The Acquisition of L2 English Morphosyntax: The Developmental Path of Acquisition by Low-Educated/Low-Literate L1 Arabic Learners and Overgeneralisation from a Generative Perspective

Abdulelah Wazen Alsulami

A thesis submitted to the School of English Literature, Language and Linguistics in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at

Newcastle University

Supervised by Prof. Martha Young-Scholten, Dr. Elaine Lopez and Dr. Gabriel Martinez Vera

November 2024

Abstract

The study is aimed firstly to examine the role of literacy in adult immigrant learners' development of English morphosyntax using the stages of Organic Grammar (or OG). Literacy—from the perspective of generative linguistics—is not thought to have a direct influence on the development of syntax. This does not mean that it plays no role in the acquisition of certain aspects of functional morphology, and emerging research on oral production of non/low-educated/literate learners shows some connection between literacy and the development of morphosyntax. To achieve this aim, 60 participants (with varying levels of literacy and amounts of native language schooling) were recruited to participate in a study. Data were collected through production tasks (picture descriptions) and comprehension tasks (computerized tasks). The results show that the developmental path of acquisition is systematic. There is a positive connection between literacy and the acquisition of morphosyntax.

The second aim is to further explore learners' overgeneralisation of functional morphemes (the use of non-target function words or multi-word utterances) during the acquisition of second language (L2) English morphosyntax and to confirm the specific Organic Grammar stage of L2 morphosyntax development at which L2 learners engage in such an overgeneralisation. The results indicate that: (1) overgeneralisation occurs after the Verb Phrase (VP) stage. (2) Overgeneralisation does not alter the general developmental path of acquisition (i.e., the stages of Organic Grammar). (3) Overgeneralisation can involve morphemes, words, and multi-word sequences. (4) Overgeneralised forms are placeholders that learners temporarily use as they work on identifying the relevant heads in the input. (5) the use of certain function words (as placeholders) seems to be peculiar to the VP-stage learners, and higher OG-stage learners use different types of place-holders (e.g., the copula be or personal pronouns), thereby supporting the recent predictions of placeholders through the stages of Organic Grammar.

Acknowledgements

First, I would like to express my deepest gratitude to my supervisors, Prof. Martha Young-Shelton, Dr. Elaine Lopez, and Dr. Gabriel Martinez Vera, who replaced Prof. Martha after her retirement, for their invaluable support, guidance, encouragement, feedback, and patience throughout my PhD research journey.

I am profoundly thankful to my wife, Dr. Taghreed Alsulami, and my children, Faisal and Tala, for their unwavering support, patience, and love, which have been a constant source of strength and motivation. I am also grateful to my youngest child, Alqays, born in 2020, for bringing joy and moments of relaxation during the challenging times of my PhD journey.

My heartfelt thanks go to my mother, Aliah, my brothers and sisters, and my uncle, Saad, for their continuous support and encouragement, which have provided me with much-needed strength.

I would also like to thank my friend, Dr. Mohammad A. Jareebi (Jazan University), for his ongoing encouragement and consultation on the statistical analyses adopted in this thesis. I am equally grateful to my friend, Mariano Gabriel, for his support with my PhD LaTeX document; his expertise in LaTeX has been invaluable, and I have learned a great deal from him.

I am sincerely grateful to my government, Saudi Arabia, for providing the scholarship that enabled me to undertake this PhD.

I extend my gratitude to Newcastle University, including the School of English Literature, Language and Linguistics, the library, IT services, and the Graduate School, for their invaluable educational support throughout my studies.

Special thanks are due to my examiners, Prof. Michelle Sheehan (Newcastle University) and Prof. Ann-Kristin Gujord (University of Bergen), for their valuable feedback and constructive criticism.

I am deeply appreciative of Dr. Egle Mocciaro (Masaryk University) for granting

me access to her PhD work both before and after its approval, which provided valuable insights for my own research.

Lastly, my heartfelt thanks go to the participants whose time and involvement made this PhD possible, and to the coordinator, Richard, for helping me connect with these participants. I am also grateful to Julia, the native speaker who recorded the stimulus, and Sharifa, the illustrator who created the images for my research methods.

Contents

List of Figures

List of Tables

List of Abbreviations

AgrP	Agreement Phrase
Aux	Auxiliary
BSM	Bilingual Syntax Measure
CAH	Contrastive Analysis Hypothesis
Cop	Copula
IP	Inflectional Phrase
L1	First Language
L1A	First Language Acquisition
L2	Second Language
LESLLA	Literacy Education and Second Language Learning for Adults
MSBA	Modulated Structure Building Approach
OG	Organic Grammar
PT	Processability Theory
TP	Tense Phrase
UG	Universal Grammar
VP	Verb Phrase

Chapter 1. Introduction

The route of language development is argued to be uniform regardless of (1) the learner's native/first language (L1), (2) the learner's age at initial exposure to the target language, (3) the type of exposure, and (4) the learner's educational background (Hawkins 2001). This conclusion is partly drawn from the morpheme order studies of second language acquisition and the extensive second-language acquisition studies of naturalistic adult immigrants conducted in the 1990s and 2000s. However, the learner's educational background/L1 literacy has received less attention in the field of SLA, either in the generative or the Usage-based approaches. Using the Noticing hypothesis, Tarone, Bigelow, and Hansen (2009) looked at production data from non-/low literate L2 learners and argued that literacy has an undeniable effect on the acquisition of morphosyntax. This line of research has important implications for the generative approach, where non-linguistic factors, including literacy, play no role in developing linguistic competence. Vainikka and Young-Scholten (2007) took steps to investigate the role of literacy under their Organic Grammar, which adopts a generative approach and claims that literacy has an indirect role in the development of morphosyntax. However, this role of literacy does not affect the route of acquisition but rather the rate. Therefore, the first aim of this study was to confirm the route of acquisition under Organic Grammar with a larger sample and a more detailed analysis, including implicational analysis for assigning OG stages to the learners and statistical analysis to investigate any potential role of literacy in the development of morphosyntax.

More recent studies on low-/non-educated immigrant adults have looked at overgeneralisation within different approaches. Applying the Minimalist Program, Julien, Van Hout, and Van de Craats (2016) tested a claim from previous studies (see, e.g., Van de Craats and Van Hout 2010) in which L2 non-literate learners use dummy auxiliaries (i.e., overgeneralised forms) during their acquisition of Dutch. Julien *et al.* argued that learners' misuse and overgeneralisations of certain morphemes mirror their developmen-

tal stage. The learners used dummy auxiliaries to mark subject-verb agreement and tense before attaching the correct Dutch suffixes to the main verbs. In a similar vein, Vainikka, Young-Scholten, Ijuin, and Jarad (2017) looked at the acquisition of L2 English by adult learners with no or low levels of L1 literacy. As with Julien *et al.*, Vainikka *et al.* argued that learners' oral production included overgeneralisations only at specific developmental stages of acquisition beyond the Verb Phrase (VP) stage. The researchers also reported that learners not only used single words as overgeneralised forms but also used multi-word utterances (e.g., 'in the'). In a follow-up study, Vainikka and Young-Scholten (2019) argue that overgeneralisation occurs as a result of the learner's search for the relevant syntactic head in the input. Once learners find the relevant head, they exhibit fewer overgeneralisations. A similar claim was made under the Basic Variety, which takes a functionalist approach. Mocciaro (2019b) examined the acquisition of L2 Italian morphosyntax by low-educated adult learners. In line with existing L2 research, Mocciaro claimed that learners use overgeneralised forms as soon as they move from the basic variety. Therefore, the study seeks, as a second aim, to explore overgeneralisation under Organic Grammar through production tasks and newly designed comprehension tasks.

This study contributes in three ways to what is already known under Organic Grammar. First, the study includes a larger sample compared to Vainikka *et al.*'s (2017) and provides a more detailed analysis through the application of implicational analysis, which has been previously used in both longitudinal and cross-sectional data, alongside statistical analysis to investigate the correlation between OG stages and literacy levels. The second contribution is the inclusion of comprehension data to explore whether it reflects the production data in terms of the difficulties learners face with bound morphemes compared to free morphemes. The third contribution is that the study explains how the OG stages influence the choice of overgeneralised forms, as the production data show that participants use different classes of overgeneralised forms. The comprehension data were also utilised to test the claim of Organic Grammar that overgeneralised forms serve a syntactic function, specifically examining whether overgeneralisation occurs as a result of learners searching for the syntactic head in the input.

In this thesis, I will use the following definitions listed in (1.1):

(1.1) a. Non-literate (or: illiterate): an adult who never went to school and cannot

read and write, neither in his/her first language, the standard language of the country of origin or the second language.

b. Low-literate: an adult who has attended school, but who has a reading level below the average primary school level.

c. Low-educated: an adult who has at most ten years of education in the country of origin. For many adult immigrants and refugees, this means at most primary education.

(Craats, Kurvers, and Young-Scholten 2005: 8)

1.1 Adult migrants with limited education or literacy

Globally, illiteracy rates are gradually declining; now, 89.9 percent of males and 82.6 percent of females over the age of 15 are able to read and write. Yet, 75 percent of the world's population, or 750 million people over the age of 15, live in poverty, rural areas, or war-torn countries and have little access to formal education and low or no literacy in their native languages. Women make up two-thirds of the 750 million (World Demographics Profile 2018).

Educators who deal with adult migrants with little education and literacy skills note the many linguistic barriers and literacy challenges their students confront as they attempt to adapt to their new countries. Meeting these challenges is complicated by various factors, ranging from visa requirements and restrictions to the availability of basic skills lessons offered by instructors and tutors specially trained to deal with such a population (see, e.g., Tarone *et al.* 2009).

Educational provision for adults with little or no formal schooling or home language literacy is much more alike than different worldwide. For instance, in most countries, learners' earliest steps toward attaining a certificate are not formally recognised (e.g., Entry 1 in the United Kingdom). In some countries, unpaid volunteer teachers play a critical role in one-on-one and small-group sessions and classroom instruction and may have the responsibility for their own courses.

Immigrant adults who arrive in highly literate societies without the skills necessary for living-wage employment and who expect to support their younger family members in numerous critical aspects (e.g., education and well-being) require at least as much attention from educators as those who arrive with sufficient schooling and high levels of literacy.

These individuals may benefit from the opportunity to learn in informal educational settings with well-trained, qualified teachers. Based on similarities in educational provision and a pressing need to share accumulated knowledge, academics and practitioners from the countries where these adult migrants settle have begun collaborating, leading to a new organisation. In 2005, Tilburg University in the Netherlands hosted the first workshop for adults with little or no home language schooling or literacy. From this workshop, LESLLA (Literacy Education and Second Language Learning for Adults; formerly known as Low-Educated Second Language and Literacy Acquisition) was founded.

LESLLA grew from a small workshop in Tilburg with fewer than 25 academics, practitioners, politicians, and students to a formal international and multidisciplinary organisation with more than 200 members worldwide. The purpose of this workshop was to exchange ideas about how to aid this population in learning to read and write in a new language for the first time by sharing research results, practical pedagogical approaches, and information on literacy policy and second language acquisition research and instruction. To date, researchers, educators, and policymakers have devoted significantly less attention to this population of adult learners.

LESLLA is unique in that it is the only organisation committed entirely to investigating issues pertaining to low-literate adult migrants. This database includes related topics of research (e.g., reading and oral language development and practice, as well as technology-based learning strategies).

Researchers have found that, compared to adult migrants who have some knowledge of their native language and literacy, these learners' literacy development progress is substantially slower (Condelli, Wrigley, and Yoon 2002; Kurvers, Stockmann, and Craats 2010), and can also be slower for oral language development (Tarone *et al.* 2009). This observation questions the assumption that learners, regardless of age, amount of education, or type of exposure, can reach high levels of oral linguistic competence in a new language (Hawkins 2001). However, only a tiny proportion of adult migrants who begin learning to read and write for the first time in their lives succeed, and the majority struggle to improve these literacy skills and linger around or below A1 in the Common European Framework of Reference for Languages (Kurvers *et al.* 2010; Schellekens 2007).

