puppet – Retrieved from https://m.media-amazon.com/images/I/81DhjIYUv-L.AC_SY606.jpg
Felt pizza – Retrieved from https://www.melissaanddoug.com/...
Wooden furniture – Retrieved from https://m.media-amazon.com/images/I/71UYn8JKhuL.jpg
Playmobil playground – Retrieved from https://i5.walmartimages.com/...
Playmobil ski lodge – Retrieved from https://www.amazon.co.uk/...
Playmobil adventure treehouse – Retrieved from https://modelshop.com.mt/...
City roads mat – Retrieved from https://media.4rgos.it/...
Two-storey parking lot – Retrieved from https://www.webby.toys/...
Toy vehicles – Retrieved from https://m.media-amazon.com/...
Animal figurines – Retrieved from https://m.media-amazon.com/...
Noah's Ark wooden set – Retrieved from https://isaacstreasures.com/...
Farmer's market sorting set – Retrieved from https://i5.walmartimages.com/...

Appendix I: Protocol instructions

These are examples of brief researcher scripts used to introduce and guide participants through each experimental task described in Section 3.3.2.2.

Colour-sorting activity

Tista' tpoġġi l-affarijiet oranġjo kollha hawn ġew?
'Can you put all the orange things in here?'
(Researcher pointing at a plastic bowl)

L-oranġjo biss. L-oħrajn agħmilhom hawnhekk.
'Just the orange ones. Put the rest here.'
(Researcher pointing at another plastic bowl, both in full view and reach of the child).

Quantifier comprehension activity (one vs all)

Hawn erbgħa minn kull animal.
'There are four of each animal.'

Issa l-kelb se jgħidlek liema annimali tagħmel fuq il-vapur, ok?
'Now the dog is going to tell you which animals to put on the ship, ok?'

Tista' tpoġġi ġiraffa waħda fuq il-vapur?
'Can you put one giraffe on the ship?'
(Researcher using dog hand-puppet)

Issa tista' tagħmel l-iljunfanti kollha?
'Now can you put all the elephants?
(Researcher using dog hand-puppet)

Real-word plural task

Ara, ballun! ‘Look, a ball!’

(Visual stimulus – 1 ball – appearing on the screen)

Issa hawn ħafna _______. ‘Now there are many _______.’

(Visual stimulus – 4 balls – appearing on the screen)

Real-word plural task – Bingo

Ha nurik kif tilgħabha din. Jien għandi sett stampi u int għandek sett differenti.

‘I’ll show you how to play. I have a set of pictures, and you have a different set.’

(Research pointing at flashcard packs)

Kull darba li ngħidu waħda, nimmarkawha t-tnejn.

‘Every time a picture is called out, we both get to mark it.’

(Research pointing at markers)

Min jgħaqqad line l-ewwel jirbaħ.

‘The first person to get a full line wins.’

Jista’ jkun line hekk jew line hekk.

‘It can be a line going this way, or this way.’

(Researcher indicating vertical or horizontal on the bingo sheet)

Ha ngħidlek waħda jien l-ewwel. Imbagħad imiss lilek tgħidli waħda minn tiegħek.

‘I'll call out one of my pictures first, and then it's your turn to call out one of yours.’

L-ewwel wieħed minn tiegħi huwa… fenek. *Sibu fuq tiegħek u mmarkah.*

‘My first one is... a rabbit! Find it on your board and mark it.’

(Researcher revealing flashcard with an image of a rabbit and showing it to the child)

Novel-word plural task

Mela dal-kelb mar fuq pjaneta differenti.

'So, this dog went to a different planet.'

(A planet in outer space appearing on the screen)

U meta wasal hemm iltaqa' ma' ħafna affarijiet strambi, bħal annimali strambi.

'And when he got there, he met lots of really strange things, like weird animals.'

Qatt ma rajna bħalhom qabel.

'We've never seen anything like them before.'

Jien ħa ngħidlek x'inhu wieħed u int se tgħidli ħafna x'ngħidulhom, ok?

'I'm going to tell you what one is and you're going to tell me what lots of them are called, ok?'

M'hemmx tajjeb u ħażin għax qatt ma smajna bihom dawn.

'There's no right or wrong answer because we've never heard of them before.'

Mela ara, il-kelb qal "Ara, lepa".

'So look, the dog said, "Look, a lepa".'

(Visual stimulus – 1 strange creature – appearing on the screen)

Issa hawn ħafna ___________.

'Now there are many _______________.'

(Visual stimulus – 4 of the same strange creature – appearing on the screen)

Opposites task

X'qed tara hawnhekk?

'What do you see here?'

(One clean monkey and one dirty monkey appearing on the screen)

X'hemm differenti bejniethom?

'How are they different?'

(Researcher pointing at the monkeys on the screen)

Spot-the-difference

Din qisha logħba, ara.

'Look, this is like a game.'

(Researcher placing the two laminated sheets in front of the child)

Għandna dawn l-istampi. Fihom ħafna affarijiet kważi l-istess imma hawn ftit affarijiet differenti.

'We have these two pictures. In them, lots of things are almost the same, but a few things are different.'

Ħa naraw kemm se ssib differenzi inti.

'Let's see how many differences you'll find.'

Narrative Elicitation

Tista' tgħidli x'qed jiġri hawnhekk? 'Can you tell me what's happening here?'

(Researcher pointing at the narrative visual stimuli and using dog hand-puppet)

X'hawn iktar? 'What else is there?'

U issa, x'ġara? 'And now, what happened?'

Appendix J: Real-word plural task items

The criteria for item selection and the corresponding list of items are provided below. All stimuli followed the same layout and visual format. A representative example of the visual display is shown in Figure J.1.

The full version features a total of 20 test items. The items selected are:

a) Easy to depict, integral child-language vocabulary items, appearing in one or both of two early-years vocabulary lists (Gatt, 2010; Muscat et al., 2017). The following did not feature in these lists but were also used: (i) *tablet* 'tablet' (electronic device), now widely used in Maltese; (ii) *mistoqsija* 'question', not included presumably because both lists mainly consist of concrete nouns; and (iii) *papru* 'drake – male duck', not included although both lists include *papra* '(female) duck'. To date, no research has examined word/POS frequency in Maltese CDS.

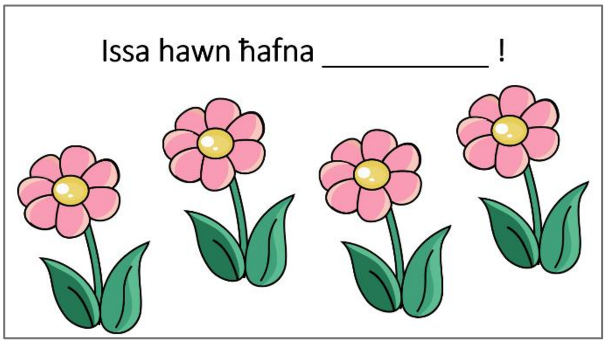
b) Equally divided between masculine and feminine nouns in the singular. One noun is marked for the feminine with the suffix *-a* (*kelb* '(male) dog', *kelba* 'female dog') and one for the masculine with the suffix *-u* (*papra* '(female) duck', *papru* 'drake – male duck'); the rest are unmarked. Note that grammatical gender does not determine whether a word takes SP or BP.

c) Equally divided between those that take SP and BP. The SP suffixes targeted include *-i* (e.g., *platt–platti* 'plates'), *-ijiet* (e.g., *siġġu–siġġijiet* 'chairs'), and *-s* (e.g., *kamera–kameras* 'cameras'), and the BP patterns targeted include CCVVCVC (e.g., *blalen* 'balls', the plural of *ballun*), CCVVC (e.g., *klieb* 'dogs', the plural of *kelb*), and CVCVC (e.g., *stilel*, the plural of *stilla*) – intuitively high-frequency suffixes and patterns in child-language vocabulary. To date, no research has examined the frequency of pluralisation patterns in spoken Maltese or in Maltese child language.

List of items presented in the order of appearance in the full real-word plural task:

1. *fjura – fjuri* 'flowers'
2. *ballun – blalen* 'balls'
3. *kalzetta – kalzetti* 'socks'
4. *lapes – lapsijiet* 'pencils'
5. *borma – borom* 'pots'
6. *siġġu – siġġijiet* 'chairs'
7. *tablet – tablets*
8. *kelba – klieb* 'dogs'
9. *ġakketta – ġkieket* 'jackets'
10. *ktieb – kotba* 'books'
11. *platt – platti* 'plates'
12. *mistoqsija – mistoqsijiet* 'questions'
13. *serp – sriep* 'snakes'
14. *farfett – friefet* 'butterflies'
15. *nanna – nanniet* 'grandmothers'
16. *papru – papri* 'ducks'
17. *ħabel – ħbula* 'ropes'
18. *mejda – imwejjed* 'tables'
19. *kamera – kameras* 'cameras'
20. *stilla – stilel* 'stars'

Figure J.1 *Screenshot of a real-word plural task trial. Each screen displayed a carrier phrase and four identical, easily recognisable clip-art-style images. All images were sourced from open-access repositories that allow free reuse for research and educational purposes.*



Appendix K: Novel-word plural task items

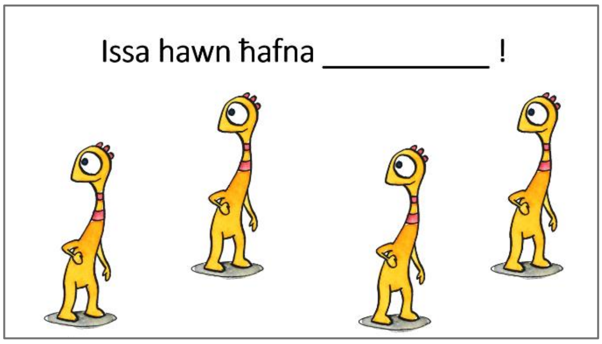
The full set of items used in the task is listed below. All stimuli followed the same layout and visual format. A representative example of the visual display is shown in Figure K.1.

All novel words were phonotactically legal singular nouns in Maltese. They were originally constructed by Nieder et al. (2020) and are reproduced here with the authors' kind permission.

List of items presented in the order of appearance in the novel-word plural task:

1. *lepa* (*nota* – *noti* 'notes')
2. *qetties* (*qattus* – *qtates* 'cats')
3. *peġġa* (*tazza* – *tazzi/tazez* 'glasses')
4. *vieren* (*żiemel* – *żwiemel* 'horses')
5. *gadma* (*dabra* – *dabriet/dbabar* 'ulcers/sores/patches')
6. *torħa* (*terħa* – *terħiet/trieħi/teraħ* 'sashes')
7. *seff* (*saff* – *saffijiet/safef* 'layers')
8. *nedimist* (*laburist* – *laburisti* 'labourites')
9. *qonda* (*kelba* – *klieb* 'dogs')
10. *tampi* (*parti* – *partijiet* 'parts')
11. *pertut* (*partit* – *partiti* 'matches/games/parties')
12. *rulipa* (*minuta* – *minuti* 'minutes')
13. *fana* (*sala* – *swali* 'halls')
14. *kerem* (*qamar* – *qmura* 'moons')
15. *delq* (*bank* – *bankijiet/banek* 'benches/banks')
16. *rebta* (*rabta* – *rabtiet* 'ties/agreements')
17. *bellin* (*ballun* – *blalen* 'balls')
18. *pannat* (*tallab* – *tallaba/tlaleb* 'beggars')
19. *kaffif* (*qassis* – *qassisin* 'priests')
20. *tekkina* (*takkuna* – *tkaken* 'heels')

Figure K.1 *Example trial screen from the novel-word plural task showing the carrier phrase and four identical images. The novel images used in this task were those originally created for van de Vijver and Baer-Henney (2014), later employed by Nieder et al. (2020), and used here with their kind permission.*



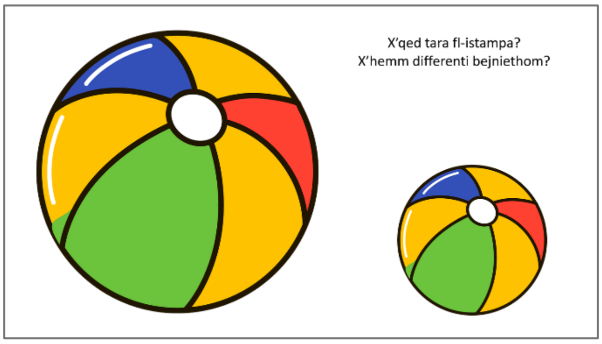
Appendix L: Opposites task items

The full set of items used in the task is listed below. All stimuli followed the same layout and visual format. A representative example of the visual display is shown in Figure L.1.