Oral language development is also slower for illiterate adults, as their initial source of language input is limited to listening (Tarone and Bigelow 2005). If they have no literacy

skills, they will be unable to decode printed materials. Older migrants may process oral language input more slowly and recall things less effectively than younger migrants (Tarone *et al.* 2009).

With respect to the acquisition of morphosyntax, researchers found that more literate learners were better than those with lower literacy levels at recalling and accurately repeating and at supplying inflectional morphology (Bigelow, Delmas, Hansen, and Tarone 2006; Bigelow and Tarone 2004; Tarone *et al.* 2009). Under Organic Grammar (or OG), Young-Scholten and Strom (2005) observed a positive correlation between literacy and the OG stages.

1.2 Outline of the thesis

This thesis is organised as follows. Chapter 2 first reviews the early studies on L1-L2 functional morpheme orders (Bailey, Madden, and Krashen 1974; Brown 1973; De Villiers and de Villiers 1973) and the developmental stages of negation (e.g., Cancino, Rosansky, and Schumann 1978; Ravem 1974), followed by influential studies on the acquisition of morphosyntax by low-educated adult L2 learners (see, e.g., Clahsen, Meisel, and Pienemann 1983; W. Klein and Dittmar 1999; W. Klein and Perdue 1997; Vainikka and Young-Scholten 1994, 1996a,b) and their subsequent developmental stage theories: Processability Theory (Pienemann 1998), the Basic Variety (W. Klein and Perdue 1997), and Organic Grammar, formerly known as Minimal Tree (see, e.g., Vainikka and Young-Scholten 2005). Findings from these studies indicate that the route of language development is uniform and unaffected by general cognitive factors. For example, Hawkins (2001) reviewed these studies and concluded that the developmental path of L2 morphosyntax is fixed, regardless of learners' L1, type of exposure, or education. Using Schmidt's (1994) Noticing Hypothesis, Tarone and Bigelow (2005) and Tarone *et al.* (2009) contest this view and argue that literacy has an undeniable effect on the acquisition of morphosyntax. This chapter also includes studies (Vainikka and Young-Scholten 2007; Vainikka *et al.* 2017; Young-Scholten and Strom 2005) on the relationship between literacy and the acquisition of morphosyntax. These studies were carried out under Organic Grammar (Vainikka and Young-Scholten 1994, 1996a,b, 2011), which takes a generative approach. Again, analysis of these studies shows that the route of L2 morphosyntax is fixed and unaffected by literacy.

Chapter 3 begins by briefly summarising overgeneralisation in L1 (see, e.g., Berko 1958), followed by overgeneralisation in L2 (see, e.g., Fleta 2003; García Mayo, Ibarrola, and Liceras 2005; Ionin and Montrul 2010; Lakshmanan 1993/94) and the relevant accounts, supported by evidence from several studies. This section also addresses the overgeneralisation of multi-word strings (see, e.g., N. C. Ellis 2003; N. C. Ellis 2012; Eskildsen and Cadierno 2007; Myles 2004; Myles, Hooper, and Mitchell 1998). Closely connected to the second aim of this thesis, this chapter also explores overgeneralisation by adult L2 immigrants, which has been investigated under three different theoretical frameworks: the Minimalist Program (Chomsky 1995), Organic Grammar (Vainikka and Young-Scholten 1994, 1996a,b, 2010), and the Basic Variety (W. Klein and Perdue 1997).

Chapter 4 describes the methodology used in this thesis, including the participant recruitment procedure, participant descriptions, and instrument descriptions (tests and tasks). Data were collected through oral production tasks (picture descriptions) and comprehension tasks (online auto-picture matches). Literacy was measured through a reading assessment, and the route of morphosyntax was investigated through Organic Grammar. This chapter also details the analysis adopted in this thesis, including counting procedures, implicational analysis used for assigning OG stages to learners, and statistical analysis used to investigate any potential relationship between literacy and OG stages, as well as any potential relationship between lower morphosyntax in comprehension and the use of overgeneralisations.

Chapter 5 will first present the results of reading/literacy levels, as the first aim of this thesis is to examine the role of literacy in adult immigrant learners' development of English morphosyntax using the stages of Organic Grammar (specifically, the role of literacy in the developmental path of acquisition). This chapter will also present the results of the verbal morpheme order through production and comprehension data, followed by the developmental stages of acquisition. To investigate the developmental stages, I used Vainikka and Young-Scholten's (1996a,b) criteria for assigning each learner an OG stage and Déprez and Pierce's (1993) predictions with respect to word order in L1 English negative sentences. This OG placement will, in turn, help to investigate the use of overgeneralised forms introduced in the last section of this chapter. To interpret whether overgeneralised forms are placeholders as claimed by Vainikka and Young-Scholten (2019), I looked at whether these forms precede inflected verbs, and I also looked at whether

learners who produced overgeneralised forms struggled with the comprehension of the third-person singular *-s* or the past tense marker *-ed*.

Chapter 6 will discuss the results of morphosyntax in terms of (1) the route of morphosyntax and (2) the rate of acquisition. Based on existing research, the route of acquisition, rather than the rate, seems to be unaffected by literacy. This consistency in the route of acquisition applies to all learners in this study. However, education or literacy levels correlate with OG stages, suggesting that literacy may impact the rate of language development.

In this chapter, I will also discuss the use of overgeneralised forms in the structure-building of Organic Grammar. Specifically, I will explore the choice and frequency of overgeneralised forms in relation to OG stages, which seem to impact both the choice and frequency of these forms. I will explain how OG stages influence the selection of overgeneralised forms. Later in this chapter, I will delve into the use of multi-word strings as placeholders compared to chunks in Exemplars. I will also propose possible reasons why VP-stage learners tend to use multi-word strings as placeholders and argue that the overgeneralised forms in my data do not feed into grammar but rather serve as placeholders for the relevant syntactic head. Lastly, before concluding this chapter, I will address whether L1 transfer accounts for the overgeneralisation of subject personal pronouns.

Chapter 7 provides a summary of the results in relation to the aims and the hypotheses investigated in this thesis, stating the main research findings, limitations, and suggestions for future research work.

Chapter 2. Stage-like development in morphosyntax and the role of literacy

2.1 Introduction

This chapter reviews the major trends in second-language acquisition research on the acquisition of morphosyntax over the past four decades. This research has focused on the order of functional morphemes as well as the stages of language development. Both morpheme order and the stages of language development studies have important implications for the developmental stages or route of acquisition for low-literate/low-educated L2 adult immigrants. Therefore, Section 2.2 briefly reviews the history of second language acquisition prior to the advent of Universal Grammar, with an emphasis on morpheme acquisition order studies traced back to Brown's (1973) study. In Section 2.3, I cover the first developmental longitudinal studies conducted in the 1970s, which dealt with the acquisition of negation. Section 2.4 covers the seminal studies on the developmental stages of low-/non-educated adult immigrants, including subsequent developmental stage theories and their critiques. Section 2.5 presents the generative approach to language acquisition, which first discusses Strong Continuity and the Structure-Building approach, both of which take a generative perspective on language acquisition, showing how Strong Continuity stands in opposition to Weak Continuity or the Structure-Building hypothesis, based on which Vainikka and Young-Scholten propose Organic Grammar. Section 2.6 covers the theory of Organic Grammar adopted in this thesis, both in its syntactic and developmental stage theories. This section also discusses the idea of morphosyntactic triggers developed under Organic Grammar, followed by supporting and counter-evidence to Organic Grammar. For the sake of completeness, Section 2.7 presents the Modulated Structure Building Approach, again adopting a generative perspective and taking a similar view to Organic Grammar.

Under the generative approach, all these studies (morpheme order, early developmental stages, and seminal studies on the developmental stages of low-/non-educated adult immigrants) have been taken as evidence that learners' routes of acquisition are uniform regardless of age, type of exposure, and L1 education or literacy (Hawkins 2001). Therefore, section 2.8 covers the opposing view that literacy plays a role in the development of morphosyntax (Tarone and Bigelow 2005; Tarone *et al.* 2009) and the stance of the generative approach, where non-linguistic factors play no role in acquisition where Vainikka and Young-Scholten (2007), based on their studies in the 1990s and Young-Scholten and Strom's (2005) study, investigated the role of literacy in the L2 development of morphosyntax under their Organic Grammar theory (see, e.g., Vainikka and Young-Scholten 2010). Vainikka and Young-Scholten's (2007) conclusion is that 'literacy indeed plays a key—but likely very complex—role' (p. 127). However, this role of literacy appears not to affect the route of acquisition. Vainikka *et al.* (2017), again under Organic Grammar, explored the L2 route of English by non-low-educated adult learners. Vainikka *et al.* concluded that the route of acquisition is fixed and unaffected by non-linguistic factors such as literacy.

In generative linguistics, input is a key factor in the acquisition of morphosyntax; however, low-educated learners are assumed to be far from interactions with native speakers (see, e.g., Mocciaro 2019b; Vainikka and Young-Scholten 2011). Therefore, Section 2.9 will discuss the role of input in language acquisition. Before closing this chapter, I will discuss some difficulties that LESLLA learners could potentially face in Section 2.10. These include trauma, ongoing stress (migration-related stress), and motivation in LESLLA learners. As indicated in Chapter 1, the first aim of this thesis is to investigate the developmental path of acquisition by low-literate learners under Organic Grammar.

2.2 Early approaches to second language acquisition

2.2.1 Morpheme order studies for L1 and L2 acquisition of English

The morpheme order studies refuted the Contrastive Analysis Hypothesis (CAH) by demonstrating that children do not master the language through habit formation but rather suggest that they bring an internal mechanism to the task of language acquisition (i.e., based on a systematic order of acquisition). These morpheme order studies have

significant implications for stage-like development and the path of acquisition, which I intend to explore in low-literate/low-educated adult immigrants. The literature on the common order of acquisition for grammatical morphemes is rooted in the framework proposed by Brown (1973), who conducted a longitudinal empirical study of three monolingual American children up to the age of four. The author examined the acquisition of 14 grammatical morphemes in their first language, which had not been acquired when Brown began to collect data. The researcher engaged each child in play several times a month, and the sessions were audio-recorded and subsequently transcribed. To measure the children's speech samples, Brown calculated the mean length of utterance. This measurement instrument made it possible to appreciate how new knowledge expanded the length of utterances. To determine whether the children had acquired the target morphemes, these morphemes had to emerge 'in 90 percent of all obligatory contexts for three successive two-hour samples' (ibid.: 398). This is known as suppliance in an obligatory context. For instance, a child might say 'Daddy counting toy'. The missing auxiliary 'is' is obligatory in this context, but the child does not supply it; similarly, the missing plural *-s* is also obligatory, but the child does not produce it. The three children's acquisition of morphemes was categorised into five stages. In Stage I, the children produced only content words, that is, words without grammatical morphemes. From Stage II to Stage V, Brown observed that the 14 morphemes were acquired in a specific order, although each child required varying amounts of time. At the end of data collection, the author observed that all three learners had acquired English grammatical morphemes in a systematic manner, albeit at varying rates and ages.

This order was examined by De Villiers and de Villiers (1973) using a cross-sectional design. The researchers collected data from 21 monolingual children acquiring English, whose ages spanned those included in Brown's study. The participants ranged from 16 to 40 months old. Similarly, the researchers interacted with the children to gather hundreds of utterances from each child. They analysed these utterances, focusing on 8 of the 14 morphemes, and used the same technique of suppliance in an obligatory context to calculate a percentage showing morpheme accuracy. De Villiers and de Villiers concluded that there was a high degree of similarity between the orders studied in their analysis and those studied by Brown.

Despite the fact that both studies agreed on the order, De Villiers and de Villiers

argued that there are three potential determinants to account for this order: morpheme frequency, grammatical complexity, and semantic complexity. Brown (1973: 255), on the other hand, did not consider frequency a factor because he claimed that 'the order of acquisition is dependent upon relative complexity, grammatical and/or semantic'.

Similar to non-adult utterances produced by young children, the utterances produced by L2 child learners have demonstrated the use of linguistic mechanisms rather than general cognitive mechanisms to acquire morphosyntax. Prior to the studies of the 1970s, Corder (1967) put forward the idea that L2 learners' errors are systematic, and Selinker (1974) advanced the term 'interlanguage' to characterise L2 learners' linguistic systems. The earliest empirical investigations into the acquisition of grammatical morphemes in a second language revealed some surprising results. Dulay and Burt (1973, 1974b) conducted the first of two studies involving 151 Spanish-speaking 5- to 8-year-olds and 60 Spanish-speaking and 55 Chinese-speaking 6- to 8-year-olds who received different types and amounts of L2 English exposure. Dulay and Burt divided these children into three groups: the first group consisted of 95 children from Sacramento, California, who attended a US monolingual school where they received formal English instruction. The second group comprised 26 children from San Ysidro, California, who attended an English school but spoke Spanish at home. Learners in the third group were 26 children from East Harlem, New York City, who attended a bilingual school where both English and Spanish were spoken; however, they did not receive formal instruction in English. Unlike the methodology used by Brown (1973) and De Villiers and de Villiers (1973), which included play sessions, Dulay and Burt used the Bilingual Syntax Measure (commonly abbreviated as BSM) tool to capture children's oral production. This measure aimed to elicit simple syntactic structures through seven cartoon images and 33 questions designed to evoke spontaneous speech. Data analysis was similar to that used by De Villiers and de Villiers, focusing on morpheme accuracy.

To determine whether learners' errors were related to developmental linguistic mechanisms or due to the interference of the L1, Dulay and Burt (1974a) applied the BSM method on 179 L1 Spanish children aged 5-8 to extract 513 utterances containing errors. The researchers analysed only utterances that contained errors, which were then categorised into three groups. The researchers showed that all groups made similar errors. Dulay and Burt's (1973) results demonstrated that children did not rely on their L1 to

learn the L2. Indeed, 87.1 percent of errors were attributed to the type of development reported in Brown's (1973) study.

Regarding adult learners of L2, Bailey *et al.* (1974) included adults and expanded the variety of L1s: 73 participants from 12 different linguistic backgrounds, ranging in age from 17 to 55. Thirty-three participants were L1 Spanish speakers, whereas the other 40 participants came from a variety of typologically diverse linguistic backgrounds. These participants were learning English in the United States; some were enrolled in English as a Second Language (ESL) programmes to prepare for college studies, while others participated in continuing education ESL programmes. Bailey *et al.*, as noted by Hawkins (2001), do not mention the amount of previous exposure to English that their participants had. Unlike research involving young children, these studies also did not include play sessions; instead, they required learners to verbally describe visuals. Data analysis was then conducted in a manner similar to that adopted by De Villiers and de Villiers for the morpheme accuracy score. Although the accuracy profiles of the Spanish participants and non-Spanish participants varied, significant commonalities were observed. Progressive *-ing*, the contractible copula, and plural *-s* again emerged as the most accurately used morphemes, followed by possessive's and the third-person singular *-s*.

No.	Brown; De Villiers and de Villiers L1 children	Dulay and Burt L2 children	Bailey *et al.* L2 adults
1	Plural *-s*	Plural *-s*	Progressive *-ing*
2	Progressive *-ing*	Progressive *-ing*	Contractible copula
3	Irregular past	Contractible copula	Plural *-s*
4	Articles	Contractible auxiliary	Articles
5	Contractible copula	Articles	Contractible auxiliary
6	Possessive *'s*	Irregular past	Irregular past
7	third-person singular *-s*	Third-person singular *-s*	Third-person singular *-s*
8	Contractible auxiliary	Possessive *'s*	Possessive *'s*

Table 2.1: Morpheme order for L1 and L2 acquisition of English, adopted from Young-Scholten and Naeb (2020: 86)

However, there were statistically significant differences between L1 children, L1 adults, and child and adult L2 learners, prompting researchers to question the reasons for these discrepancies. Researchers also sought to determine the causes of the L2 order. For example, Goldschneider and De Keyser (2001) proposed an account for the L2 order based on the more prominent perceptual salience of the earlier-acquired grammatical morphemes. Additionally, there were criticisms of the data collection methods used, which were not comparable across the three populations (child L1 and L2 learners, and adult L2 learners).

Criticisms were also directed at data analysis methods, which included the use of suppliance in an obligatory context. In a longitudinal study of a young boy named 'Homer', acquiring English as a second language, Wagner-Gough (1975) observed that while the young boy produced the suffix *–ing* in obligatory contexts, he also used it in contexts where it was incorrect. For example, he used verbs plus *–ing* to refer to a completed action that happened in the past (e.g., 'Mark and Fred going' to mean 'Mark and Fred went'). This indicates that the function of this suffix in Homer's utterances was not yet similar to that of native English speakers of the same age. This implies that researchers should not only look at what the learner fails to supply in obligatory contexts when measuring interlanguage. Instead, researchers should consider where learners do not make mistakes and whether the function of a particular morpheme reflects the target language. Rosansky (1976) also criticised the nature of cross-sectional studies. As mentioned earlier, the groundbreaking studies in L2 morpheme order research are primarily cross-sectional, rather than longitudinal. I will revisit this topic in Chapter 4, where I further discuss the limitations of cross-sectional studies and the implicational analysis adopted in this thesis, which has been suggested in the literature as a potential solution for these limitations.

Researchers have continued to conduct longitudinal and cross-sectional studies to search for observable patterns in the acquisition of this group of English morphemes (e.g., Cox 2005, Kahoul, Vainikka, and Young-Scholten 2018, E. Klein, Stoyneshka, Adams, Pugach, Solt, and Rose 2004, Lardiere 2003, Haznedar 1997, Bliss 2006). The variable use of a given morpheme by learners remains a topic of lively debate. The key research questions are whether the morpheme and the syntax have been acquired and what effect the phonology has on whether the learner's production of a morpheme results in a consonant cluster that is not permitted in their L1 phonology (e.g., 'goes' vs. 'likes'). In early

studies, data were typically derived from various types of production, usually oral but sometimes written (e.g., free compositions).

2.3 Early developmental stage studies

During the same period, a group of L2 researchers focused on the development of specific linguistic phenomena over time. Unlike the morpheme order studies, their approach was longitudinal. The two major phenomena studied during that time were the development of negation and question formation. Since question formation is beyond the scope of this thesis, I will only discuss the development of negation. Before presenting the early research on stages, I will define what I mean by a stage.

2.3.1 Definition of stages

While the concept of stages has multiple applications, Ingram (1989) addresses all aspects of child language development, including phonological and semantic development, as well as the development of morphology and syntax. Ingram begins this chapter by defining the idea of stages and suggesting possible criteria for these stages: (1) plateau; (2) transition to a further stage; (3) acceleration; (4) implicationality; (5) internal consistency; (6) underlying principle. When a given behaviour manifests consistently over time, we can isolate specific points from a lengthy developmental continuum. These points are referred to as stages. The rationale for defining these points is to highlight periods when the behaviour remains temporarily stable on a plateau. According to Ingram, development cannot be permanently halted if this behaviour is characterised as a stage. Instead, there must be a transition from that plateau to the next stage. Additionally, Ingram introduced a criterion concerning acceleration. Research findings on children's development have shown that before the temporary stabilisation at a plateau, the frequency of behaviour often increases dramatically. Although babies might skip a stage, they do not regress; for instance, babies do not walk before crawling. Ingram also introduced the concept of implicationality to explain the systematic order: Stage 1 precedes Stage 2, which logically precedes Stage 3, and so forth.

A stage can apply either to a single behaviour or to an associated cluster of behaviours. An early example of the latter is the stages of Piaget's intelligence development, in which sensorimotor and cognitive structures are thought to be tightly related (Piaget, Inhelder,

and Fraise 1969). With that in mind, an additional criterion is internal consistency, meaning that the behaviour or cluster of behaviours displayed at a specific stage follows a uniform pattern or group of patterns.[1]

One of the criteria for stages, as discussed above, is the transition from one stage to the next. This raises the question of how the learner progresses to the next stage. This concept is linked to the idea of triggers, which are components in the input that effectively push the learner forward, possibly because the triggering words or grammatical forms in the input cannot be integrated into the learner's existing grammar. In the following section, I present the early developmental stage studies on the development of negation.

2.3.2 The development of negation

Researchers suggest that second language English learners acquire sentential negation in a systematic manner. Decades ago, Ravem (1974) studied his Norwegian-speaking son, who was six and a half years old. Data were gathered over a three-month period and included both spontaneous creation and translation tasks. Unlike in English, where the negative clause element typically precedes the verb in the main clause, in Norwegian it follows the verb. Ravem maintained that there was no indication of L1 transfer in the early development of negation because the particle 'not', as in (2.1), precedes the verb in his son's L2 English utterances.

(2.1) I not looking for edge.

Cancino *et al.* (1978) proposed an early descriptive generalisation based on longitudinal data gathered from six Spanish-speaking L2 English learners—two children aged 5, two adolescents aged 11 and 13, and two adults. Data collection included spontaneous speech and repetition (or elicited imitation). The children and adolescents were exposed to English in a formal educational setting, while one adult worked in a factory and the other served as a babysitter for an English-speaking family. Cancino *et al.* established a four-stage developmental model through which all six learners progressed, as illustrated in Table 2.2.

[1]For instance, different word order patterns are observed during L1 or L2 German acquisition (Vainikka and Young-Scholten 2011). At a particular stage, one of these orders may dominate, while alternative word order choices may be less dominant.

Stage	Type of negation		Example
1	'no'	+ verb	You no tell your mother But no is mine, is my brother I no can see
2	'no' 'don't' (unanalysed)	+ verb + verb	He no like it He don't like it I don't can expalin
3	copula/ auxiliary	+ 'no/not'	It's not danger No, he's not skinny Somebody is not coming in He can't see
4	'don't' (analysed) ('no' + verb disappears)	+ verb	He doesn't laugh like us I didn't even know

Table 2.2: Stages of development in the acquisition of English negation

In reviewing Cancino *et al.*'s (1978) suggestion, Hawkins (2001) notes that Cancino *et al.* did not make more nuanced syntactic distinctions to establish a link between the absence or underspecification of inflection (I) in learners' mental grammar at this stage of development. More importantly for my data, as we will see later in Chapter 5, Cancino *et al.* do not, as Hawkins argues, make a distinction between tense and agreement when the sentential negator 'don't' becomes an analysed form (e.g., 'doesn't' for agreement).

Interestingly, they reported the phrase 'I don't know' in their studies and identified it as nonmorphic, therefore excluding it from the analysis. Based on their studies, which aimed to establish a path of negation development, this exclusion is practical because such a chunk would not provide accurate information regarding the use of 'don't'. These instances are similar to what my learners do, as we will see in the results chapter.

.

2.4 Influential studies on L2 acquisition by adult immigrants and stage-like development

In the early 1980s, the focus of converging language developments in syntactic theory shifted away from morpheme orders and the developmental path of various constructions to the Principles and Parameters of Universal Grammar (or UG). This shift in focus has neglected the possible role of literacy in the development of morphosyntax.

At the same time, with significant economic expansion after the Second World War,

Central-Western Europe became home to the largest number of migrants from all regions of the world, totalling around 70.6 million immigrants, mostly from North Africa and Turkey. From the early 1950s to the early '70s, the overall number of foreign workers in Europe increased from around 4 million to 10 million. In Germany, the most common destination for migrants, the proportion of foreign workers known as *Gastarbeiter* ('guest workers') increased from 0.6% to 11.2%, reaching 3 million at the start of the 1970s. The 1973 oil crisis and the ensuing global recession led to more stringent immigration policies and a significant decrease in the recruitment of foreign workers. However, the migration flow did not come to a halt. As a result of the high concentration of immigrant workers, there was a surge in second language acquisition research focused on this learner population in the 1970s and 1980s. These workers often received little or no native schooling and were therefore unlikely to receive L2 courses in their new countries due to their frequent status as temporary workers. Consequently, this section will present the underlying influential studies on L2 acquisition by adult immigrants and subsequent theories of the developmental stage, including some key criticisms. To maintain a logical order in this section, these theories are not arranged chronologically but are organised based on their foundational research.