List of items presented in the order of appearance in the novel-word plural task:

1. ballun kbir – ballun żgħir 'a big ball – a small ball' (ball.SGM)
2. qattus oħxon – qattus irqiq 'a fat cat – a thin cat' (cat.SGM)
3. raġel twil – raġel qasir 'a tall man – a short man' (man.SGM)
4. karozza ġdida – karozza antika/qadima 'a new car – an old car' (car.SGF)
5. kaxxa miftuħa – kaxxa magħluqa 'a closed box – an open box' (box.SGF)
6. xadina nadifa – xadina maħmuġa 'a clean monkey – a dirty monkey' (monkey.SGF)
7. basket ġdid – basket antik/qadim 'a new bag – an old bag' (bag.SGM)
8. mejda kbira – mejda żgħira 'a big table – a small table' (table.SGF)
9. baqra ħoxna – baqra irqiqa 'a fat cow – a thin cow' (cow.SGF)
10. bieb magħluq – bieb miftuħ 'a closed door – an open door' (door.SGM)
11. mara twila – mara qasira 'a tall woman – a short woman' (woman.SGF)
12. ħanżir nadif – ħanżir maħmuġ 'a clean pig – a dirty pig' (pig.SGM)

Figure L.1 *Screenshot of an opposites task trial. Each screen displayed two question prompts and easily recognisable clip-art-style images. All images were sourced from open-access repositories that allow free reuse for research and educational purposes.*



Appendix M: Sample transcription conventions and post-transcription cleaning procedures

This appendix provides a sample overview of the transcription conventions and post-transcription data-cleaning procedures applied prior to MLUm calculation and analysis. The sample is representative of the conventions followed throughout.

A. Transcription Conventions

A.1 Structure and Formatting

- Font: Never bold or underlined. Italics used only as specified (Column 4 below).
- Metadata fields:
    - #age_in_months
    - #child_s_initials
    - #age_day_of_recording (years;months.days)
    - #recording_date (day.month.year)
    - #recording_duration
    - #relevant_information
    - #activities
    - #abbreviations (#INV = investigator; #CHI = child; #MOT = mother; #FAT = father; #SIB = sibling)
- Table layout: four transparent columns

Column	Content	Description
1	Duration marker	Every five minutes (e.g., 05.00, 10.00, 15.00) marked on the second.
2	Speaker	Investigator, child, mother, father, or sibling. Child utterances numbered sequentially (CHI1, CHI2…).
3	Utterance	Lower case unless grammar requires a capital (e.g., proper nouns).
4	Comments	Brief notes in English only when necessary, typically referring to video data (e.g., *CHI points at green car*). Orthographic target items appear in *italics* for clarification.

Practical notes:

Check child-utterance numbering, maintain consistent speaker tags, and insert additional rows as needed without altering existing time codes.

A.2 Data Sensitivity and Anonymisation

- All recordings and transcriptions are private and confidential. Participant data may not be shared outside the research team and must be deleted from personal devices once transcription is complete.
- Anonymisation rule: first letter of the participant's name followed by the remaining CV pattern (e.g., *Lippu* → *Lvccv*).

A.3 Annotation Conventions

Table L.1 *Annotation Symbols Used in the Transcription Corpus*

Symbol	Definition and Use
[[]]	Overlapping speech
(.)	Prosodic/intonational boundary (pause length not specified)
…	False start/repair/phrase cut-off (e.g., En. whe…; Mt. me…)
'	Ellipsis (e.g., En. f'ower ~ flower; Mt. *haw'hekk* ~ *hawnhekk* 'here')
//	Non-Maltese item (e.g., Mt. *ġib it-/toy/* 'get the toy')
[]	Insertion/variation (e.g., En. s[n]eep ~ sleep; Mt. *ka[j]ozza* ~ *karozza* 'car')
<>	Non-verbal vocalisation/noise (e.g., <sneeze>)
()	Uncertain but identified item
(?)/???	Uncertain unidentified item/sequence
.	Pragmatically complete utterance/long final pause
?	Question structure
“”	Direct speech (speaker assumes another voice)
_	Elicitation cue (e.g., Mt. *il-ktieb il-*_ 'the _ book')

A.4 Notes on Special Symbols and Pauses

- Stress marks are used only where needed to distinguish identical orthographic forms with different stress (e.g., *'dadi* 'dice' vs *da'di* 'father').
- Pauses and punctuation: (.) indicates brief pause; final periods mark pragmatically complete utterances; no pause is marked after ?.
- Examples of overlaps and mixed-language utterances:
  - Full overlap: [[/I went to the beach/]]
  - Partial overlap: /I went [[to the]] beach/
  - Mixed Maltese–English question: *il-/green/* 'the green'
- Examples of fillers and onomatopoeic items: *mm, aa, eqq, wijj, bumm*
- Examples of non-verbal sounds: <laughter>, <gasp>, <cough>
- Length of speech sounds not reflected orthographically (e.g., *wow* not *wooow*).

A.5 Lexical and Activity Labels

- Example toy names: *dog hand-puppet, Barbie dress-up game, city roads mat.*
- Activity names: *colour-sorting task, novel-word plural task, opposites task.*

B. Post-Transcription Data Cleaning

This stage ensured that only spontaneous, meaningful child utterances were retained for MLUm calculation. Utterances meeting any of the criteria below were removed in their entirety.

B.1 Exclusion Criteria

- Yes/no responses in Maltese or English (e.g., *iva* 'yes', *le* 'no', *yep, yes, no*).
- Utterances consisting solely of fillers (e.g., *eqq, mm*).
- Incomplete or aborted utterances containing only truncated forms (e.g., *whe…*).
- Utterances composed only of non-verbal sounds (e.g., <gasp>).
- Unintelligible speech, marked as (?) for single items or ??? for longer sequences.
- Immediate full adult repetitions, where the child repeats the adult's preceding utterance verbatim.

Note. All transcriptions were produced and annotated in Praat (Boersma & Weenink, v. 6.x) following standard TextGrid procedures. The full transcription manual and transcriber training materials are available upon request.

Appendix N: Sample MLUm Calculation and Scoring System

This appendix provides a sample overview of the scoring system and calculation procedures used in determining MLUm. It illustrates how morpheme values were assigned and how specific within-utterance elements were treated. The conventions presented here are representative of the procedures applied throughout the dataset.
Definitions of 'utterance' and 'morpheme' are provided in Sections 3.6.1. and 3.6.2.

A. Note on Within-Utterance Treatment

As detailed in Appendix M, entire utterances were excluded when they met criteria for omission during transcription cleaning. For analysis, the internal structure of all retained utterances was otherwise preserved. However, the following within-utterance elements were excluded from the morpheme count (i.e., assigned a score of *0*):

- fillers,
- non-verbal elements (e.g., laughter, vocalisations),
- repeated segments within a child's own utterance (e.g., *she walks like… like a duck* → *like…* = 0),
- unintelligible speech,
- passages deemed by the researcher to be rote-learned, and
- partial repetitions of the adult's immediately preceding utterance (e.g., Adult: *he climbed up the tree*; Child: *he climbed up there and fell down* → repeated segment *he climbed up* = 0).

B. Sample MLUm scoring system

The scoring framework below was adapted from the Arabic Morphemes Per Utterance (MPU) system (Tallas-Mahajna & Dromi, 2023) described in Section 3.6.3, with minor modifications to reflect the morphological characteristics of the present corpus. Each morpheme type was assigned a numeric value as shown.

Nouns

Category/Form	Score	Example
Uninflected (no gender variation)	1	*siġġu* 'chair', *papà* 'dad'
Gender-variable (M/F)	2	*tifel – tifla* 'boy – girl'
Plural (sound or broken)	2	*fjuri* 'flowers' (SP), *blalen* 'balls' (BP)
Dual	2	*widnejn* 'two ears'
Collective	2	*siġar* 'trees'

Adjectives and Participles (Active, Past)

Category/Form	Score	Example
Masculine	1	kbir 'big'
Feminine	2	kbira 'big'
Plural	2	kbar 'big'
Comparative	2	ikbar 'bigger'
Superlative (including *l*-)	3	l-ikbar 'biggest

Verbs

Category/Form	Score	Example
Perfective, masculine, singular ('unmarked')	1	*ħolom* 'he dreamt'
Perfective/Imperfective (all other gender/person inflections)	2	*ħlomt* 'I dreamt', joħolmu 'they dream'
Imperative singular	1	*oħlom* 'dream'
Imperative plural	2	*oħolmu* 'dream'
Future marker	1	***se*** *noħlom* 'I'm going to dream
Continuous marker	1	**qed** noħlom 'I'm going to dream'

Negation

Element	Score	Example
mhux	1	***mhux*** *sabiħ* 'not nice'
ma	1	***ma*** *mar ħadd* 'nobody went'
-x	1	*ma żlaqt**x*** 'I didn't slip'

Objects

Type	Score	Example
Direct	1	sieq**i** 'my foot'
Indirect	1	ktibt**li** 'you wrote to me'

Pronouns, Demonstratives, Determiners, Uninflected Prepositions, Time Adverbials, Interrogatives, Quantifiers

Each assigned a score of *1* (e.g., *jien* 'I' = 1 *ta'* 'of' = 1).

Other Categories

Category	Score	Examples
Onomatopoeic words	1	*woof*
Proper nouns	1	*Beauty and the Beast*
Collocations/ construct phrases acquired as bound lexical items	1	*ħobż biż-żejt* 'bread with oil – traditional Maltese bread filled with a variety of ingredients'

C. Computation

MLUm values were computed manually from the clean transcripts using the morpheme scores described above. The total morpheme score per utterance was divided by the number of utterances included to yield a mean MLUm value per participant.

Appendix O: Sample Transcription Excerpt Scored Using the Present MLUm System and Portelli (2005)

This appendix provides a short excerpt from the corpus, illustrating the practical application of the MLUm scoring system described in Appendix N. For comparative purposes, the same excerpt has also been scored using Portelli's (2005) MLUm system, the only other established framework applied to Maltese child language data.

The example serves two purposes:

- to demonstrate how the present scoring conventions (described in Appendix N above) were operationalised in practice, and
- to highlight key differences in morpheme assignment and resultant MLUm values between the two systems.

The table below presents a short excerpt from a child language transcription, with all adult utterances excluded. The excerpt represents the post-cleaning dataset; that is, any utterances removed prior to analysis (see Appendix M) are not included. The data are taken from the second recording session with Duminku at age 3;00.09. Each of the 20 child utterances is shown with the original transcription conventions, followed by segmentation and morpheme scoring under both the present MLUm system and Portelli's (2005) framework. Scores marked with an asterisk (*) indicate cases not explicitly addressed in Portelli's system (English lexical items, false starts or cut-offs, and conjunctions) and are therefore scored in accordance with the present study's conventions to avoid confounding effects.