2.4.1 The Heidelberger Pidgin Projekt

The Heidelberger Pidgin Projekt (see W. Klein and Dittmar 1999) was the first cross-sectional study of 48 immigrant manual workers. The participants were acquiring L2 German in naturalistic settings: 24 (16 male, 8 female) of them were L1 Italian speakers, and the other 24 (16 male, 8 female) were L1 Spanish speakers. All were over 18 years old. Their exposure to German was limited due to the marginalisation that characterised their migration experience. The researchers divided the participants into four groups by length of German residence: up to 2 years, 2–4 years, 4–6 years, and more than six years. The research project was motivated by the hypothesis that the immigrant learners were creating some sort of pidgin (as it turns out, they were not), given the large number of adult immigrants recruited to fill labor shortages. To test this hypothesis, the researchers collected production data through directed conversations and then extracted speech samples from 100 sentences for each learner for analysis. Thus, the researchers placed the learners into four proficiency groups. The study involved a control group

whose proficiency matched that of the highest L2 group. The lowest group supplied utterances with non-finite elements, the main verb, or a subject. Analysis of the data revealed a typical developmental progression of morphosyntax that various studies have since confirmed.

The researchers were able to account for variation in the rate of acquisition after two years of residence by including non-linguistic factors: the type of job, the amount of native schooling, the location where the immigrants resided, the intensity of interaction with native speakers, marital status, mobility, gender, age at testing, and attitudes. The adult immigrants in the Heidelberg project were manual workers who had received at least some primary native schooling and were socially, politically, and linguistically isolated. This was the first project to show, aside from the aforementioned sub-stage, that naturalistic adult learners 'do not act idiosyncratically; rather, stages of development are mostly typical for all learners, and they progress through these stages in a well-defined route' (1977: 45; MYS translation). However, when the rate of development of learners was compared to their proficiency level, length of residence, and other cognitive factors, the rate of development varied. Although learners with higher levels of schooling acquired German to varying degrees, those with lower education almost always fell into the lowest proficiency category. Although learners with higher levels of schooling acquired German to varying degrees, those with lower education almost always fell into the lowest proficiency category. We now move on to the second seminal study on adult immigrants, which was both cross-sectional and longitudinal, and it also included its subsequent developmental stage theory: Processability Theory, along with some key criticisms from two studies.

2.4.2 ZISA (Zweitspracherwerb italienischer und spanischer Arbeiter)

ZISA involved 105 learners from three different language backgrounds: Spanish, Portuguese, and Italian. This study examined the acquisition of L2 German word order by working-class adults and other immigrants to Germany who received no formal German instruction but were instead acquiring the language naturalistically. The researchers conducted two types of studies: cross-sectional and longitudinal. In the cross-sectional study, data were collected from 45 participants (20 L1 Italian, 19 L1 Spanish, and 6 L1 Portuguese speakers) through interviews. In the longitudinal study, data came from 12

adults every two weeks over two years. ZISA researchers, based on word order, established a systematic order of acquisition, beginning with what the researchers termed 'canonical', unmarked SVO, but also identified two variational features of acquisition within these learners. While all L2 learners followed the same developmental route in relation to word order in L2, they varied in the extent to which they produced functors (or grammatical morphemes) in their interlanguage. The first group favoured a standard orientation prioritising accuracy, while the second group favoured a simplifying orientation prioritising communication effectiveness (Clahsen 1983). In other words, the first group produced grammatical morphemes, while the second group appeared to omit them. Contrary to L1 children, Clahsen and Muysken (1986; 1989) argue that UG does not affect adult interlanguage development. They base their argument on learners moving the non-finite verb to the final position at Stage 3, which goes against certain syntax theories. Clahsen's (1991) study found no relationship between the development of agreement and verb raising in adult learners versus children. Pienemann (1998) extended this analysis to his Processability Theory, which suggests predictable stages. Before moving on to the next influential study, I will present Processability Theory.

2.4.2.1 Processability Theory (PT)

Pienemann identifies PT as the framework that elucidates the computational mechanisms required for language processing and acquisition. Early in the learning process, learners are unable to manipulate the L2 in a target-like manner due to limitations in these computational mechanisms. Within this framework, the primary challenge for learners is learning to exchange grammatical information across sentence elements.

> One of the main features of LFG is the unification of lexical features, which captures a psychologically plausible process consisting of (1) recognising grammatical information in the lexical entry, (2) temporarily storing that information, and (3) using it at another point in the constituent structure (p. 73).

This procedure of exchanging grammatical information is referred to as 'feature unification' in the LFG model. In languages where this is applicable, learners must ensure that a verb and its subject share the same number feature, or that a noun and its article share the same gender, number, and case features. The capacity to match features across elements and to move elements within a sentence develops gradually. The primary

premise of Processability Theory is that learners possess a Hypothesis Space that extends over time as they progress through a hierarchy of processing procedures. The theory's core premise is this hierarchy of processing mechanisms, but it also incorporates the concept of perceptual salience. According to Pienemann, perceptual salience implies that sentence beginnings and endings are easier to remember and manipulate. This predicts that learners will initially move elements to (or from) sentence-initial or sentence-final positions before moving elements within the sentence. Table 2.3 presents five stages of procedural development. These stages are lexically-driven (i.e., the operation of UG is not involved).

Processing procedures	Structural outcomes
5. Sub.-clause procedure	main and sub clause
4. S-procedure	inter-phrasal info, exchange
3. Phrasal procedure	phrasal info, exchange
2. Category procedure	lexical morphemes
1. Word/lemma access	'words'

Table 2.3: Processing procedures and structural outcomes (Pienemann 1998: 9)

In his Teachability Hypothesis (the second part of Processability Theory), language development is implicational. Acquisition of a higher procedural stage is contingent upon the acquisition of lower procedural stages, which cannot be skipped as a result of formal instruction. Instruction is most effective when it emphasises structures from 'the next stage' (ibid.: 250).

Although the Teachability Hypothesis has received some support, it has also been challenged by other researchers. One early challenge comes from Liu (1991), who conducted a longitudinal study of a five-year-old Chinese boy, named Bob, on multiple occasions, involving interactions with peers, teachers, and the investigator. Liu reported that Bob acquired English questions at Stages 4 and 5 before acquiring those at Stage 3. Tarone (2007) regarded this phenomenon as startling counter-evidence to Pienemann's hypothesis, arguing that the social setting was a major factor affecting Bob's order of acquisition. Following this, Spada and Lightbown (1999), in their study of one hundred and fifty Canadian francophone ESL students, found that learners at Stage 2 demonstrated some knowledge of English questions typical of Stages 4 and 5 during the post-test and delayed post-test. A more recent challenge comes from Bonilla's (2015) work. In one of

his studies, he randomly assigned beginner learners to higher Stages (Stage 4 and Stage 5). The learners at Stage 4 received instruction on Stage 4 grammar, and similarly, the learners at Stage 5 received instruction on Stage 5 grammar. The findings reveal that instruction targeting higher Stages enables learners to produce target verbal morphology beyond the immediate next Stages, thus questioning the validity of PT. For example, learners who are at Stage 2 are able to produce the target grammar of Stage 4, skipping Stage 3. Moving on, the next project on uninstructed adult immigrants was carried out under the European Science Foundation (ESF).

2.4.3 European Science Foundation (ESF)

The ESF conducted a large-scale study (W. Klein and Perdue 1997) in five different European countries: France, Germany, Sweden, the Netherlands, and the UK. It focused on naturalistic adult learners of English as an L2. Data were collected from 40 learners with Arabic, Finnish, Italian, Spanish, and Turkish backgrounds. The data collection was longitudinal, spanning a period of 30 months, during which learners were tested at regular intervals using various methods, such as personal narratives, depicting a picture, or retelling a story from a short clip (namely, C. Chaplin's Modern Times). Different tasks were used to collect various types of data. For example, temporality analysis was primarily focused on personal narratives, utterance structure on retelling a story, and spatiality on picture descriptions. The project—and the subsequent intensive research effort—was motivated by a function-to-form approach, which sought to investigate how notional categories or functions (e.g., time, space, etc.) were encoded by learners as they progressed along the acquisition path. The most long-lasting outcome of research conducted in this theoretical context was undoubtedly the development of the Basic Variety theory by proposed W. Klein and Perdue (ibid.).

2.4.3.1 The Basic Variety theory

W. Klein and Perdue regard the Basic Variety as being comparable to Universal Grammar in the sense of Chomsky's Minimalist Program, with the caveat that the Basic Variety allows for only weak (or unmarked) formal features. They propose three developmental stages: pre-Basic Variety, Basic Variety, and post-Basic Variety. In the pre-Basic Variety Stage, the utterances are relatively simple, mainly consisting of nouns (and occasionally

adverbs and adjectives) but lack verbs. The second stage—the Basic Variety—introduces verbs, but they always appear in their base form as bare verbs, indicating the absence of inflectional morphology. According to W. Klein and Perdue, this challenges the assumption that the acquisition process is incremental and developmental, as it shows a developmental discontinuity at the earliest Stages and combines syntactic and non-syntactic principles. Similarly to W. Klein and Perdue's earlier works, BV can capture the initial states of interlanguage but does not account for the SVO word order.

2.4.3.2 Constraints on the structure of the Basic Variety

The Basic Variety is largely characterised by the interplay of three types of constraints: phrasal, semantic, and pragmatic. The phrasal constraints allow for only three main utterance structures, as shown in Figure 2.1.

$$
\begin{array}{ll}
\text{PH1a} & \text{NP}_1 - \text{V} \\
\text{PH1b} & \text{NP}_1 - \text{V} - \text{NP}_2 \\
\text{PH1c} & \text{NP}_1 - \text{V} - \text{NP}_2 - \text{NP}_2 \\
\text{PH2} & \text{NP}_1 - \text{Cop} - \left\{ \begin{array}{c} \text{Adj} \\ \text{NP}_2 \\ \text{PP} \end{array} \right\} \\
\text{PH3} & \left\{ \begin{array}{c} \text{V} \\ \\ \text{Cop} \end{array} \right\} - \text{NP}_2
\end{array}
$$

Figure 2.1: The Basic Variety

In these structures, NP_2 must be lexical, while NP_1 may be represented by a personal pronoun or an empty element. All three structures can be preceded or followed by time or space adverbials. As previously mentioned, verbs are always non-finite (i.e., not inflected).

The key semantic constraint suggested for the Basic Variety relates to the idea of control. For verbs that are linked with more than one 'actant' (argument),

> a semantic asymmetry is observed in that one actant has a higher, and the other(s) a lower degree of control over the situation ... This asymmetry is a continuum ranging from clear "agent-patient" relations down to cases of real or intended possession

The suggested semantic constraint on the utterance structure dictates that the argument with the highest control (controller) should be stated first. The personal pronoun

'she' and the proper noun 'Charlie' in the representative examples in (2.2) act as controllers; therefore, they are mentioned first.

(2.2) a. she $_{[\text{controller}]}$ pushin the policeman
b. Charlie $_{[\text{controller}]}$ hittin the head

All learners in the ESF research appeared to acquire Basic Variety, and some subsequently reached a stable stage during which their production remained ungrammaticalised. However, there are some criticisms against the BV. For example, Schwartz (1997) evaluated the assumption of phrasal constraints and argued that the BV ignored L1-based order differences. For instance, Punjabi and Turkish speakers use their L1 word order OV, while Arabic and Italian speakers use their L1 word order VO. These L1-based order differences, as we will see shortly, are adopted in all hypotheses in the generative approach to L2A. Other researchers argue that the BV overlooks the complexity of second language acquisition by demonstrating that other learners progressed toward what is referred to as the post-Basic Variety. The primary sign of this progression was the acquisition of finiteness, indicating the gradual appearance of inflectional morphology, such as tense marking preceding aspect marking and irregular forms preceding regular ones (W. Klein, Dietrich, and Noyau 1995). Parallel developments were observed in the pronoun system, the acquisition of complex sentences such as cleft structures ('is not the man who stole the bread, it is the girl': W. Klein and Perdue 1992), subordination ('they think about one house for live together', p. 322), and the development of means to express more complex topics (Dimroth 2002). W. Klein and Perdue argue for the learner's subjective need to adapt to his or her social environment regardless of the meaning of the forms (e.g., verbal morphemes) they use.