Table 0.1 *Sample Child Language Excerpt Scored Using the Present MLUm System and Portelli (2005)*

	CHI Utterance	This study's MLUm scoring system	Score	Portelli's (2025) scoring system	Score
1	i'id nilgħab [n]logħba	irrid(2) nilgħab(2) logħba(1)	5	irrid(2) nilgħab(2) logħba(1)	5
2	trid nerġgħu nibdew minn il-bidu?	trid(2) nerġgħu(2) nibdew(2) minn(1) il-(1) bidu(1)	9	trid(2) nerġgħu(3) nibdew(3) minn(1) il-(1) bidu(1)	11
3	dil-puzzle tal-logħba [[bħal di']]	dil-(2) puzzle(1) tal-(2) logħba(1) bħal(1) din(1)	8	dil-(2) puzzle(1) tal-(2) logħba(1) bħal(1) din(2)	9
4	ara dik	ara(1) dik(1)	2	ara(2) dik(2)	4
5	di... di' bħal /game/ ok?	di...(0) din(1) bħal(1) game(1) ok(1)	4	di...(0*) din(2) bħal(1) game(1*) ok(1)	5
6	imma... imma ma nafx nagħmilha	imma...(0) imma(1) ma(1) nafx(3) nagħmilha(3)	8	imma...(0*) imma(1*) ma(1) nafx(2) nagħmilha(3)	7
7	naħseb di'	naħseb(2) din(1)	3	naħseb(2) din(2)	4
8	tazza u p... platt	tazza(1) u(1) p...(0) platt(1)	3	tazza(1) u(1*) p...(0*) platt(1)	3
9	m... man-nanna	m...(0) man-(2) nanna(2)	4	m...(0*) man-(2) nanna(2)	4
10	erm (.) farfett wijj(0)	erm(0) farfett(1) wijj(0)	1	erm(0) farfett(1) wijj(0)	1
11	tal-/ball/	tal-(2) ball(1)	3	tal-(2) ball(1*)	3
12	/doggy/	doggy(1)	1	doggy(1*)	1
13	im... mhu[s] hekk di'	im...(0) mhux(1) hekk(1) din(1)	3	im...(0) mhux(1) hekk(1) din(2)	4
14	di' ħażina hekk	din(1) ħażina(2) hekk(1)	4	din(2) ħażina(2) hekk(1)	5
15	/fro[k]/	frog(1)	1	frog(1*)	1
16	[[u di']] minn haw' di'	u(1) din(1) minn(1) hawn(1) din(1)	5	u(1*) din(2) minn(1) hawn(1) din(2)	7
17	dik hemmhekk	dik(1) hemmhekk(1)	2	dik(2) hemmhekk(1)	3
18	fejn hemm tad-/doggy/?	fejn(1) hemm(1) tad-(2) doggy(1)	5	fejn(1) hemm(1) tad-(2) doggy(1*)	5
19	di' u di'	din(1) u(1) din(1)	3	din(2) u(1*) din(2)	5
20	mad-/doggy/	mad-(2) doggy(1)	3	mad-(2) doggy(1*)	3
		MLUm=3.85(77/20)	**77**	**MLUm=4.5 (90/20)**	**90**

As anticipated, the present system yielded slightly lower MLUm values. Across the 20 utterances, total morpheme counts differed by 13 (77 vs 90), resulting in MLUm scores of 3.85 and 4.50 respectively. The main divergence arises from Portelli's (2005) treatment of demonstrative pronouns, each assigned a score of 2, which inflates scores given their frequency in spontaneous speech. In a few limited cases, however, her scores are lower. For example, while both systems assign a value of 2 to verbs in the perfective first person singular, Portelli adds an additional morpheme for the negative markers *ma* and *-x* (total = 3), whereas in the present study these are each scored separately (1+1). This decision reflects both the fact that *ma* and *-x* do not always co-occur in Maltese and the developmental possibility that a child may acquire one component of the negation system before the other.

Appendix P: Sample POS-Tagged Transcription

This appendix presents a sample excerpt tagged with the five-tag POS scheme developed for this study (see Section 3.7, Table 3.8 for tag definitions). The example, included for illustrative transparency, demonstrates how tags were applied.

Foundational principles:

- Consistency over granularity: Apply the same five tags throughout; avoid one-off or ad hoc categories.
- Interpret intent: Prefer the child's likely communicative intent in context when forms are non-target-like.
- Context first: Disambiguate with immediate discourse context; avoid reanalyses beyond the utterance frame.

The same excerpt used in Appendix O above is employed here for POS tagging, that is, the data are taken from the second recording session with Duminku at age 3;00.09. The tag legend is provided below with token-level tagging inline as *token|TAG.*

Tag legend (five-tag scheme; see Section 3.7, Table 3.8)

- N = Noun
- V = Verb
- ADJ = Adjective/Participles functioning adjectivally
- OTH = Other word classes (e.g., determiners, pronouns, pseudoverbs)
- CLS = Child-language-specific items (e.g., onomatopoeia, neologisms)

Table P.1 *Sample Child Language Excerpt Tagged Using the Present Five-Tag POS Scheme*

	CHI Utterance	This study's POS tagging system
1	i'id nilgħab [n]ogħba	irrid|V nilgħab|V logħba|N
2	trid nerġgħu nibdew minn il-bidu?	trid|V nerġgħu|V nibdew|V minn|OTH il-|OTH bidu|N
3	dil-puzzle tal-logħba [[bħal di']]	dil-|OTH+OTH puzzle|N tal-|OTH+OTH logħba|N bħal|OTH din|OTH
4	ara dik	ara|V dik|OTH
5	di... di' bħal /game/ ok?	di... din|OTH bħal|OTH game|N ok|OTH
6	imma... imma ma nafx nagħmilha	imma... imma|OTH ma|OTH nafx|V nagħmilha|V
7	naħseb di'	naħseb|V din|OTH
8	tazza u p... platt	tazza|N u|OTH p... platt|N
9	m... man-nanna	m... man-|OTH+OTH nanna|N
10	erm (.) farfett wijj	erm farfett|N wijj|CLS
11	tal-/ball/	tal-|OTH+OTH ball|N
12	/doggy/	doggy|N
13	im... mhu[s] hekk di'	im... mhux|OTH hekk|OTH din|OTH
14	di' ħażina hekk	din|OTH ħażina|ADJ hekk|OTH
15	/fro[k]/	frog|N
16	[[u di']] minn haw' di'	u|OTH din|OTH minn|OTH hawn|OTH din|OTH
17	dik hemmhekk	dik|OTH hemmhekk|OTH
18	fejn hemm tad-/doggy/?	fejn|OTH hemm|OTH tad-|OTH+OTH doggy|N
19	di' u di'	din|OTH u|OTH din|OTH
20	mad-/doggy/	mad-|OTH+OTH doggy|N
		Tag totals: OTH=41, N=14, V=9, ADJ=1, CLS=1

This excerpt demonstrates the consistent application of the five-tag POS scheme to Maltese child language data. Tag assignment adheres to the principles outlined above, balancing interpretive accuracy with methodological simplicity.

Appendix Q: Proportions of Maltese and English Tokens per POS Category

This appendix presents the proportions of Maltese and English tokens observed for each POS category across the 8 analysed transcriptions. Values are expressed as percentages of total tokens per transcription. See Sections 4.2 and 4.3 for a description and interpretation of these distributions.

Child	Age	MLUm	Nouns %		Verbs %		Adjectives %		Others %		Child-lang. specific %	
Anna	18m	1	**64.4**	65.5	**0**	0	**0**	0	**11.1**	40	**24.4**	En.
				34.5		0		0		60		Mt.
Anna	36m	3.54	**13.4**	41	**20**	5.2	**2.4**	71.4	**62.4**	2.2	**1.7**	En.
				59		94.8		28.6		97.8		Mt.
Stiefnu	18m	1.59	**25.5**	2.9	**11.7**	0	**2.2**	0	**29.9**	0	**30.7**	En.
				97.1		100		100		100		Mt.
Stiefnu	36m	2.45	**16.1**	21.2	**16.6**	8.8	**4.4**	0	**58**	25.2	**4.9**	En.
				78.8		91.2		100		74.8		Mt.
Kelina	36m	2.89	**16.6**	47.4	**24**	0	**9.2**	19	**47.2**	4.6	**3.1**	En.
				52.6		100		81		95.4		Mt.
Kelina	60m	4.18	**11.7**	17.6	**24.7**	6.9	**0.3**	0	**61.2**	6.7	**2.1**	En.
				82.4		93.1		100		93.3		Mt.
Reno	36m	2.68	**26.3**	57.7	**10.1**	10	**6.6**	53.8	**54**	7.5	**3**	En.
				42.3		90		46.2		92.5		Mt.
Reno	60m	5.18		6.2		0		20		0.4		En.
			17.2	93.8	**25.3**	100	**1.2**	80	**54.8**	99.6	**1.7**	Mt.

Appendix R: Sample Transcription Excerpts

This appendix presents selected transcription excerpts drawn from the study corpus. The complete transcriptions are not reproduced here, as they comprise a substantial body of material (each approximately sixty pages in length). Instead, concise representative snippets are included to illustrate both longitudinal development and cross-participant variation.

For one focal participant, *Lippu*, an excerpt from each of the five recording sessions is provided to demonstrate developmental progression over time. In this case, the excerpts are between three and four pages in length. For the remaining seven participants, excerpts from their first and final recording sessions are included to capture individual variation across the data collection period. Each of these excerpts spans one to two pages.

While full metadata accompany the complete transcriptions (see Appendix M), the metadata presented here have been substantially modified and all identifying details have been omitted or abbreviated to protect participant anonymity. The transcription data themselves remain unaltered. The digitisation and open availability of the full corpus are planned as part of the postdoctoral phase of this research, constituting a next step toward broader accessibility and potential reuse.

All recordings included in this appendix were conducted in each child's living room area designated for play. A parent was present during the session. Each description below provides a brief contextual overview of participant characteristics, session dynamics, and the focus of the excerpt.

#Lippu 27 months

#Child’s age on the day of the recording 2;03.21

#The mother was present and participated occasionally. The child, an only child who does not attend childcare, is highly Maltese-dominant with clear and intelligible speech. He was observant, assertive, and engaged. Background sounds include the family dog. In this excerpt, the child shifts from looking at a book to playing with a Noah’s Ark set, inventing playful “punishments” for the animals.

#CHI Child #INV Investigator

CHI	ilbist ta’ Daisy int	CHI spots Daisy Duck on the INV’s t-shirt
INV	x’inhi?	
CHI	ta’ Daisy	
INV	ta’ Daisy?	
CHI	dik hi Daisy (.) dik t’hemm	
INV	ee di’ fuq il-flokk tiegħi (.) eżatt (.) bravu	
INV	din ta’ Donald Duck hu’? Daisy Duck	
CHI	eħe	
INV	ee [[tħobbhom]]?	
CHI	[[Daisy]] Duck dik	
INV	tħobbha?	
CHI	inħobbha	
INV	u haw’hekk haw’ kelb	CHI and INV skimming through a book
INV	u x’għandu l-kelb? <noise>	
INV	taf x’inhi din? <noise>	
CHI	<vocalisations>	
INV	x’għandu? x’għandu haw’hekk?	
CHI	noħorġu di’ u dik (.) mhux il-/books/	CHI wants to play with toys rather than look through the book
INV	ee t-/toys/ hu’?	
CHI	[[mhm]]	
INV	[[ejja]] ħa nurik mela /toy/ ieħor li għandi	
CHI	/new toy ./	
INV	di’ vera sabiħa ara	

INV	ejja haw'hekk ħa nurihielek <noise>	INV setting up Noah's Ark
INV	mela din [[ara]]	
CHI	[[waqajt]]	
INV	waqajt?	
INV	eqq (.) bumm	
INV	poġġi haw'hekk ħalli ma tweġġax (.) fuq [[it-tapit]]	
CHI	[[bumm]] bumm (.) waqajt .	
INV	ara kemm haw' annimali haw' ġew	
INV	mela daw' ħa noħorġuhom kollha (.) tajjeb? <noise>	
INV	u ħa nagħmluhom haw'hekk	
CHI	fejn hemm annimali?	
INV	ħafna haw' annimali	
CHI	fejn?	
INV	haw'hekk (.) haw'hekk	
INV	arahom fejnhom l-annimali	
INV	issa ara ħa nagħmlu (.) l-iljuni	
INV	ħa nagħmluhom kollha ħdejn xulxin haw'hekk	
INV	x'qed jagħmel dak l-iljun? <noise>	
INV	daħal ġewwa? <noise>	
CHI	mhm	
INV	eqq <noise>	
CHI	nagħmillu /punishment/	CHI closes lion in the lower deck of the boat
INV	ħa tagħmillu /punishment/ miskin	
INV	għalfejn? x'għamel?	
CHI	għamiltlu /punishment/	
INV	jaħasra [[imma dak...]]	
CHI	[[għalaqtu...]] għalaqtlu d-/door/	
INV	għalaqtlu d-/door/?	
CHI	eħe	
INV	imma għalfejn tajtu /punishment/?	
INV	x'għamel?	
CHI	xi ħaġa tal-pastażi	
INV	ee għamel xi ħaġa tal-pastażi	

INV	’ma dak ħa jkollu joqgħod hemm ġew?	*imma*
INV	ma jistax jiġi ħdej’ sħabu?	
CHI	le	
INV	ee miskin (.) orrajt	
CHI	ma jistax jiġi ħdej’ sħabu	
INV	fhimt	
CHI	ma jistax ħa ji... li... jista’ jo’ ħdej’ (.) [j]il-mamà (.) ħdej’ il-mamà tiegħu	*joqgħod* reduced to *jo’*
INV	ee ma jistax immur ħdej’ il-mamà tiegħu	
CHI	a’ fejnha l-mamà tiegħu?	*allura* reduced to *a’*
CHI	liem hi?	
INV	liem hi [[l-mamà tiegħu]]? <noise>	MOT in the background
CHI	[[jaqaw]] di’?	
INV	bravu (.) illa’ kemm int bravu	
CHI	dik hi [[l-mamà tiegħu]]	
INV	[[għaraftha]] (.) dik hi .	
INV	hi tatu /punishment/?	
CHI	li... jien tajtu /punishment/	
INV	inti tajtu /punishment/?	
CHI	jien tajtu /punishment/	
INV	u issa ħa nħalluh joħroġ jew irid joqgħod hemm?	
CHI	i[n]id joqgħod hemm	*irid*
INV	ee orrajt .	
INV	issa ara (.) ħa nara kemm int b...	
CHI	qed fil-/boat/	*qed* – dialectal
INV	qiegħed fil-/boat/ hu’? [[bħala /punishment/]]	
CHI	[[qiegħed fil-/boat/]] .	
INV	il-mamà tiegħu taf li qiegħed hemm ġew jew qed tfittixu?	
CHI	taf fej… qiegħda... q... qed tfittxu	
INV	ee qed tfittxu?	
CHI	[[hemm ġew?]]	
INV	[[ara xi...]] (.) hemm ġew	
INV	hemm ġew qiegħed?	
CHI	le (.) hemm ġe... hemm [ġ]arra	*barra*

INV	hemm t... hemm barra (.) hu’?	
INV	hemm taħt qiegħed (.) <noise>	
INV	iva (.) <noise>	
CHI	daħlet	
INV	daħlet?	
CHI	daħlet ġo rasha	
INV	[[ħabtet]] rasha?	
CHI	[[daħ...]]	
CHI	le mhux ħabtet rasha (.) daħlet	
INV	ee daħlet rasha hemm ġew (.) fhimt	
CHI	daħlet rasha	
INV	u din iż-żebra ara (.) di’ fej’ qiegħda?	
CHI	nagħm[il]lha /punishment/ ukoll	
INV	le x’tagħm[il]lha /punishment/ ukoll	
INV	x’għamlet di’?	
CHI	ħar... ħaġ... xi ħaġa tal-pa... tal-pastażi	
INV	ukoll?	
CHI	eħe	
INV	illami	
CHI	fe’n qiegħda l-mamà tiegħu?	CHI looking for a big zebra figurine
INV	m’għandix mamà żebra .	

#Lippu 33 months

#Child's age on the day of the recording 2;09.05

#The child was curious, talkative, and confident, frequently asking questions and retelling familiar stories. Speech was fully intelligible, even when briefly using a pacifier. He remains Maltese-dominant and demonstrates strong narrative skills and memory. This snippet captures his lively recounting of *Beauty and the Beast* and *Aladdin*.

#CHI Child #INV Investigator

INV	naqra (.) mqareb hu'?	CHI and INV talking about Gaston
CHI	naqra mqareb (u jiżaħħu) daqqa 'l Beast	
INV	x... jagħti daqqa lil Beast?	
CHI	mhm	
INV	ħażin hu' jagħmel hekk?	
CHI	mhm	
INV	għalfejn itih daqqa?	
INV	tiftakar?	
CHI	a... għax hu ħa... (.) għax ħa jiġi gi... għax ħa jiġi /prince/	
INV	ee għax ikun ħa jsir /prince/?	
CHI	eħe	
INV	[u]mbagħad fl-aħħar isir /prince/ hux?	
CHI	mm	
INV	nibdew nerfgħuhom issa?	
CHI	xħin... xħin tin... [t]ittih /kiss/ i'-Belle ji... jiġi dak il-glitter	*il-*
INV	jiġi dak il-_	
CHI	dak jiġi dak il-glitter u jiġi /prince/	
INV	eżatt (.) mela Belle ittih /kiss/ [u]mbagħad jiġi l-glitter u jsir /prince/ (.) eżatt	
CHI	dak l-ewwel	
INV	eh?	
CHI	Gaston tah daqqa ta' sikkina	
INV	tah daqqa ta' sikkina?	
CHI	mhm	

INV	eh?	
CHI	[u]mbagħad waqa' mal-art	
INV	eħe	
CHI	[u]mbagħad Belle ta[ħ]tu /kiss/	*tatu*
CHI	u ġie dak il-glitter u ġie /prince/	
INV	eżatt	
INV	u taf għalfejn kien /beast/ fil-bidu?	
CHI	għax ma kienx jixxerja [[fil-bidu]]	
INV	[[ee għax]] (.) fil-bidu ma kienx jixxerja	
INV	u min għamlu?	
INV	tiftakar min daw... min għamlu /beast/?	
CHI	dik il-/witch/	
INV	dik il-/witch/ hux vera?	
INV	u kellu dik il-fjura (.) hux?	
CHI	għax l-ewwel mill-bidu beda	
INV	eh	
CHI	u kien hemm /story/ biex ma jixxerjax u ġie /beast/	
INV	eħe	
CHI	[u]mbagħad (.) għamlu Be Our Guest u Tale As Old As Time	CHI singing the words 'tale as old as time' as if he's about to sing the song
INV	u Tale As Old As Time (.) eħe	
CHI	[u]mbagħad tefgħuh ġol-gaġġa	
INV	eħe?	
INV	min kien ġol-gaġġa?	
CHI	il-w... il-papà u l-Belle	
INV	il-papà u l-Belle kienu fil-ħabs (.) eħe bravu .	
CHI	dik tal-/cartoons/ qed ngħidlek	
INV	eżatt (.) [[eħe]]	
CHI	[u]mbagħad [[dejjem]] jgħidu Tale As Old As Time	CHI singing the words 'tale as old as time' as if he's about to sing the song
INV	eħe (.) sabiħa dik	
CHI	[u]mbagħad bd... wara kienu qed jgħidu Tale As Old As Time	
INV	eħe	

CHI	[u]mbagħad ż... ġie Gaston	
INV	eħe	
CHI	u tah daqqa ta' [kissina]	*sikkina*
INV	mhm	
CHI	[u]mbagħad ta[ħ]tu /kiss/ Belle	*tatu*
INV	tatu /kiss/ Belle (.) iva	
INV	[[lil Beast hux vera?]]	
CHI	[[[u]mbagħad ġie l-gli...]]	
CHI	xħin tatu /kiss/	
INV	eħe	
CHI	ġie dak il-glitter	
INV	eħe	
CHI	[u]mbagħad ġie bea... ġie qi...	
INV	ġie /prince/	
CHI	ġie /prince/	
INV	neħħiha l-gażaża għax ma [[nifhmekx biha]]	
CHI	[[hekk]] l-istorja tagħhom	
INV	hekk l-istorja tagħhom?	
CHI	mhm	
INV	[u]mbagħad f'Aladin /two/ x'jiġri?	
INV	għax jiena l-/one/ biss rajt	
INV	x'hemm f'Aladin /two/?	
INV	hemm Yas... hemm Jasmine?	
CHI	hemm	
INV	u x'tagħmel fih?	
CHI	in-/number two/?	
INV	eh	
CHI	u hemm /number three/	
INV	eh x'jiġri fihom dawk?	
CHI	dawk tal-/istory/ tal-Gafa... Jafar[m]	*Jafar*
INV	l-istorja tal-_	
CHI	Jafar[m]	
INV	eh veru?	
CHI	'bagħad xħin mill-bidu ġie Jaf... ġie Jas... Ala... dawn... dawk il-/horses/	

INV	eħe?	
CHI	imbagħad ġie għal dak (.) ġie Jafar	
INV	eħe?	
CHI	u ġie Aladin	
CHI	l-ewwel ġew... marru dawk il-/horses	
CHI	[u]mbagħad dak (.) Abu	
INV	Abu?	
CHI	[u]mbagħad ġie Aladin fuq rashom	
INV	eħe	
CHI	u qabeż ħdej' Abu	
CHI	[u]mbagħad kien qed jiġ'i u qabeż fuq il-fo... /horse/	*jiġri*
INV	eħe	
CHI	[u]mbagħad kien qed jieħu dawk il-/treasures/	
INV	ee kien qed jieħu t-/treasures/ (.) eh	
CHI	[u]mbagħad dak (.) dak il-/birdie/ kien ġo gaġġa	
INV	mm	
CHI	[u]mbagħad kien qed jiftaħ Aladin u ġiet Jasmine	
INV	mhm	
CHI	u [dħalu] hemm ġew	*daħlu*
INV	u x'għamlet?	
CHI	u... u idħol hemm ġew (.) ġol-/lamp/	
CHI	u l-ġ... il-/genie/ il-... il-... il-Jafar għamlu ġol-/lamp/	
INV	ee għamlu ġol-/lamp/	
INV	u Abu min kien?	
INV	ix-xadina?	
CHI	mm	
INV	Abu x'kien?	
INV	xadina jew jew [[għasfur]]?	
CHI	[[ejja]] naraw in-/number two/	
INV	naraw in-/number two/?	
INV	jien ma nistax noqgħod haw' narah miegħek <chuckle>	

CHI	le naraw waħ... it-/two/	
INV	eh?	
CHI	[u]mbagħad noq'għdu nilagħbu	*noq'għdu* – dialectal
MOT	<chuckle>	
INV	imma naħseb twil eqq it-/two/ jiena	
CHI	mhux twil	
INV	eh	
CHI	in-/number one/ twil	
INV	u n-/number two/ kemm hu twil?	
CHI	mh[i]x twil	*mhux*
INV	issa narah id-dar ħalli darb'oħra li niġi ngħidlek għoġobnix	
INV	tajjeb?	
CHI	ħa nurihulek	
CHI	agħmilli l-film	CHI addressing his MOT

#Lippu 38 months

#Child's age on the day of the recording 3;02.07

#The mother was nearby but minimally involved. The child was enthusiastic, confident, and fully intelligible. Some dialectal traces appear in his Maltese, likely influenced by his grandmother. He excitedly welcomes the investigator and shares his toys, celebrating his recent third birthday milestone.

#CHI Child #INV Investigator #MOT Mother

CHI	ġibt dawn [[ukoll biex nilgħab]]	
INV	[[ġibt dawk]] ukoll? <noise>	
INV	illami kemm hi sabiħa dik	
INV	dik ta' xiex inhi?	
CHI	ta' Cindirella	
INV	ta' Cindirella	
CHI	u tar... ta' Cindirella di'	
CHI	ara x'tagħmel di' <noise>	
INV	x'tagħmel?	
INV	urini	
CHI	tqila	CHI carrying it and placing it on the floor next to the INV
CHI	ara kemm hi sabiħa	
MOT	<chuckle>	
INV	ħa nara	
CHI	minn dawn il-/buttons/	
INV	eħe?	
CHI	tixgħel	
INV	<gasp> wow <noise>	
INV	dik min tahielek?	
CHI	u għandi xi ħaġ'oħra (.) [[x'nurik]]	
INV	għandek [[xi ħaġ'oħra]]?	
MOT	mm	
INV	li ħa turini?	
INV	[[urini]]	
MOT	[[x'ġej?]]	