2.4.4 Lexikalishes Lernen (LexLern)

The last project that recruited adult immigrants was LexLern (Clahsen *et al.* 1983). The researchers in this project recruited 17 learners with L1 Korean and Turkish and seven with L1 Spanish who were acquiring German outside the classroom. Data were collected using interview techniques and tasks designed to evoke specific sentence forms, enabling the analysis of subjects, verb placement, and person and number markers (see Vainikka and Young-Scholten 2011). The analysis from this study contributed to shaping the

initial stages of acquisition under the theory of Organic Grammar (Vainikka and Young-Scholten 1994, 1996a,b, 2011). Table 2.4 provides a summary of these influential studies on the naturalistic acquisition of L2 morphosyntax by adult learners. I will revisit this study and VYSA's study later in Section 2.6.3 when presenting the supporting evidence for Organic Grammar.

Study	L1 and L2	Description	Ideas introduced
ZISA 1980s Clahsen (1983)	Spanish, Portuguese and Italian immigrants L2 German	cross-sectional: 45; 2-year longitudinal: 12 learners	L2 development moves in stages; there is debate on whether adults use the same linguistic mechanisms as children.
ESF 1990s	Immigrants learning 5 European L2s Korean	2 $\frac{1}{2}$ yrs: 40 learners	L2 learners start with a 'Basic Variety'; some don't go further L2
LexLern (Lexikalishes Lernen) Clahsen (1990)	Korean and Turkish immigrants L2 German	cross-sectional: 17 learners	L2 learners follow a natural order that is indeed largely independent of their L1 except for at the very start.
VYSA 1990s-2000s	L1 English exchange students L2 German	1 year longitudinal: 3 learners	Educated exchange students not instructed in the L2 follow the same stages as less educated immigrants.

Table 2.4: Influential studies of naturalistic adult learners' acquisition of L2 morphosyntax (Vainikka and Young-Scholten 2011)

Moving on to the following section, which presents the generative approach to language acquisition. This section will first discuss Strong Continuity and the Structure-Building approach, both of which adopt a generative perspective on language acquisition, highlighting how Strong Continuity contrasts with Weak Continuity or the structure-building hypothesis from which Vainikka and Young-Scholten propose Organic Grammar.

2.5 Generative approach to language acquisition

Studies within the framework of generative linguistics reveal no general consensus that learners follow a specific path of acquisition, and these are formulated under two main approaches. The first approach assumes Strong Continuity and argues for the availability of functional projections at the initial state of acquisition, thus dismissing the concept of stages. In contrast, the second approach assumes Weak Continuity, or the Structure-Building model, and explicitly argues for the incremental development of stages and the absence of functional projections at the initial state. Therefore, the following section will

introduce these two approaches.

2.5.1 Strong Continuity in language acquisition

When children begin the process of acquiring their first language, it is easy to note that they omit a substantial portion of the necessary materials from their utterances. Their first single-word utterances provide little direct evidence of syntax, except for the systematic lack of functional elements, despite the high frequency of these elements in the input. However, the first multi-word utterances quickly indicate that children acquire head-complement relations. For instance, Milles reported that the utterance of a young German learner (Meike; age: 1.10) provided evidence that Meike established the non-finite verb (where -INF = infinitive suffix) follows its object complement (i.e., German has an object-verb (OV) in the VP order as in (2.3).

(2.3) a. *Hause gehen.*
home go*-INF
'(I) go home.'
b. *Give doggie.*

Young children often omit subjects and all functional (inflectional) features; here are just two examples. For a long time, researchers have debated how much (if any) syntactic structure can be inferred from a child who speaks only two words, and how the child progresses from combining merely two words to forming complex sentences. This debate concerns the issue of continuity, or the link between early and later grammars, which refers to the hypothesised mechanisms that propel the child from one non-adult grammar or developmental stage to the next.

Since the early 1990s, the Strong Continuity Hypothesis, based on Steven Pinker's 1984 continuity assumption, has been regarded as a potential resolution to this issue. In the 1960s and 1970s, it was commonly believed that the child's initial linguistic system was not syntactic but purely semantic, governing associations between words in the child's utterances (see, e.g., Braine 1963; McNeill 1966; Slobin 1966).

After it was discovered that a child's development demonstrated cognitive discontinuity, research focus shifted towards understanding how children transition from a pre-syntactic to a syntactic system (see Brown 1973; Pinker 1982, 1984; and the chapters in

Wanner and Gleitman 1982). Chomsky's presentations in the 1980s on the specific nature of Homo sapiens' hard-wiring for language facilitated the hypothesis that children's early language is syntactic from the beginning, despite the existing semantic relationships between words produced by children in their early two- and multi-word utterances. In other words, the assumption that the Principles and Parameters of Universal Grammar represent this hard-wiring implies that syntax, together with semantics, constrains development from the outset. Since its inception, advocates of the Strong Continuity Hypothesis have argued that Universal Grammar offers a single, uniform, and syntactic tree shared by all languages (see, e.g., Boser, Lust, Santelmann, and Whitman 1992; Hyams 1992; Poeppel and Wexler 1993; Weissenborn 1990). From this perspective, the CP (Complementiser Phrase) is always projected, incorporating all other potential projections into this tree. Consequently, no further acquisition processes are necessary for acquiring structure since Universal Grammar provides the entire structure. The child merely fills in specific lexical elements in response to input while acquiring the ambient language over time. Extending the tree metaphor, the child adds leaves to the selected branches of the universal tree. This perspective on acquisition assumes that, although children produce many utterances that are not at all adult-like and that these utterances represent the child's pre-adult syntactic system, their immature grammar does not violate the principles of Universal Grammar. As a result, children's subsequent pre-adult grammar, or stages of acquisition, indicate potential grammar. However, the Strong Continuity Hypothesis holds that stages of acquisition reveal little about syntactic structure since all structures (i.e., all functional projections) are believed to be established from the start of acquisition. Therefore, two-word utterances such as (2.3a) and (2.3b), which were discussed above, indicate full syntactic structure under the Strong Continuity approach. However, both content and function words/affixes will be missing where the child has not yet learnt them.

The equivalent of the Strong Continuity hypothesis within the framework of Universal Grammar in second language acquisition is the Full Transfer (of L1) and Full Access (to Universal Grammar) Hypothesis, proposed by Schwartz and Sprouse (1996) and their subsequent work. Similar to Strong Continuity in first language acquisition, proponents of the FT/FA hypothesis assert that the representation of syntax in learners' minds consists of a single syntactic tree, inclusive of all functional projections. At the onset of second

language acquisition, L2 learners transfer their entire L1 syntactic tree. Furthermore, as development progresses, elements from any branch of the tree can be transferred.

The study by Schwartz and Sprouse (1994) on the acquisition of German word order and nominal case by an adult L1 Turkish speaker, Cevdet, offers empirical support for the FT/FA hypothesis. Cevdet was enrolled in a German language classroom for one year, receiving over 30 hours of instruction each week, totaling 26 months. Cevdet's data, displayed in Table 2.5, illustrates that his proficiency in German had developed through three stages by the end of the data collection period.

Pronominal subjects								
	Preverbal				Postverbal			
stage	SVX		XSV		...VS		Total	
1	18	86%	3	14%	0	0%	21	100%
2	109	50%	38	18%	69	32%	216	100%
3	48	41%	2	2%	67	57%	117	100%
Nonpronominal subjects								
	Preverbal				Postverbal			
stage	SVX		XSV		...VS		Total	
1	11	92%	1	8%	0	0%	12	100%
2	93	78%	26	21%	1	1%	120	100%
3	46	75%	7	11%	8	13%	61	100%

Table 2.5: Declarative main clauses with two or more nonverbal constituents

According to Schwartz and Sprouse (1996), Cevdet's initial L2 German SOV (Subject-Object-Verb) order corresponds to his L1 Turkish word order, with interlanguage word order for pronominal and non-pronominal subjects reaching 86% and 92%, respectively. Furthermore, Schwartz and Sprouse (1996) note that subordinate clauses in Cevdet's data are not overtly marked in lexical terms, as exemplified in (2.4a). In the second stage, Cevdet altered his word order, inverting the pronominal subjects, as illustrated in (2.4b). At this stage, Cevdet also began to mark subordinate clauses with lexical morphemes acting as complementisers. In the third stage, there was also an indication of VS (Verb-Subject) inversion for non-pronominal subjects.

(2.4) a. *Der Mann seine Frau* gekußt.
The man his wife kissed
'The man kissed his wife.'
(Schwartz and Sprouse 1994: 335)

b. *dann trinken wir bis neun Uhr*
Then drink we until nine o'clock
'Then we will drink until nine o'clock.'
(Schwartz and Sprouse 1994: 336)

c. *das hat eine andere Frau* gesehen.
That has an other woman seen
'Another woman saw that.'
(ibid.: 338)

Schwartz and Sprouse argued that case-checking mechanisms constrained Cevdet's word order and nominative case. His interlanguage word order cannot be construed by comparing the L2 system with a target-like grammar. Schwartz and Sprouse maintain that interlanguage should be studied separately, as Bley-Vroman (1983) suggested.

To test the validity of the FT/FA hypothesis, Haznedar (1997) performed a longitudinal study of an L1 Turkish child's acquisition of L2 English IP, his optional use of infinitives, and his CP system. More recently, Özçelik (2009) studied the acquisition of quantificational scope by nine L1 English learners of Turkish and by 19 L1 Turkish learners of English.

Prévost and White (2000) questioned whether the lack of inflectional morphology in L2 learners' data reveals an absence of mental representation. They investigated the finiteness of verbs produced by two adult L1 Arabic learners of French, one adult L1 Spanish learner of German, and one adult L1 Portuguese learner of German to determine whether the absence of tense or agreement morphology indicates the lack of a Tense Phrase (TP) or an Agreement Phrase (AgrP). According to their findings, Prévost and White argue that 'L2 learners have unconscious knowledge of the functional projections and features underlying tense and agreement' (p. 103).

To explain L2 learners' inconsistent application of inflectional morphology, several researchers (e.g., Epstein, Flynn, and Martohardjono 1996; Haznedar and Schwartz 1997; Lardiere 1998b) have proposed the Missing Surface Inflection Hypothesis (MSIH) as a subhypothesis of Full Transfer. Under the MSIH, the omission of inflectional morphology and the relevant syntactic projections is attributed to non-linguistic factors, such as learners' slow processing. This strong continuity in acquisition sharply contrasts with the Weak Continuity approach, which has been explored under the Structure-Building approach in

L2 acquisition.

2.5.2 The Structure-Building approach and the Minimal Trees Hypothesis

The Structure-Building approach was initially developed for first language acquisition by several researchers starting in the late 1980s. This approach was studied for English in the UK by Radford (1988), in the USA by Lebeaux (1989) and Vainikka (1993/4), and in Canada by Guilfoyle and Noonan (1992). For German, see Clahsen (1990) and Clahsen and Penke (1992); and for Swedish, see Platzack (1990). Subsequent contributions on various languages also include Clahsen, Eisenbeiss, and Vainikka (1994) and Rizzi (1993/4), Wijnen (1994) on Dutch, and Radford (1995) on English, with modifications. All these researchers concur on a broad approach in which children begin their acquisition with few or no functional projections and gradually develop more (functional) projections. Indeed, cross-linguistic evidence suggests that the Inflection Phrase (IP) is acquired before the CP at the level of 'coarse' functional projections. There is weak evidence for developmental stages preceding IP (i.e., from Verb Phrase (VP) to IP), as shown in Table 2.6, but some stages can be identified (see, e.g., Vainikka 1993/4).