CHI	fil-kamra tiegħi (.) <noise>	CHI in the distance, walking to his room
CHI	fejnha Ariel li ddoqq?	CHI in the distance, cannot find the toy he wants
MOT	mhux haw' taħt Lippu	
CHI	dawk qegħdin f... taħt it-televixin	
MOT	ee għax kont poġġejtha darba fil-kamra	
MOT	imma naħseb haw' qiegħda ta	
MOT	tini ċans <noise>	MOT looking for the toy the CHI wants
CHI	ħa nurik xi ħaġa vera differenti [[issa]]	
INV	[[xi]] ħaġa vera [[differenti]]?	
MOT	[[haw' ara (.) <laughter>]]	
INV	[[ħa nara <laughter>]]	
CHI	[[ħa nurik]]	
MOT	<laughter>	
CHI	din [[għandha l-ilma]] u tkanta	
INV	[[wow (.) dik għandha l-ilma ġo fiha	
CHI	smajtha kif...	
CHI	u jien għandi waħd' oħra	
INV	għandek oħra?	
CHI	għandi [[din]]	
INV	[[illami]] kemm għandek	
CHI	għandi din	
INV	eħe?	
CHI	din	
INV	kemm huma sbieħ	
INV	liem waħda tħobb inti l-iktar minnhom?	
CHI	Moana	
INV	Moana?	
INV	anke jien Moana jogħġobni ħafna	INV referring to the movie
INV	imma di' vera sabiħa bl-ilma	
INV	u għandha daw' ukoll bil-glitter hu'? <noise>	
CHI	u tkanta <noise>	
INV	u tkanta wkoll <noise>	
INV	daw' ikantaw?	

CHI	jien irrid n... nilgħab magħha issa <noise>	CHI addressing his MOT
MOT	[[mel' ilgħab magħha]]	
INV	[[trid tilgħab miegħi?]]	
CHI	[[mm]]	
INV	[[mela mela]]	
INV	trid nerfgħuhom lura haw' ġew daw'?	
CHI	um... [u]mbagħad nilgħab wara l... b'tiegħek daw'	
MOT	[[chuckle]]	
INV	[[eżatt]] (.) wara tiegħi nilagħbu b'daw'	
INV	tajjeb?	
INV	ejja nerfgħuhom [[għalissa]]	
CHI	[[tajjeb?]]	
INV	tajjeb ħafna .	
INV	[[mela...]]	
CHI	[[u]] bid-/doll/ biex nagħmlu dik?	
INV	bil-xiex?	
CHI	biex nagħmlu di'	CHI pointing at his Cinderella carriage toy that lights up
INV	dik tad-dawl tiegħek?	
CHI	biex nagħmlu dik tad-dawl	
INV	mela mela mela	
CHI	ara x'ħa nurik iktar [[fuqha din]]	
MOT	[[<chuckle>]]	
CHI	araha issa	
INV	urini	
CHI	ara issa x'tagħmel	
MOT	ara l-laptop	
MOT	bil-mod	
INV	x'tagħmel? x'tagħmel?	
CHI	ara x'tagħmel <noise>	
CHI	<struggling sound> di' nista' ntellagħha?	
INV	ħa nara .	CHI pressing a button for the toy carriage to start move
INV	ii	
CHI	din tagħmel iktar	

INV	tagħmel iktar?	
INV	kemm hi f... sabiħa di' eh	
CHI	di' (.) vera?	
CHI	jekk irridu (.) hekk tixgħel u hekk inwaqqfu	
INV	ee twaqqfu	
INV	tagħfas il-buttuna (.) tixgħel u twaqqfu hux vera?	
CHI	ijwa	
INV	um... id-dawl?	
INV	minn fejn id-dawl?	
CHI	id-dawl (.) minn hemmhekk	
INV	minn hemmhekk?	
CHI	nistgħu nilagħbu bit-/toys/ tagħha?	
INV	trid tilgħab bit-/toys/ tiegħi?	
INV	mela mela mela	
CHI	x'ġibt illum differenti? <noise>	
INV	x'ġibt differenti?	
INV	ħa nurik kollox li għandi orrajt jiena <noise>	
INV	[[ara...]]	
CHI	[[u]] jien għandi dawn differenti	
CHI	ta[j]hom... ta[j]homli għall-/birthday three/	
CHI	jien għall-/birthday three/ armejt id-/dummies/ kollha	
INV	armejt id-/dummies/ kollha?	
INV	veru?	
INV	illa' kemm int bravu	
INV	vera jew?	
CHI	vera	
INV	u int meta [[għalaqt /three/]]?	
CHI	[[ħa nurik]] il-kaxxa tal-... taf... ta' fuq is-sodda	
INV	x'ħa turini?	
CHI	il-kaxxa ta' [[fu...]]	
CHI	[[il-ko...]] taf... taf... taħt is-sodda nett	
INV	eħe?	

MOT	[[x'għandu?]]	MOT in the distance
CHI	[[eħe biex]] nurik	
INV	qed jurini li rema [[d-/dummies/ kollha]]	
MOT	mela	
INV	[[għandu xi kaxxa /dummies/?]]	
MOT	[[x'ħa turiha imma?]]	
MOT	kaxxa?	
CHI	ħa nuriha dik	
MOT	ee jaħasra l-kexxun fej' kellu d-/dummies/ u issa vojt [[<chuckle>]]	
INV	[[ee żvujtajtu?]]	
CHI	ħa nara kif ħa nurik biex veru	
INV	ħa turini biex nara [[li veru]]?	
MOT	uriha (.) uriha	
INV	<chuckle>	
INV	ħa[n] nara	
INV	<gasp> bla /dummies/	
CHI	bla /dummies/	
CHI	[[daqshekk]]	
INV	[[wow]]	
INV	daqshekk hu' għax issa inti qed tikber hu'?	
CHI	u bqajt nieħu l-biċċa biss	
MOT	[[<laughter>]]	
INV	[[il-biċċa biss]]	
INV	wow (.) kemm int bravu .	
INV	mela (.) issa għandek /three/ .	
CHI	issa rrid nilgħab bit-/toys/ tagħha	

#Lippu 44 months

#Child's age on the day of the recording 3;08.25

#The mother participated briefly at the beginning and end. The child, now attending kindergarten, was confident, cooperative, and imaginative, enjoying the session thoroughly. The family dog can be heard occasionally. This excerpt captures play with the Playmobil playground set as the session draws to a close.

#CHI Child #INV Investigator #MOT Mother

CHI	A[j]iel (.) fej' ħa toqgħod?	*Ariel*
INV	Ariel? tista' tinżel [[miż-żurżieqa wkoll]]	
CHI	[[miq... ma' di']] <noise>	
INV	ee għax dik naqra kbira hu'?	
INV	tgħid toqgħod fuq il-bandla?	
CHI	toqgħod imma ma ridtx	
INV	ee ma tridx tmur fuq il-bandla (.) orrajt (.) <noise>	
INV	eqq	
CHI	qabżet <chuckle>	
CHI	din ħa tkun... din ħa tkun il-b... din ħa tkun il-/baby/ ta' dik	
INV	ee orrajt .	
CHI	għax... għax l... din... [i]din ħa tmur ġol-banju	
INV	eh? <noise>	
CHI	għax dik meta tmur ġol-banju 'tiġi /mermaid/	'*tiġi* – dialectal
INV	ee	
CHI	u issa	
INV	eh?	
CHI	xħin marret ġol-banju (.) ġiet (.) hekk .	
CHI	ġiet hekk	
INV	ee marret ġol-banju u saret żgħira qisha?	
CHI	u saret /mermaid/	
INV	u saret /mermaid/	

INV	fhimt (.) ok .	
INV	tajjeb (.) trid nibdew nerfgħuhom issa?	
CHI	x'nerfgħu?	
INV	kollox	
CHI	biex?	
INV	biex imbagħad immur jiena	
INV	imbagħad x'ħa toqgħod tagħmel inti meta mmur jiena? <noise>	
CHI	nilgħab bit-/toys/ tiegħi	
INV	tilgħab bit-/toys/ tiegħek?	
CHI	ejja nilagħbu bit-/toys/ tiegħi ftit	
INV	ftit? orrajt	
INV	mela jiena ħa nerfa' daw' (.) tajjeb?	
CHI	imma ħalli l-bandli	
INV	il-bandli?	
INV	imma kollox ħa nerfa' l-bqija (.) orrajt?	
CHI	kollox nerfgħu imma barra l-bandli <noise>	INV putting toys away
INV	tajjeb hekk <noise> .	
CHI	ħa nagħm'l iż-żurżieqa <noise>	*nagħmel*
INV	agħmilha (.) agħmilha (.) [[waħħalha]]	
CHI	[[għax...]] għax qabżet di' Ariel <noise>	
INV	eh x'għamlitilna <noise>	
CHI	x'għamlitilna <noise> .	
CHI	int Michela jew M... Mixela?	
INV	Michela (.) Michela jiena <noise> .	
CHI	ara l-bandla fej' weħħiltha	
INV	ee għamiltha iktar 'il fuq	
INV	sabiħ ta' hemm fuq (.) togħġobni jiena	
CHI	jew haw' isfel	
INV	ukoll jekk trid	
INV	ara waqagħlek dan ħa	
CHI	mi... x'waqagħli?	
INV	dak (.) ħa naħ... ħa .	
CHI	dik tiegħek	
INV	eża' .	*eżatt*

CHI	din kif tweħħilha?	
INV	ejja ħa nagħmilhielek jiena .	
CHI	din titneħħa? <noise>	
INV	titneħħa iva imma...	
CHI	kif?	
INV	ma nafx kif naqilgħuha <noise>	
CHI	din xorta titneħħa?	
CHI	din titneħħa? <noise>	
INV	dik iva (.) dik ġibha haw' ara <noise>	
INV	bravu <noise>	
CHI	imma jien ħa nibniha issa <noise>	
INV	ibniha (.) ibniha (.) f'idejk <noise> .	
CHI	din tinqala'	
INV	dik (.) ukoll	
INV	u din ar... inti haw' fuq weħħiltha?	
CHI	ee (.) eħe	
INV	haw' fuq tridha jew haw'hekk?	
CHI	ee emm fuq .	
INV	ħa (.) eħe?	
CHI	<noise> imma... imma din-naħa fej' jistgħu joq'du	*joq'du* – dialectal
CHI	fej' jistgħu j... imorru haw' <noise>	
INV	orrajt	
CHI	mamà w'ra mel... tista' tilgħab miegħi	*wara*
INV	<chuckle>	
MOT	x'ħa tagħmel ħi?	
INV	jiena ħa mmur dalwaqt	
MOT	mel' issa noqgħod nilgħab miegħek jien (.) tajjeb?	
CHI	eħe	
INV	tajjeb? <noise>	
CHI	emm mamà tista' ġġibli dak is-/soft toy/ ta' Stitch?	
MOT	/soft toy/ ta' Stitch?	
INV	[[/soft toy/ ta' Stitch?]]	
MOT	[[ħa turihulha lil]] Michela?	

CHI	Mixela jew Michela?	
INV	[[Michela (.) Michela]]	
MOT	[[le Michela]]	
MOT	ajma	
INV	mela	
CHI	emm it... na... reġgħet [i]tneħħietli di’	
INV	haw’ ħa <noise> .	
CHI	fejn toqgħod?	
INV	min?	
CHI	int	
INV	jiena?	
CHI	[[eh]]	
INV	[[Ħal]] Balzan noqgħod	
CHI	fejn?	
INV	fejn?	
INV	eqq ġo dar hu’?	
INV	ii [[kemm hu ħelu]]	MOT brining CHI’s soft toy
MOT	[[ħa Lippu]]	
CHI	dan tiegħi	
INV	dak min tahulek?	
CHI	mhux minn Disneyland	
INV	ee minn Disneyland	
CHI	ar’ int qed tara buttuna?	*ara*
INV	eħe?	
INV	qed jixgħel (.) qed jixgħel	
CHI	haw’hekk	
INV	kemm hu sabiħ	
MOT	mal-bandli ħa grazzja	

#Lippu 50 months

#Child’s age on the day of the recording 4;02.05

#The mother was present but not involved, occasionally heard in the background while cooking. The family dog was nearby. The child was lively, outspoken, and highly engaged, showing excellent command of Maltese. He displayed strong storytelling and pretend-play abilities. This excerpt begins with the *Frog Story* (Mayer, 1969) narrative task and transitions to play with a Playmobil ski lodge.

#CHI Child #INV Investigator

INV	mela din ħa tkun storja (.) dwar tifel li kellu żrinġ	INV introducing *Frog Story*
INV	imma daż-żrinġ intilef	
INV	tajjeb?	
INV	[[mela]]	
CHI	[[u]] kelb	
INV	u kelb ukoll	
INV	mela (.) bonġu (.) nies	
INV	illum għandna lil Lvcv magħna u ħa jgħidilna storja (.) dwar (.) żrinġ (.) mitluf	INV pretending to address TV viewers
INV	Lippu f’idejk <noise>	
CHI	darba waħda (.) kien hemm tifel b’żrinġ ġo bott	
CHI	u kien hemm kelb barra .	
CHI	[u]mbagħad (.) kien rieqed u k... il-... i[l]-żrinġ ħarab .	*iż-żrinġ*
CHI	[u]mbagħad (.) qam u ma sabx il-kelb .	
INV	il-kelb sabu (.) hux [[vera]]?	
CHI	[[mhm]]	
INV	il-kelb qiegħed ħdejh	
INV	lil min ma sabx?	
CHI	i[l]-żrinġ .	*iż-żrinġ*
CHI	’bagħad fittex kull m’hemm u ma sabux .	*imbagħad*
CHI	[u]mbagħad iċċekkja barra t-tieqa (.) u rah .	
CHI	’bagħad ħarab il-kelb .	
CHI	[u]mbagħad qabeż .	
INV	it-tifel qabeż mit-tieqa (.) [[hux]]?	

CHI	[[mm]]	
INV	u reġa’ sab il-kelb .	
CHI	mhm	
CHI	imbagħad qed jgħajjat għa[ll-i]żrinġ	*għaż-żrinġ*
INV	eżatt (.) qed ifittxu	
INV	u x’haw’ haw’hekk?	
INV	haw’ ħafna _	
CHI	siġar	
INV	iva (.) u qed ifittxu	
INV	qed jgħajjat	
CHI	’bagħad sab iż-żunżan .	
INV	iva	
CHI	[u]mbagħad (.) sab żrinġ imma ma kienx tiegħu	
INV	ee	
CHI	[u]mbagħad baqa’ [jfittx] .	*jfittex?*
CHI	[u]mbagħad (.) waqa’ għax kien hemm balla żnanż[e]n .	*żnaż[e]n* – dialectal
CHI	[u]mbagħad (.) sab /owl/ u beda jiġri wara’h .	*wara’h* – dialectal
CHI	[u]mbagħad (.) kompla jgħajjat .	
CHI	[u]mbagħad qabditu .	
INV	min qaba... qabdu hemm?	
CHI	ma n... x’inhi?	
INV	dik qisha ta’ Father Christmas naqra hu’?	
INV	qisha ċerva jew [[/reindeer/]]	
CHI	[[mhm]]	
INV	qabditu	
INV	ara fej’ t[a]la’ (.) fuq rasha	
INV	u issa?	
CHI	kien ħa jaqa’	
INV	kien ħa jaqa’	
CHI	’bagħad hu xor... waqa’ huwa u l-kelb .	
CHI	imbagħad waqa’	
INV	ġo xiex waqa’ taħseb?	
CHI	ġol-ilma	
INV	bravu	

CHI	imbagħad qam u sab...	
INV	bravu (.) x'sab?	
INV	ħa[n] naraw	
CHI	[s]okk	*zokk*
INV	zokk (.) iva	
INV	u x'qed jgħidlu 'l-kelb?	
CHI	x'qed jgħidlu?	
INV	qed jgħidlu joqgħod _	
INV	kwiet (.) hux veru?	
CHI	għal xiex?	
INV	ejja ħa[n] naraw	
INV	għax x'kien hemm wara z-zokk?	
CHI	iż-żrinġ	
INV	żrinġ	
INV	wieħed?	
CHI	tnejn	
INV	u issa?	
CHI	ġie ferħan	
INV	iva .	
CHI	[u]mbagħad kellu wieħed f'ide'h (.) 'bagħad kollha wara'h	*ide'h, wara'h* – dialectal
INV	u x'ta... fej' taħseb li mar [u]mbagħad?	
CHI	id-dar	
INV	mar id-dar	
INV	bravu	
INV	/high five/ (.) inti tajjeb għal fuq it-televixin eh	
INV	biex tgħid l-istejjer	
INV	kemm int tajjeb Lippu	
INV	mela ara ħa ngħidlek .	
INV	għandi hekk u għandi hekk	INV showing CHI stickers to choose two
INV	agħżel inti liem tixtieq	
CHI	mm	
INV	tista' tieħu waħda minn haw' u waħda minn haw'	
INV	jew tnejn minn haw'	

INV	agħżel tnejn .	
INV	ħu ħu (.) mela [[mela]]	
CHI	[[hemm]] il-/favourite colours/ qi'd wi.. w[a]ħed minn isfel u wieħed minn fuq	*qiegħed, wieħed*
INV	[[mela]]	
CHI	[[ma]]	CHI calling MOT to see the stickers and stick them on chart
INV	naħseb il-/mummy/ qiegħda fuq it-telefown bħalissa għax qed nismagħha	
CHI	[w]uq iċ-/chart/	*fuq*
INV	agħmilha fuq iċ-/chart/	
CHI	ħa nġibha miegħi	CHI in the distance bringing his sticker chart
INV	mela mela	
INV	orrajt	
INV	mela trid noħorġu t-/toys/ issa?	
CHI	mhm	
INV	għandi xi ħaġa vera [[ħa]]...	
CHI	[[ħa]] ngħolli naqra 'l Stitch	CHI propping up his Stitch soft toy
INV	għax miskin tah ħass ħażin	
INV	iv' eżatt (.) agħmlu hemm u di' agħmilha ħdejh	
CHI	fin-nofs	
INV	fin-nofs (.) l-għarusa tiegħu	
CHI	mhm	
INV	għandi di' naħseb ħa tkun togħġbok lilek	
CHI	għal xiex taħseb li togħġobni?	
INV	għax tal-Playmobil <noise>	INV bringing out Playmobil set
CHI	żgur li togħġobni <noise>	
INV	hux veru? <noise>	
CHI	ħa [nippijkja]	CHI peeping into bag
INV	mela mela mela <noise>	
CHI	dar .	
INV	din taf x'inhi?	
INV	inti qatt mort /skiing/?	
CHI	mort (.) xi darba <noise>	
INV	veru?	

CHI	eħe	
INV	aħjar minni għax jien qatt ma mort /skiing/	
INV	din dar li jkun hemm ħdejn l-/iskiing/ <noise> tajjeb?	
CHI	għaliex ma ġibtx il-bandli?	

#Anna 18 months

#Child’s age on the day of the recording 1;03.05

#Both parents were present and participated to ease the child’s shyness. She was at the one-word stage and parents produced mostly English nouns within Maltese utterances. Engagement was mainly through gestures and animal sounds while playing with toy animals and a playground set.

#CHI Child #INV Investigator #MOT Mother #FAT Father

INV	di’ x’inhi din?	INV showing CHI a toy cow
CHI	[[moo]]	
INV	[[x’iktar?]]	
INV	ee ħa jinżel ukoll?	
CHI	[[oqqo]]	
INV	[[ejja ħa nagħmilhie...]]	
INV	oqqo .	
CHI	<vocalisation>	
FAT	<gasp>	
MOT	eqq	
FAT	dak min hu Anna?	
INV	dak x’inhu dak?	
CHI	kħħ	FAT showing CHI a toy lion
FAT	[[mm kif jagħmel il-/lion/ hu’ dak?]]	
INV	[[mm (.) hekk jagħmel hu’?]]	
MOT	[[<chuckle>]]	
FAT	kif jagħmel Anna?	
CHI	/[ai]/	*slide*
MOT	ii ħa jitla’ fuq is-/slide/ il-/lion/?	
INV	ii	
FAT	Anna	
MOT	<chuckle>	
FAT	kif jagħmel il-/lion/ Acv?	
CHI	[[kħħ]]	
INV	[[kif jagħm...]]	
MOT	<chuckle>	

INV	hekk hu'?	
CHI	/[ai] ./	*slide*
INV	u dan?	
INV	da' x'inhu dan?	INV showing CHI a toy pig
CHI	<snorts>	
INV	[[iva hekk jagħmel (.) brava .]]	
MOT	[[<chuckle>]]	

#Anna 36 months

#Child’s age on the day of the recording 3;00.20

#The parents were present but non-participatory. The child was talkative, energetic, and code-switched frequently between Maltese and English. Background noise includes her baby brother. In this excerpt, she shows the investigator her toy kitchen and recounts a minor injury while baking.

#CHI Child #INV Investigator

CHI	jien għandi din ara <noise>	
INV	ee (.) dik x’inhi?	
INV	ee qisha kċina [[wkoll]]	
CHI	[[hekk]] bħal di’ u hekk	
CHI	biex twaħħalhom	
INV	biex issajjar fuqha hux?	
CHI	eħe	
INV	u x’issajjar fuqha dik inti? <noise>	
CHI	jew doughnut <noise>	
INV	eħe <noise>	
CHI	jew /fi[s]/ <noise>	*fish*
INV	jew?	
CHI	/fish/	
INV	ee /fish/ (.) orrajt <noise> .	
INV	eh issakkru l-vann (.) sabiħ	
CHI	hawnhekk fej’ tagħmel...	
CHI	mi’ haw’ tiftaħ	*minn*
INV	tiftaħha u [[tagħlaqha minn hemm]]	
CHI	[[u /close/ hekk]]	
INV	eżatt	
CHI	u jsuq haw’	
CHI	hawnhekk il-(?) <noise>	CHI’s brother can be heard in the background
CHI	hawnhekk bie[s] jar... nara	*biex*
INV	biex jara hu’? [[il-mirji]]	
CHI	[[eħe]] <noise> .	
INV	[u]mbagħad għandhom daw’ ara wkoll	

CHI	x'inhu dak?	
CHI	[sikkini]	*sikkini ~ skieken*
INV	eżatt	
CHI	għa[s] jien għandi [xikkini] wkoll (.) haw'	*xikkini ~ skieken*
CHI	jien hawnhekk weġġajt bil-/biscuits/ tal-/coconut ball/ hawnhekk	
INV	ee miskina <chuckle>	
CHI	[u]mbagħad (.) [[għamiltha]] stikk	

#Stiefnu 19 months

#Child’s age on the day of the recording 01;07.08

#Both parents were present, and the mother supported comprehension when needed. The child, at the two-word stage, used gestures and sounds communicatively and showed good attention. Maltese is dominant at home, though English prevails in childcare. The excerpt involves shared reading of an animal board book.