	IP-elements acquired earlier [before or around age 2]	**CP-elements acquired later** [after age 2]
English	tense auxiliary verbs	relative clauses sentential complementation
Polish	tense/aspect	relative clauses complex sentences
Scandinavian	negation	relative pronoun
French	clitic pronouns tense negation	subordinate clauses relative clauses
Hebrew	tense negation agreement	relative clauses causal and temporal linking of clauses
Turkish	verbe inflection	conjunctions
Georgian	agreement inflection	two-clause constructions
Mandarin Chinese	modals aspect marking	topicalisation discourse particles
Japanese	verbal inflections	relative clauses
Kaluli	tense	discourse particles subordination
Sesotho	tense/aspect	relative clauses topicalisation
K'iche' Maya	aspect negation	yes/no question particle topicalisation

Table 2.6: L1 acquisition of IP and CP (Vainikka and Young-Scholten 2010: 105)

The relative scarcity of evidence for early developmental stages is linked to the lack of relevant data collected from children at this stage, typically between one and two years old. Guilfoyle and Noonan (1992) clearly described and advocated for the Structure-Building approach, coining the term 'Structure-Building Hypothesis'. Furthermore, Slobin (1997) summarised production data from 12 languages in Table 2.6. These findings have significant implications for the acquisition of IP-related elements and CP-related elements, suggesting that IP-related elements develop before CP-related elements. While the original data in the table are presented theoretically and encyclopaedically, there is no evidence in these publications to support the reverse sequence of acquisition, i.e., CP before IP.

In terms of L2 acquisition, Vainikka and Young-Scholten (1994, 1996a,b) propose that the initial state is a grammar with early lexical representations derived from the L1, but that the L1 grammar only partially constitutes this initial state. This view is based on the Weak Continuity Hypothesis for L1 acquisition (see, e.g., Clahsen *et al.*

1994; Clahsen and Penke 1992; Vainikka 1993/4), which posits that the initial state of grammar lacks the full complement of functional categories, though these categories are present in the Universal Grammar inventory. Consequently, the initial state of grammar in L2 acquisition consists only of lexical categories and lacks functional categories from both the UG inventory and any other sources, including the L1.

2.6 Organic Grammar

Organic Grammar was initially developed as a Weak Continuity approach, known as the Minimal Trees Hypothesis and the Structure-Building Hypothesis. It strongly advocates for the gradual introduction of functional projections by L2 learners. The term Organic Grammar first appeared in Vainikka and Young-Scholten (2004), followed by a series of publications (see, e.g., Vainikka and Young-Scholten 2005, 2007, 2010, 2011). Vainikka and Young-Scholten define Organic Grammar in these publications as both a syntactic theory and a theory of second-language acquisition. As indicated in Chapter 1, Organic Grammar is based on the analysis of influential studies on adult L2 learners acquiring their target language naturalistically.

The fundamental concept of OG is that children and L2 learners start their acquisition with the core of the sentence, a VP stage that includes the verb and its arguments. They then progress to an intermediate stage, the IP, and eventually to the CP stage. This progression typically involves Structure-Building, a concept that has gained prominence in recent syntactic theory as part of Minimalism, where syntactic trees are constructed from the bottom up. This contrasts with the earliest versions of syntactic theory, where no derivation occurs. The bare VP of Organic Grammar serves as a universal starting point for language acquisition. Vainikka and Young-Scholten (2005, 2011) note that a bare VP stage may only be evident in adult L2 acquisition, as development through this stage is usually more prolonged than in L1 child acquisition. The bare VP is also apparent in cases where L1 development is slow. For example, in situations where language processing or access to its input is compromised, such as in children who are cognitively or linguistically impaired (Tager-Flusberg 1994) or in the oral language acquisition by deaf children (Goldin-Meadow and Mylander 1993; Kegl 1994).

2.6.1 Organic Grammar as a syntactic theory

Since the 1990s, Vainikka and Young-Scholten have conceptualised language structure, particularly grammatical syntax, as a composition of functional and lexical projections within phrase structures. At the core of this architecture is the lexical projection, VP, which is headed by V (Verb). The VP consists of the verb and its complement, with the subject positioned as the specifier. Positioned immediately above the lexical projection are various levels of functional projections, such as IP or CP layers. Each functional projection encapsulates the acquisition of a functional phrase that is headed by a specific functional morpheme. Following Pollock (1989), functional projections at the IP level in Organic Grammar involve the division of functional projections into TP, NegP (Negation Phrase), AgrP, and AspP (Aspect Phrase). At the CP level, *wh*-markers or complementisers of embedded clauses introduce additional functional projections.

Similarly, X-bar theory (Jackendoff 1972, 1977) suggests that above the lexical projection, in this case, VP, are functional projections, one above the other, each matching an overtly marked functional morpheme. Vainikka and Young-Scholten (2010: 25) take up their stance in relation to Government and Binding (GB) and Minimalism:

> The theory of Organic Grammar is strictly speaking, post-1990s and post-Minimalist; however, OG shares perhaps more features with the Principles and Parameters approach and Government and Binding (GB) Theory (Chomsky 1981[a]) than with Minimalism (1990 and beyond).

Vainikka and Young-Scholten align more closely with post-1990 versions of generative grammar than with pre-1990 ones. Their position can be defined in two ways. Firstly, Organic Grammar is consistent with Chomsky's pre-1990 view of language structure, in which Universal Grammar is innate and dedicated to the language faculty in human minds. Organic Syntax, however, rejects a syntactic tree with fully-fledged projection branches for all of the world's languages that lacks the concept of developmental stages. Instead, Organic Syntax presupposes that learners gradually project a complete tree of a particular language via the lexical projection VP. Organic Grammar also incorporates elements of the Principles and Parameters approach, such as the style of movements, including head movement, A-movement, and A-bar movement, as well as X-bar theory. In other words, it rejects the assumption that the movement of a linguistic item is used to check for agreement or tense features; rather, it posits that movement is solely used to fill syntactic slots.

Second, Vainikka and Young-Scholten argue that OG is post-minimalist in nature. As it does with Minimalism, OG values bottom-up structure derivation, economical syntactic architecture design, lexical universals, and computational mechanisms. Organic Syntax is in contrast to the Minimalist Programme, which limits previous language design to lexicon operation through universal computational components such as Merge and Move. Vainikka and Young-Scholten (2011: 1) believe that Minimalism reduces syntax to 'a state of non-existence'. According to Vainikka and Young-Scholten (2011), Minimalism sheds light on the language's design and syntactic structure derivation. However, one major drawback is that it ignores two fundamental phenomena: stages of development in L2 data and the presence of syntactic structures in connection to functional projections or movement, such as long-distance *wh*-movement. Additionally, Minimalism falls short of adequately describing language acquisition. Vainikka and Young-Scholten (2011) propose that, since the lexicon is the source for phonological, morphological, semantic, and syntactic information in Minimalism, learning a language is expected to be seen as acquiring the lexicon's specific information. However, Vainikka and Young-Scholten unequivocally emphasise that the only similarity between Organic Grammar and Minimalism is the structure's bottom-up derivation. In other words, in Minimalism, a lexicon with certain features is selected (Select), combined (Merge) with another lexicon with the same feature, and then repositioned (Move) for feature checking, which eventually enables the configuration of syntactic and lexical head features. This is analogous to the structural construction of functional projections using the X-bar theory.

The assumptions of Organic Syntax are mainly congruent with the Minimal Trees and Structure-Building hypotheses. The primary distinction is the inclusion of the Master Tree, which highlights the uniqueness of each language. Organic Grammar is composed of the following ten assumptions listed in (2.5):

(2.5)
- ***Assumption 1:*** Each language has a Master Tree that includes all possible projections occurring in that language.
- ***Assumption 2:*** All and only those projections occur in the Master Tree for which there is evidence in the language.
- ***Assumption 3:*** Universal Grammar provides the tools for acquiring the Master Tree, based on input.
- ***Assumption 4:*** The Master Tree is acquired from the bottom up.
- ***Assumption 5:*** The Acquisition-Syntax Correspondence: syntax mirrors acquisition.
- ***Assumption 6:*** Actual instantiations of the tree are projected from the bottom up, based on the Master Tree.

Assumption 7: Partial trees may be projected for constructions which do not involve the full Master Tree structure.
Assumption 8: Lexical and functional projections differ in terms of how they are represented in the grammar.
Assumption 9: Cross-categorial generalisations about structure are possible.
Assumption 10: Only as much adjunction is posited as necessary.

The idea of a Master Tree as the defining feature of each language is the central aspect of Organic Grammar. The Master Tree includes all possible projections in a given language, and these projections must be evidenced in that language. As outlined in the list of assumptions, Assumption 5 is a logical consequence of Assumptions 1–4 and describes the close relationship between syntax and language acquisition. Assumptions 6 and 7 describe how the suggested mechanism is used in syntax. Assumption 8 states that the Master Tree does not include functional projections since they are provided by Universal Grammar; rather, the Master Tree includes lexical projections. Furthermore, lexical categories are cross-linguistically comparable. Assumption 9 allows for the possibility of a cross-categorical generalisation between specifier and complement positions. The final assumption is that adjunctions headed by adverbs are posited at the lowest level of the tree in Organic Grammar. The authors note that partial trees may be projected for constructions not involving the Full Master Tree. Organic Grammar holds that it is Universal Grammar that enables the acquisition of the Master Tree based on the L2 input. The acquisition of the Master Tree is a bottom-up process that is paralleled by the suppliance of functional morphemes.

2.6.2 Organic Grammar as a developmental stage theory

Vainikka and Young-Scholten (2005) use Organic Grammar to challenge both the Basic Variety and Processability Theory regarding the earliest stages of development. While they agree with the idea of incremental and developmental stages adopted in the Basic Variety, their main criticism revolves around the discontinuity observed at the earliest words and the mixing of syntactic and non-syntactic principles. They also argue that while the Basic Variety can account for the initial states of interlanguage, it cannot explain L1-based word order differences—an issue also raised by Schwartz (1997), as discussed in Section 2.4.3. According to Vainikka and Young-Scholten, the Basic Variety is derived from data representing three developmental stages: verb-less utterances, utterances with

L1 word order, and then L2 word order. However, they contend that only the third stage is fully represented by the Basic Variety, leaving discontinuities in the earliest stages and L1-based differences unaddressed.

In their scrutiny of Processability Theory, Vainikka and Young-Scholten (2005) note that it uses the same data as the Basic Variety to describe interlanguage developmental stages and how primary L2 input is processed by learners. While Vainikka and Young-Scholten (ibid.) agree with the concept of stage-like development and the incomplete influence of L1 at early stages, as well as the uniformity of L2 development regardless of L1, they highlight the absence of UG operation in Processability Theory.

According to Organic Grammar's ten assumptions, all language learners, whether L1 or L2, develop phrase structure similarly, beginning with lexical projection. For L2 learners, the initial state includes L1 lexical projection or is built on the minimal trees of L1 (i.e., L1-based differences). The tree starts to develop by using principles of innate mechanisms (UG operation) that drive language acquisition. As shown in Table 2.7, Organic Grammar suggests that language acquisition begins with the VP—the core or bottom part of a syntactic tree (Assumption 4)—which initially follows L1 word order and later adapts to L2 word order. This is followed by an intermediate stage (IP, first TP, then AgrP), and finally the CP stage, as a result of the interaction between UG principles and input.

Stage	word order	Verb types	agreement/tense	pronouns	syntax
VP	L1 order, then L2 order	thematic (main) verbs	none	subject, object pronouns absent	none
NegP	resembles the L2 apart from complex syntax	thematic verbs; copula 'is'	none	pronouns forms begin to emerge	Negation; single clauses; formulaic or intonation-based Qs.
TP	resembles the L2 apart from complex syntax	thematic verbs, modals; copula forms beyond 'is'	no agreement; some tense, some aspect, but not productive	more pronoun forms, but they can still be missing	Conjoined clauses. Formulaic wh-Qs; yes/no Qs w/o inversion.
AgrP	resembles the L2 apart from complex syntax	thematic verbs, modals, copula forms beyond 'is'; auxiliaries in all forms and tenses	productive tense, aspect; some agreement, esp. forms of 'be'	pronouns obligatory, 'there' and existential 'it'	Simple subordination; wh-Qs but all Qs may lack inversion
CP	always resembles the L2 thematic	complex tense, aspect forms; passives; range of thematic verb, modal, auxiliary forms	forms usually correct, apart from newly attempted ones	use of 'there' and 'it' beyond stock phrases	Complex subordination. All Qs with inversion.