#CHI Child #INV Investigator #MOT Mother

CHI	a[n]a	*ara*
MOT	ara x’haw’	
MOT	di’ x’inhi?	MOT pointing at a fish in the picture book
CHI	il-kokka	
MOT	[[le]] mhux kokka	
FAT	[[le]]	
MOT	di’ x’inhi?	
CHI	ħuta	
MOT	[[ħuta]]	
INV	[[ħuta]]	
INV	illa’ kemm int bravu	
MOT	bravu	
MOT	eqq di’ x’inhi?	MOT pointing at a tiger in the picture book
CHI	qq hemm it-tig’a	*tigra*
MOT	tigra	
INV	tigra	
MOT	bravu .	
MOT	eqq (.) fejn hu l-farfett?	MOT pointing at a butterfly in the picture book
CHI	baħħ	
MOT	baħħ [[<laughter>]]	
INV	[[baħħ?]]	
MOT	sibu ħi (.) sibu sibu l-farfett	
MOT	fejn hu l-farfett?	
CHI	hemm	CHI pointing at the butterfly in the picture book
MOT	hemm (.) [[bravu]]	

#Stiefnu 36 months

#Child's age on the day of the recording 03;30.00

#The child's parents were both present; however, the mother mostly participated as the father was tending to the child's baby sister. The child's vocabulary and focus continued to develop, and Maltese remained the dominant language at home. Here, he prepares a "pizza" for his father using a felt food set.

#CHI Child #INV Investigator #MOT Mother

CHI	<vocalisations>	
CHI	'ħja' nagħmel zalza	*aħjar*
MOT	aħjar tagħmel hu'? <laughter>	
INV	<chuckle>	
INV	eħe?	
MOT	u x'se nagħmlulu 'ssa?	
CHI	tadama	
MOT	mhm?	
MOT	x'iżjed?	
CHI	u din	
MOT	dik basla	
MOT	x'iżjed?	
CHI	mm (.) ok	
INV	eħe?	
INV	x'iktar?	
INV	xi jħobb il-papà?	
CHI	mm	
INV	kollox iħobb?	
CHI	mm (.) /with this/	
CHI	/one two t'ee/	*three*
CHI	/one two three fou' f[o]ve six/	*five*
INV	<laughter>	
CHI	<vocalisations>	
INV	daqshekk?	
INV	lesti?	
CHI	mm (.) emm	

MOT	eqq (.) issa daw’ x’se nagħmlu [[bihom?]]	
CHI	[[għad]] fadalli (da')	
MOT	[[nista’]] ni[e]kolha?	*nikolha*
CHI	[[qq]]	
MOT	[[nista’]] nikolha di’?	CHI shaking his head
INV	le?	
CHI	ħa npoġġiha haw'	
MOT	fej’ ħa tpoġġ... hemm?	

#Karmena 24 months

#Child's age on the day of the recording 2;00.20

#The mother and older sister participated throughout due to the child's shyness. The child code-switched frequently as her preferred play language is English since this is the language spoken at childcare. In this excerpt, the child engaged in pretend cooking play with her mother using her own toys. The investigator simply observed as the child was still warming up to her.

#CHI Child #MOT Mother

CHI	ġibtha	
MOT	x'għamiltli haw'?	
CHI	din	
MOT	xi xtrajtli ħi?	
CHI	di'	
MOT	dik x'inhi?	
CHI	dik tat-/toast/	
MOT	agħmilhieli jekk jogħġbok .	
MOT	/thank you/ (.) hekk	
MOT	issa xi naqra ikel għandek?	
CHI	eħe għandi	
CHI	ħa mmur ċ... ċekja	
MOT	ħa tmur tiċċekja (.) ok .	
CHI	għandi <noise> .	
CHI	għandi banana (.) jien ma	
MOT	ok	
CHI	għandi banana	
MOT	għandek banana	
MOT	erm (.) xi ħaġa (.) sħuna għandek? [[tajba?]]	
CHI	[[eh]] ee għandi	
CHI	tihieli l-banana	
MOT	x'inhu?	
CHI	/please/	
CHI	tihieli	
MOT	ħa ntihielek imma stenna naqra	

MOT	għidli naqra x'għandek l-ewwel ħalli [[nagħżel]]	
CHI	[[għandi]] /chicken/	
MOT	iva	
CHI	/blueberries/	
MOT	x'iktar?	
CHI	m'għandix /blueberries/	

#Karmena 44 months

#Child's age on the day of the recording 3;08.15

#Family members again participated to help the child feel at ease. Initially reserved, she became increasingly confident and playful. Although Maltese is used at home and school, her preferred play language is English. This excerpt captures her tending to *ġulbiena* (vetches), grown in Malta to decorate cribs around Christmas time.

#CHI Child #INV Investigator #MOT Mother #SIB Sibling

CHI	ara /how much/ ma	
MOT	<laughter> kibret hux?	
CHI	kibr... kibret /this size/ <panting>	
INV	x'għandek hemmhekk inti?	
CHI	rummiena	CHI referring to *ġulbiena*
MOT	hux rummiena (.) rummiena	
CHI	<laughter>	
CHI	ma (.) kibret	
INV	mhux rummiena naħseb	
MOT	x'jisimha dik (.) x'jgħidulha?	
MOT	il-ġu_	
CHI	dik ġu[nb]liena	*ġulbiena*
MOT	ġu[nb]liena <laughter>	
CHI	<laughter>	
INV	/wow/	
MOT	issa dik mhux suppost turiha dawl	
CHI	ma /it's this size/	
MOT	kibret sew (.) issa nagħmluha quddiem il-Bambin ta	
INV	mela dik ma kinitx kbira?	
INV	kif kienet qabel?	
CHI	żgħira	
INV	żgħira?	
INV	[u]mbagħad x'għamiltu?	
CHI	l-ilma [u]mbagħad kibret ħafna	
INV	imma xi trid tagħmel?	
INV	trid tħalliha barra din jew le?	

CHI	le (.) fil-... fil-/dark/ (.) fid-dla... fid-dlam	
INV	ee (.) tħalliha fid-dlam?	
INV	u inti għamiltilha tajjar?	
CHI	ija	
MOT	inti għamilthom mal-papà	
MOT	x'għamiltu biex għamiltuha?	
INV	ħalli [[issa la mmur id-dar nagħmel]]	
CHI	[[l-ilma]]	
INV	l-ilma (.) eħe (.) x'iktar?	
CHI	u t-tajjar u... u /those circles which are black what are they called/?	CHI referring to the seeds
MOT	iż-ż_ iż-żer_	
CHI	iż-żerriegħa	
INV	ee (.) iż-żerriegħa?	
INV	u tfajtuhom fuq xulxin?	
INV	ġo xiex għamiltuhom?	
CHI	ma /can I get one/?	
MOT	imma x'in... għid lili qabel iġġibha	
MOT	fiex għamiltha?	
CHI	fil- ... /something/	
MOT	x'inhi imma?	
MOT	fi platt?	
MOT	fi platt għamiltha?	
CHI	<chuckle> le	
MOT	għidli fiex	
CHI	/can I get it?/	
CHI	ħa nġibhom	
SIB	fil-/containers/	

#Duminku 30 months

#Child's age on the day of the recording 2;06.12

#Conducted at the kitchen table, with both parents present but not involved. The child, who attends an English-dominant childcare centre, communicates well in Maltese and enjoys teasing the investigator. Speech is mostly intelligible. The excerpt captures an interaction with toy vehicles.

#CHI Child #INV Investigator

INV	ġibhom haw'hekk	
INV	kemm inti bravu (.) mhm	
CHI	u dan mhux[t] vapur	
CHI	da' mhux[t] vap'	*vapur*
INV	x'inhu mela da'?	
CHI	daw' mhux[t] vapuri	
CHI	da' tra'	*trakk*
INV	dak trakk (.) eżatt (.) bravu	
CHI	u da' tra[r]	*trakk*
CHI	tra[r] ukoll	*trakk*
INV	ukoll ukoll	
CHI	m'la daz... tra'	*mela, trakk*
CHI	da' mhux [p]apur	*vapur*
INV	x'inhu dak?	
CHI	da' qisu…	
CHI	da' qisu vann	
INV	qisu naqra vann	
INV	vera (.) għandek raġun	
CHI	u das-...	CHI shaking the toy boat because he thinks it's making a noise
INV	eħe jagħmel ħoss	
INV	qed tisimgħu?	
CHI	mel haw' haw' toqba żgħira	
CHI	<playing vocalisations>	CHI emulating the sound the toy boat is making
INV	hemm toqba żgħira (.) iva bravu <noise> .	

INV	iss’ ara l-vapuri ħa nerfgħuhom haw’ ġew	
CHI	da’ [t]sejjer < playing vocalisations>	CHI emulating the sound of a car engine being started
CHI	qabbad di’	

#Duminku 54 months

#Child’s age on the day of the recording 4;06.24

The father was nearby but did not participate. The child was talkative, imaginative, and articulate, using deliberate word play for humour. Speech was fully intelligible. In this snippet, he discusses his pet rabbit and school activities.

#CHI Child #INV Investigator

INV	x’inhi?	
CHI	meta toqgħod tkellimha	
CHI	toqgħod tagħmel	
INV	toqgħod tagħmel hekk (.) b’widnejha	
CHI	toqgħod iċċaqlaq widnejha	
INV	[[<chuckle>]]	
CHI	[[<chuckle>]]	
CHI	ara x’għandna	
INV	/wow/ (.) dik minn fej’ ġiet?	
CHI	tathieli t-/teacher/	
INV	tathielek it-/teacher/?	
INV	tħobb tmur l-iskola inti?	
INV	[[tieħu pjaċir?]]	
CHI	[[ija]]	
INV	veru?	
CHI	l-aħħar [[darba (.) ġiet]] il-mamà u kilna ġelat	
INV	[[mm]]	
CHI	u n-[nannas]	*nannas ~ nanniet*
INV	l-iskola?	
CHI	issa ħa mmorru l-/aquarium/	
CHI	/two days/ jew /one day/	
INV	x’ħa tmorru?	
CHI	l-/aquarium/	
INV	ee (.) l-/aquarium/	
INV	qatt mort inti (.) l-/aquarium/?	
CHI	ee (.) ħa... ħafna ta (.) drabi	
CHI	ma nafx kemm	

INV	ee (.) [[ġieli mort]]	
CHI	[[xi /two/]] (.) jew /three/	
INV	ma' min mort (.) qabel?	
CHI	ee (.) insejt	
INV	<chuckle> insejt?	
CHI	għax tant kien ilu	
INV	ee	
CHI	insejt	

#Toninu 30 months

#Child's age on the day of the recording 2;06.26

#The mother observed quietly throughout. The child was energetic, cooperative, and engaged, though still mastering several phonemes. The excerpt captures the closing moments of the play session, during which the child wants to continue playing with the toy vehicles.

#CHI Child #INV Investigator #MOT Mother

CHI	/it's/ taż-żibe[ċ]	*żibel*
INV	issa ħa mmur jiena Toninu (.) [[orrajt]]?	
CHI	[[imma]] nkompli nilgħab (.) b'daw'	
INV	ħa tkompli tilgħab b'dawk (.) inti?	
CHI	iwa	
INV	iva?	
INV	ee dawk issa ħa neħodhom miegħi (.) orrajt?	
INV	u mbagħad tista' tilgħab bit-/toys/ tiegħek inti (.) tajjeb?	
CHI	imma jiena [jj]id nilgħab (.) b'daw'	*rrid*
INV	b'dawk trid tilgħab?	
CHI	iwa	
INV	ok	
INV	ftit (.) ftit ieħor (.) tajjeb?	
CHI	u teħodhom[j]i	*teħodhomli*
MOT	imbagħad teħodhom magħha	
MOT	għax trid tagħtihom lil xi ħadd ieħor	
INV	[[eżatt (.) eżatt]]	
MOT	[[biex tilgħab magħhom]]	
MOT	eh?	
INV	tajjeb?	
CHI	[[iwa]]	
MOT	[[ok?]]	
INV	għax inti għandek ħafna /toys/ (.) hux vera?	
CHI	iwa	

#Toninu 56months

#Child's age on the day of the recording 4;08.01

#Both parents were present but uninvolved. The child was talkative, imaginative, and expressive, showing a marked increase in English use. His speech was clear with only minor articulation gaps. Here, he comments humorously on novel images during the novel-word plural task.