Table 2.7: Organic Grammar stages for L2 English

Before presenting supporting evidence and counter-evidence for Organic Grammar, as

well as discussing functional projections within Organic Grammar, I will introduce the concept of morphosyntactic triggers in adult SLA, as developed in Vainikka and Young-Scholten's work during the 1990s, to explain why L2 adult immigrant learners persistently struggle to attain target-like native competence.

2.6.2.1 Morphosyntactic triggers in adult SLA

Inextricably linked to the idea of parameters is the notion that specific parameter settings are triggered over the course of language acquisition (see, e.g., Gibson and Wexler 1994; more recently, Sakas and J. D. Fodor 2001). Fodor, on the other hand, advances the idea of a designated trigger, in which parameters specify the type of input that will drive the language learner to choose a particular parameter setting. By and large, it is considered that triggers must be robust in the primary linguistic data. For instance, Clark and I. Roberts (1993), in their model of mathematical learnability, suggest that triggers cannot be rare in the input. Given that functional elements are often thought to be the locus of parametric variation and that they are situated in the functional projections under investigation for syntactic tree growth, it is reasonable to investigate the idea that functional elements drive this development.

To begin with, the question is whether first language learners' triggers also function as triggers for second language learners. In general, existing literature on L2 acquisition shows that, although bound morphemes (e.g., the verbal morphology *-ed*) serve as triggers in L1 acquisition, free morphemes do so in L2 acquisition. For instance, Zobl and Liceras (1994) reviewed the first and second language morpheme order studies on the learning of English conducted in the 1970s to address the differences between L1 and L2 acquisition. In one of the early morpheme order studies, Bailey *et al.* (1974) reported that the order of acquisition for adult L2 learners was comparable to that of L2 children but different from that of L1 children. The order of the morphemes in Table 2.8 is determined by their acquisition within specific functional projections: children first acquire the affixes (i.e., bound morphemes) associated with DP and IP, whereas second language learners first acquire the free morphemes associated with DP and IP, followed by the affixes.

Related functional projection	Morpheme order in L1A	Morpheme order in L2A
Nominal (DP)	1. possessive	1. article
	1./2. article	2. possessive
Verbal (IP)	1. past & 3SG	1. auxiliary
	2. auxiliary	2. past & 3SG

Table 2.8: Relative morpheme order in English acquisition (Vainikka and Young-Scholten 1998, based on Zobl and Liceras 1994)

In their analysis of the morpheme order studies, Zobl and Liceras (ibid.) take a perspective similar to Vainikka and Young-Scholten's (1998), as we will see shortly: In L1 acquisition, functional projections are realised as bound morphemes, whereas in L2 acquisition, they are realised as free morphemes. Zobl and Liceras suggest, however, that morpheme order analyses indicate a further distinction between L1 and L2 acquisition—namely, that nominal functional projections are acquired earlier than verbal functional projections in L1, although this generalisation is not evident in L2 acquisition orders. They maintain this view to demonstrate that functional projections develop gradually during first language acquisition but are transferred from the native language during second language acquisition. This is based on the premise that a DP must be posited prior to an IP being established. However, syntactic theory does not imply such a constraint, as the 'nominal track' and the 'verbal track' are separate concerning functional projections. As a result, the Structure-Building approach may also be applied to L2 acquisition.

Additionally, the morpheme order studies indicate that one morpheme, *-ing*, is acquired relatively early by both L1 and L2 English learners. This might be seen as evidence against Vainikka and Young-Scholten's proposal that bound morphemes do not serve as triggers in the input for L2 acquisition. Vainikka and Young-Scholten's proposal, on the other hand, is based on a theory of structure building from the bottom up, in which elements connected with the VP, whether bound or free morphemes, are assumed to be acquired before functional elements. Assuming that the construction V+ing is a non-finite form (since the *be*-form 'is' is typically proposed in L1 acquisition; see, e.g., Radford 1990) posited in Verb rather than Inflection, the acquisition of *-ing* by L2 learners before the acquisition of other verbal morphemes indicates that the VP projection is available before functional projections.

Based on the reanalysis of the ZISA and LexLern studies, Vainikka and Young-

Scholten (1998) suggest morphosyntactic triggers for various functional projections in L1 and L2 acquisition. Vainikka and Young-Scholten observed that several learners in the ZISA and LexLern studies (both longitudinal and cross-sectional) appeared to be fossilised. They conclude that it is the different status of triggers for second language learners, rather than a lack of access to Universal Grammar, that precludes them from attaining native competence. Vainikka and Young-Scholten suggest that the difference between bound and free morphemes as triggers may stem from phonology. Typically, free morphemes consist of at least a phonological foot, whereas bound morphemes consist of units smaller than a foot. It is widely believed that aspects of L1 phonology are transferred to L2 acquisition. Furthermore, it is generally agreed that adult L2 learners experience persistent difficulties in phonology.

2.6.3 L2 German evidence for Organic Grammar

Vainikka and Young-Scholten (1994) investigated the acquisition of German by 11 Turkish and six Korean immigrants in naturalistic settings to determine whether L2 learners still have direct access to Universal Grammar. It is noteworthy that all three languages are classified as head-final; however, verbs in matrix clauses occupy the second position in German, assumed to be moved (or raised) to that position. Unlike in German, both Korean and Turkish allow empty subjects. In German and Turkish, subject-verb agreement is required, but not in Korean. Data were gathered in a cross-sectional design, primarily through interviews, narratives based on comic strips depicting hand-drawn pictures, researcher actions, and tea-making processes. In their analysis, Vainikka and Young-Scholten adopted two criteria for acquisition: production of 60 percent correct in obligatory contexts, and the suppliance of at least two out of four agreement morphemes as the acquisition of the agreement model.

Vainikka and Young-Scholten argue that learners progress through three stages: an early head-final VP stage, as in Figure 2.2, followed by an intermediate stage (see Figure 2.3), and then to an advanced stage. During the initial state of acquisition, learners' production revealed no evidence of the acquisition of functional projections; VP headedness in learners' data indicates the transfer of L1 Turkish or Korean lexical projections. At the intermediate stage, AgrP as a functional projection and pro-drop becomes optional. At a later stage, the data demonstrated non-pro-drop and verb raising to the head of

AgrP to achieve overtly expressed agreement (see Figure 2.4).

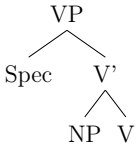


Figure 2.2: VP stage (Vainikka and Young-Scholten 1994: 285)

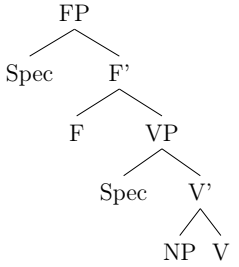


Figure 2.3: FP stage(ibid.: 290)

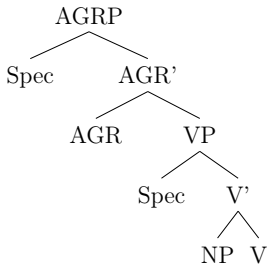


Figure 2.4: AgrP stage (ibid.: 286)

Vainikka and Young-Scholten (1996a) analysed the acquisition of German by four Italian (longitudinal) and seven Spanish (cross-sectional) learners. The Italian data were obtained as part of the ZISA project through fortnightly interviews over a two-year period, while the Spanish data were collected as part of the LexLern project using various research methods, including interviews. The two stages of developmental syntactic structure, VP and IP, are represented in the data of Italian and Spanish learners. Unlike Turkish and Korean learners, Italian and Spanish speakers' VP stages are divided into two sub-stages: the head-initial VP stage (see Figure 2.5), which is transferred from their native languages, and the head-final VP stage (see Figure 2.6), which is triggered by the German input (i.e., learners acquire German VP headedness). Data from the earliest point in time reveal no functional projections; that is, there is no evidence at the suggested Stage I that could demonstrate that the acquisition of auxiliaries or modals indicated the positioning

of projections relevant to agreement, complementisers, wh-questions, and inverted yes-no questions at this early stage.

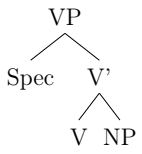


Figure 2.5: Stage 1a (Vainikka and Young-Scholten 1996b: 195)

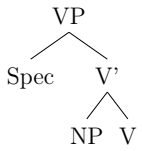


Figure 2.6: Stage 1a (ibid.: 195)

In Stage II (see Figure 2.7), learners exhibit an FP stage, representing an unidentified early stage of the head. The L2 learners examined developed at varying rates. For example, one learner progressed from the IP to the CP level, while another stabilised in the head-final sub-stage. Additionally, it was evident that the language of some learners had fossilised during the data collection period.

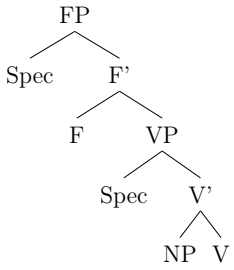


Figure 2.7: Stage II FP stage (ibid.: 195)

Vainikka and Young-Scholten (2011) investigated the acquisition of German by three *ab initio* post-puberty American exchange students during a one-year overseas exchange programme. These learners were enrolled in German schools that did not provide German as a second language, although they offered classes focused on survival skills in German. Vainikka and Young-Scholten used various elicitation techniques to collect oral production data demonstrating learners' use of VP, copulas, modals, verb suffixes, negation, questions, and word order. Based on the learners' data, the researchers found no functional projections during the early stage of acquisition, three weeks after the learn-

ers' arrival in Germany. Instead, L2 German learners begin with a transferred head-first VP and subsequently progress to a head-final VP, similar to their Italian and Spanish counterparts.

Additionally, although all three learners started with a head-initial IP or CP, only two of them, Paul and Joan, succeeded in transitioning to the final position after a year. The other learner, George, projected a head-initial AgrP and a head-initial CP but did not show evidence of acquiring the target German head-final AgrP. According to Vainikka and Young-Scholten (2011: 343), George's production is 'consistent with the L1 English pattern of head-initial projections throughout the grammar'. The data in Table 2.9 illustrate the development of the three participants' functional projections.

Projection acquired	Paul	Joan	George
Bare VP	1-2	1	1
NegP	1-2	1	1-2
FP/TP	3-4	2-3	2-3
AgrP	5	5	4
CP	7-8	6-7	6-8

Table 2.9: Three American English speakers' acquisition of German (ibid.: 342)

Furthermore, it shows that, as Vainikka and Young-Scholten noted, George's data differ from that of the other two learners to some degree, even if the difference is insignificant. Specifically, George supplied NegP-relevant forms at the bare VP stage, but unlike the other two learners, his data showed no indications of projecting the German head-final AgrP.

2.6.4 L2 English evidence for Organic Grammar

Similarly to the application of OG to L2 German, the literature on L2 English reveals three empirical studies supporting the validity of OG in L2 English. The first was conducted by Young-Scholten and Strom (2005), who investigated several aspects of L2 acquisition by 17 adult non-/low-educated immigrants (Somali and Vietnamese learners) acquiring English in natural contexts. These areas include reading skills, phonological awareness, and linguistic competence. In their test battery, the researchers assessed writing and reading skills in both the first language and the second language based on Ehri's (1994) criteria (see Chapter 4 for more details), as well as phonological awareness. In this study, the authors applied OG stages criteria to investigate the correlation between

OG stages and literacy levels. I will return to this when discussing the role of literacy in Section 2.8. To place learners at one of the OG stages, they were required to describe, for five minutes, a photograph that depicted a bystander and a stopped car, with its driver looking down at an unconscious boy lying on the ground next to his bicycle. Accordingly, the researchers were able to elicit a range of constructions. Learners who produced only single-word utterances were assigned to Stage 1. Table 2.10 presents these stages and provides examples of the corresponding constructions produced by the learners.