#CHI Child #INV Investigator #MOT Mother #FAT Father

INV	ara (.) gadma	
INV	gadma...	
CHI	/what is a/ gadma (.) /with English/?	
INV	ħmm (.) naħseb gadma tibqa' (.) jiena	
INV	gadma (.) u issa haw' ħafna _	
CHI	gadmijiet	
INV	orrajt (.) tini /high five/	
INV	ħa nara l-kelb (.) jaf itini /high five/?	
CHI	iwa (.) a[j]a	*ara*
INV	bumm	
INV	ara (.) torħa	
INV	torħa waħda (.) iss' haw' ħafna _	
CHI	torħijiet	
INV	bravu	
INV	ara (.) seff	
CHI	/but that... but tha[n]t's... but that is stinks (.) a lot/	
INV	<gasp> jista' jkun vera (.) jaqq	
CHI	/it stinks... it stinks/ għax (.) għamel pipì taħtu	the novel image has a yellow shadow which the child interprets as pee
INV	ee veru	
INV	mela dan seff	
INV	u issa haw' ħafna _	
CHI	seffijiet	
INV	bravu	
INV	ara (.) dan nedimist	
INV	[[nedimist (.) iss haw' ħafna _]]	

CHI	/[[he's... he's a little]] like a chicken/	
INV	veru u xi /penguin/ (.) hux veru?	
CHI	/yes (.) actually he's a chicken–penguin/	
INV	tajjeb	
INV	mela da' jismu nedimist	
INV	iss' haw' ħafna _	
CHI	nedimisti	
INV	ara (.) qonda	
CHI	/he… he's a little like a (.) duck–snake/	
INV	eżatt ank[a] jien (.) hekk naħseb	

#Kelina 36 months

#Child’s age on the day of the recording 2;11.24

#Both parents were present, and the mother assisted throughout. The child, an only child attending childcare, was outgoing, articulate, and imaginative, with mostly intelligible speech. This excerpt shows her in “teacher mode,” directing a class of soft toys.

#CHI Child #INV Investigator #MOT Mother

CHI	ma[[jj]et ti[jj]emetti	*marret tirremetti*
MOT	u fej’ marret tirremetti?	
CHI	ma[[jj]et hemm fuq it-/toilet/	*marret*
MOT	fit-/toilet/ marret?	
MOT	[[ee]] aħjar mela (.) ma ħamġitx mela	
INV	[[ee]]	
MOT	eh? ma ħammġet imkien ħi?	
CHI	mhm	
INV	tajjeb (.) [[aħjar]]	
CHI	ejja quddiemi	
INV	ee ħa tiġi quddiemek (.) orrajt	
MOT	ajma	
CHI	iktar minn di’ għoli	
INV	ee vera iktar għolja l-banketta	
MOT	ħ... fil-baxx ħa tkun inti hu’?	
CHI	issa kif nista’ ngħaddi? [[<laughter>]]	
MOT	[[<laughter>]]	
INV	[[kif ħa tgħaddi int? <laughter>]]	
MOT	[[inti fej trid tgħaddi eżattament ħij?]]	
INV	[[ejja nressqu naqra di’ ara]]	
CHI	haw’	
MOT	ħa toqgħod fuq il-banke...	
MOT	eh ħa tgħaddi lil haw’	
CHI	<struggling sound>	
INV	[[ee tajjeb]]	

MOT	[[hekk ħa toqgħod inti?]]	
CHI	fej’ qegħdin il-Weetabix aa?	
MOT	ee [[ee]] dal-Weetabix kien?	
INV	[[haw’]]	
CHI	il-Weetabix	
MOT	ee orrajt ħa ħi	
INV	ee [[Weetabix]]	
MOT	[[qed tagħmel]] riklam [[bil-Weeta... <laughter>]]	
CHI	[[da’ bil-frott ta]]	
CHI	għalhekk għax k... kellha l-frott irremettiet għalhekk	
INV	kellha (.) x’kellha?	
MOT	[[il-frott]]	
CHI	[[kellha frott]]	
INV	ee kellha l-frott	
MOT	ma jogħġobhiex il-frott?	
CHI	le	
MOT	imma di’ frawla .	
CHI	imma la[q]qas tħobbhom ta	

#Kelina 59 months

#Child's age on the day of the recording 4;11.28

#The mother participated throughout to facilitate interaction. The child was unusually reserved and clingy, possibly due to heightened awareness of being recorded or the fact that the mother was expecting a second child. Later that day, the mother recorded an additional interaction. The excerpt focuses on the *Frog Story* (Mayer, 1969) narrative task.

#CHI Child #INV Investigator #MOT Mother

MOT	ara l-kelb xi ġralu Kelina	
CHI	il-kelb waqa'	
INV	<gasp>	
CHI	[u]mbagħad qed jirrabja miegħu	
INV	[[[u]mbagħad (.) qed _]]	
MOT	[[qed jirrabja miegħu?]]	
INV	ee qed jirrabja [[miegħu]]	
MOT	[[għalfej']] qed jirrabja miegħu?	
CHI	għax waqa'	
MOT	għax waqa'?	
INV	ee	
CHI	<breath> [u]mbagħad... <breath> [u]mbagħad marru jgħidulu	
INV	eżatt (.) [[marru]]	
MOT	[[lil min Kelina]] (.) marru jgħidu?	
CHI	'iż-żrinġ	
MOT	'iż-żrinġ?	
CHI	[u]mbagħad (.) raw xi ħaġa <breath>	
CHI	u... u... u ma setgħux ikunu jarawh	
INV	eżatt (.) ma kinux ċerti fej' qiegħed	
INV	[u]mbagħad?	
CHI	[u]mbagħad b[e]dew jistennew iktar	*bdew*
INV	beda... x'għamlu?	
CHI	b[e]dew jistennew iktar	*bdew*
INV	b[e]dew jistennew iktar (.) orrajt	*bdew*
INV	u issa?	

CHI	[u]mbagħad it-tifel kompl[i] jfittex	*kompla*
INV	mhm	
CHI	u hemmhekk qed jagħmel hekk biex forsi jsib iż-żrinġ	
INV	tajjeb (.) orrajt	
CHI	u hemmhekk <breath> qed isib (.) iż-żrinġ biex... biex... biex... biex imorru d-dar	
INV	tajjeb (.) eżatt	
INV	u haw' kokka (.) hux vera?	
INV	qed ittir (.) fuq rasu	
CHI	u haw'hekk qed jgħajjatlu	
INV	qed jgħajjatlu	
INV	u issa?	
CHI	[u]mbagħad ra /reindeer/	
INV	eħe (.) eżatt	
INV	ra /reindeer/ (.) bħal ta' /Father Christmas/	
INV	ara fej' tela'	
INV	fej' tela'?	
CHI	haw'	
INV	[[eżatt (.) fuq]] rasu	
MOT	[[mm]]	
CHI	[u]mbagħad marru (.) ifittxuh bir- /reindeer/	
INV	eħe	
MOT	iqq	
CHI	u hemmhekk waqa'	
INV	<gasp>	
MOT	issa sew (.) <gasp>	
CHI	u hemmhekk spiċċaw ġol-ilma	
INV	spiċċaw ġol-ilma (.) eżatt	
CHI	u hemmhekk <breath> qed jisimgħu xi ħaġa	
INV	qed jisimgħuh	
INV	ank[a] jien qed narahom (.) veru	
CHI	<noise>	

CHI	u hemmhekk (.) qed jagħmlu hekk biex isibuh	
INV	biex isibuh	
INV	taħseb li ħa jkun hemm inti?	
CHI	<click>	
INV	veru?	
INV	ħa[n] naraw	
INV	eqq (.) x'sabu?	
CHI	il-/mummy/ tiegħu	
INV	il-/mummy/ tiegħu (.) iva	
INV	u haw' ħafna (.) hux veru?	
INV	mhux wieħed	
INV	u [u]mbagħad?	
CHI	u [u]mbagħad (.) qalilhom /bye/ u... u marru d-dar	

#Reno 36 months

#Child's age on the day of the recording 3;00.13

#The mother assisted intermittently. The child, who attends childcare, was initially shy but grew more confident. Speech was often unclear, with the mother repeating utterances for clarity. In this excerpt, he searches for his ball and discusses a recent fall.

#CHI Child #INV Investigator #MOT Mother

INV	tħobb tilgħab inti bil-ballun?	
CHI	ijwa	
INV	iva?	
CHI	għandna għandikun	
CHI	fej' qiegħed?	
INV	fej' qiegħed?	
MOT	x'inhu ħi?	
CHI	/ball/	
MOT	hemm ara	
CHI	fej' qiegħed?	
MOT	fit-trakk	
CHI	fej' qiegħed?	
MOT	idħol ġol-kamp	
INV	wow kemm għandek kamp sabiħ	
CHI	fej' qiegħed?	
MOT	is' isa' quddiemek	
MOT	hemm (.) proset (.) hemm	
MOT	le (.) it-trakk it-trakk	
MOT	ħares lej' it-trakk Reno (.) ġol-... ġol-kaxxa tat-trakk (.) eh	
CHI	[w]allun	*ballun*
MOT	<chuckle>	
INV	ii kemm għandek ballun sabiħ	
CHI	aħħ dahri	
CHI	aħħ dahri mamà	
MOT	aw ħi	
CHI	dahri qiegħe' juġagħni	*qiegħed*

MOT	qed juġgħu dahru għax ilbieraħ waqa' miskin	
INV	oo miskin (.) qed juġgħek dahrek?	
CHI	quddiemu da'	CHI showing INV where he fell
CHI	quddiemu da'	
MOT	toqgħodx hemm	
CHI	noħodhom?	
MOT	ħudhom ħudhom [[ħi]	
INV	[[wow]	
CHI	/look/	
INV	illa' kemm hi sabiħa dik	CHI drops a toy
INV	ara attent għandek it-/toys/ (.) eħe	
INV	ġibha ġibha haw'hekk <noise>	
CHI	bumm	
INV	bumm	
CHI	fejnha d-[[/doggy/]]?	

#Reno 60 months

#Child’s age on the day of the recording 5;00.10

#The mother was nearby but non-participatory. The child, now confident and talkative, demonstrated significant improvement in speech, which was fully intelligible. Occasional Żurrieq dialect features appear in his Maltese. This excerpt features a lively exchange about toy cars and a “cow show.”

#CHI Child #INV Investigator #MOT Mother

INV	sbieħ tiegħek	
INV	eżatt joqogħdu [[fuqha]]	CHI playing with toy cars
CHI	[[l-aħħar]] darba raj... rajt xow ta’... tal-... tal-baqar	
CHI	taf kif tahielu d-daqqa ta’ ras ’ill-ieħor? <chuckle>	
INV	qq... x... x’rajt?	
INV	xow tal- _	
CHI	xow tal-baqar	
INV	veru?	
CHI	mela (.) taf kemm kien tad-daħq?	
CHI	ill' kif tahielu 'l rasu bid-daqqa l-ieħor	
INV	ir-... il-... haw’	
INV	[[baqar ta’ veru?]]	
CHI	[[ir-raġel]]	
CHI	mela	
INV	ara ara	
CHI	ill' kif tahielu d-daqqa	
INV	qatt ma rajt jiena <laughter>	
CHI	imma konna xi mkien (.) ħdej’ l-annimali	
INV	ee	
CHI	[l]ajna ħafna ta (.) imma .	*rajna*
CHI	jistennew imma qegħdin	CHI now talking about toy cars
CHI	[[jis’ennew jku...]]	*jistennew*
INV	[[ee għax]] dawk iwaħħlu hemmhekk hux?	
INV	għax [[ftit twal]]	
CHI	[[mela ħa]] nressqu (.) 'l dawn	

INV	mhm?	
CHI	eh (.) ħe...	
INV	<laughter>	
CHI	haw' ħa nikser	
INV	mhm?	
CHI	di... di' tiġri?	
INV	tiġri ij... eżatt	
INV	tiha waħda b'lura ara (.) rankatura u tiġri	
CHI	anka jien għandi waħda	