Stage	**Description and Example**
Stage 1	L1 order/thematic verbs *Example: Car. Bicycle. One boy.*
Stage 2	Thematic verbs; copula *is* *Example: You my car hit here teacher. This is car.*
Stage 3	New functional morphology *Example: The woman is cry. Coordination: Someone's die because he have accident.*
Stage 4	Subordination emerges *Example: Car hit the kid that's lie down on the street.*
Stage 5	Nearly target-like *Example: The young boy was having fun with his bike. He doesn't did that.*

Table 2.10: Stages of development with examples

Mobaraki (2007) addressed Haznedar and Schwartz's (1997) issues raised in a longitudinal study of English morphosyntax acquisition by two Farsi children, Bernard and Melissa, from the initial state onward. At the onset of the study, the children were aged 7;4 and 8;4. Over twenty months, data were elicited through the children's spontaneous speech and a battery of tests, including translation and diaries. Mobaraki investigated the acquisition of word order, negation constructions, the copula and auxiliary 'be', modals, question formation, case marking, null subjects, subject-verb agreement, and tense marking.

The findings show that prior to the seventh data collection session, 93 percent of the sentences had the SOV order, and utterances with the SVO order followed in the subsequent sessions (between the 7th and the 13th data collection sessions). Additionally, the author observes that the headedness switched while the children were still in the VP stage, consistent with the transfer of VP directionality by Italian and Spanish learners studied by Vainikka and Young-Scholten (1994). At the 14th data collection session, the

occurrences of null subjects corresponded with SVO production. Mobaraki maintains that this is consistent with Vainikka and Young-Scholten's claim on L2 German acquisition, as raised verbs are directly related to the absence of subjects. Case was not assigned in the learners' utterances, again indicating the absence of functional projection at the IP layer, consistent with the findings in Vainikka (1993/4), Haegeman (1995), Radford (1995), and Rizzi (1993/4), among others. The decline in null subjects is accompanied by a high proportion of inflectional verbs indicated for finiteness in obligatory contexts.

Mobaraki notes that children produce INFL-related components in the following order: copula be > auxiliary be > modal > verbal inflections (tense and agreement, e.g., *-ed* and *-s*). He adds that the children struggled particularly hard to provide the third-person singular *-s*. In line with Goldschneider and De Keyser (2001), Mobaraki (2007: 230) asserts that 'semantic complexity of this morpheme which stands for number, person, tense, and aspect makes it difficult to acquire [the inflectional morpheme *-s*]', italics added. Although NegP was found in the children's early production, the supplied NegP's headedness violates both Farsi and English's NegP headedness. Mobaraki contends that negation is positioned semantically rather than syntactically, stating 'the early L2 structure is only lexical and [that] the lexical meaning of the verb plays an important role in the syntactic position of the elements' (p. 231). This evidence contradicts Schwartz and Schwartz and Sprouse's (1994) hypothesis that the entire L1 grammar constitutes the initial state of L2 acquisition and Haznedar's (1997) assumption that the NegP headedness is transferred from her participant's L1 Turkish. On the other hand, it supports Vainikka and Young-Scholten's idea that only the lexical category is transferred at the initial state. Furthermore, it is recognised that children's acquisition rates and outcomes vary significantly due to the emphasis on literacy at school and learner-internal differences in verb-less utterances and processing speed.

Kahoul (2014) investigated how adults acquire verbal morphology in L2 English. Specifically, he attempted to determine the source of inflectional morphology (subject-verb agreement and tense) supplied by L1 Chinese and Arabic learners of English at three proficiency levels: low, intermediate, and high. He recruited L1 English native speakers as a control group. His data collection included production and comprehension. Production data included a sentence elicitation imitation task, whereas perception and processing data were collected through a computerised picture-choice task. Eye movement tests

were used to measure reaction time.

The findings from the production data showed that Chinese learners at low, mid, and high proficiency levels and Arabic learners at the lower levels experienced morphological variability in the use of past tense and agreement. Such variability was not observed in the syntactic behavior of Arabic learners at the highest level. The same results were observed in the processing data. Based on these observations, Kahoul maintains that learners progress similarly through a fixed developmental path of acquisition and that persistent morphological variability stems from the absence of syntactic learners' L1s. Such a conclusion raises the question of when L1 influence is possible, as Hawkins (2001) addressed.

Chen (2017) examined the acquisition of morphosyntax by eight adult native English speakers. The study aimed to test the applicability of OG to Mandarin Chinese, a typologically different language. The analysis yielded three main findings. First, when the word order in the VP differs between English and Mandarin, the learners employed the order corresponding to their L1 English, demonstrating L1 word order transfer. Second, a stage-like development pattern emerged based on a hypothesised hierarchical functional projection structure; however, variation was found in the development of individual learners. Third, although most results may be attributed to the interplay between Universal Grammar and target language input, cognitive factors such as memorisation should not be dismissed. Overall, the findings indicate the predicted stages, thus extending the explanatory utility of OG to the syntactic analysis of L2 Mandarin.

2.6.5 Functional projections for English

This section outlines the functional projections proposed for English in Vainikka (1993/4), as cited in Vainikka and Young-Scholten 2011. As demonstrated in that article, most of the word order possibilities for English adverbs are derived from the functional projections required independently for inflectional morphology. Given the productive tense morphology of English verbs (e.g., the regular past tense *-ed*) and the agreement morphology of English (e.g., the third-person singular *-s* and the copula and auxiliary verb forms), Vainikka and Young-Scholten adhere to the long-held assumption (since Pollock 1989) that English has a TP and AgrP.

That is, given the input of English tense and agreement received by the infant, these

two projections are acquired and become part of the adult syntax. Both of these projections are obligatory in English finite clauses; without them, the construction is non-finite. Consider the verb wash in (2.6): washes in (2.6a) is a finite verb, while wash in (2.6b) is a non-finite verb.

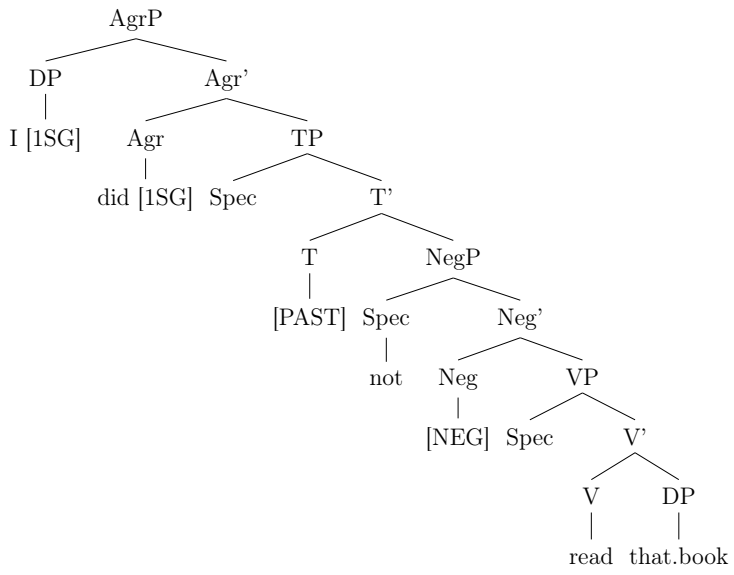
(2.6) a. Mary washes cars for living.

b. When John saw Mary, he asked her to wash his.

In terms of the path of acquisition, Organic Grammar predicts that TP (tense marking) is acquired before AgrP (subject-verb agreement). Vainikka and Young-Scholten (2011) note that this is exactly what the literature on child language acquisition for English suggests (Brown 1973; De Villiers and de Villiers 1973; Ingram 1989).

Along with Laka Mugarza (1990) and Zanuttini (1991), in terms of NegP, the sentential negation and the relevant morphemes (e.g., the sentential negator 'not/n't') are also presumed for English (Vainikka and Young-Scholten 2011). Zanuttini suggests that NegP is dominant over AgrP and TP, while Ouhalla (1993) and Benmamoun (1993) claim that the NegP position or headedness is fixed cross-linguistically. Vainikka and Young-Scholten follow the latter researchers (see, e.g., Haegeman and Gueron 1999), who propose that the English NegP is below AgrP. Whether the NegP dominates the TP or vice versa is not clear, but since the idea that TP dominates NegP fits well with the adverb analysis in Vainikka (1993/4), this will be provisionally assumed here. The example in (2.7) provides three posited functional projections of a simple English sentence.

(2.7) 'I didn't read that book.'

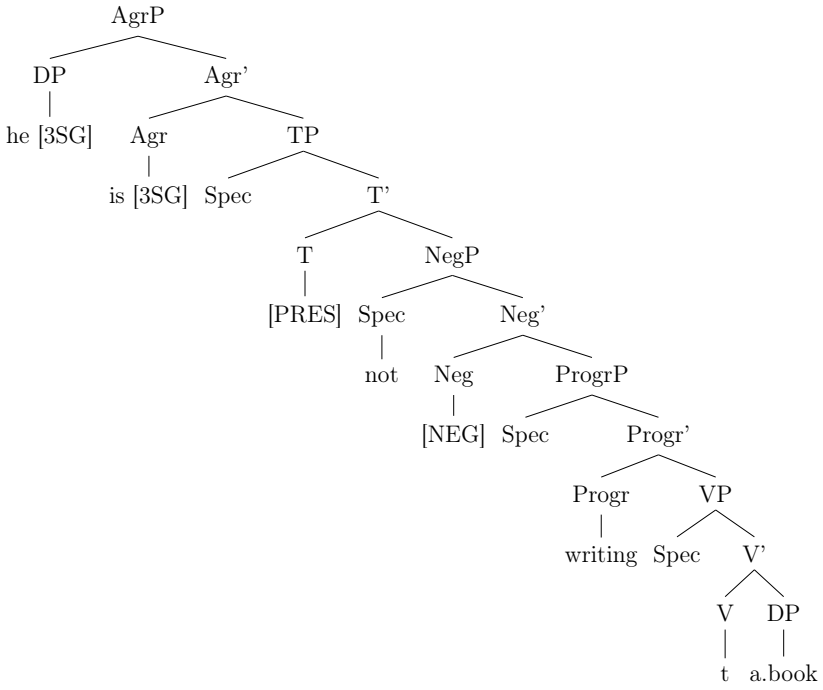


Now, we will look at more complex constructions containing auxiliary verbs other than the auxiliary 'do'. These include auxiliary verbs 'be' and 'have', which form the progressive construction ('be + V-ing') and the compound tense-aspect construction ('have + V-en'). Given the assumptions of Organic Grammar for syntax, it appears that each construction can necessitate two projections. For example, in the progressive construction, one projection is for the auxiliary 'be' and the other for the progressive morpheme *-ing*. However, one projection per construction is appropriate when using an approach to parsing in morphology developed by Brattico (2009, 2010), as applied to verb morphology in Finnish and further extended to English in Vainikka (1993/4).

The idea is similar to the earliest proposals about *do*-support in English, tracing back to Chomsky's (1957) work on *do*-transformation. For example, when a bound morpheme (e.g., the regular past tense *-ed*) in English attempts to attach to a stem that cannot accept a bound morpheme, such as a missing verb or the negator 'not', an auxiliary verb is added (higher up in the tree) to rescue the bound morpheme. Vainikka and Young-Scholten (2011) generalise this idea beyond the auxiliary 'do'.

Consider the progressive suffix *-ing* (as in 'the boy was fishing'). According to the approach detailed in Vainikka (1993/4), the main verb generates all the features (in this case, [+Progr] [+Past] [+3sg]), and one of them, [+Progr], is realised on the main verb (e.g., 'fishing'). The problem is that the verb form 'fishing' cannot have a bound morpheme attached to it. To rescue the bound morpheme, an auxiliary 'is' is inserted in T. The auxiliary 'is' is chosen based on the suffix on the main verb; that is, the auxiliary in some way binds with a given affix or at least co-refers with it. The auxiliary 'be' binds the suffix *-ing*, the auxiliary 'have' binds the past participle *-en*, and the auxiliary 'do' binds the 0-affix on the main verb.

Vainikka and Young-Scholten (2011) anticipate the following order of acquisition of English morphemes: The first verbal morpheme to be acquired is the (progressive) *-ing* without the auxiliary 'be', followed by the (participial) *-en* suffix, again without the auxiliary 'be', and the negative morpheme 'n't/not'. The latest verbal morpheme is the past tense *-ed* (and the auxiliaries 'be', 'have', and 'do' marked for Tense), followed by the third-person singular *-s*, and the agreement forms of the verbs ('be', 'have', and 'do'). This order of acquisition is also predicted by the morpheme order studies discussed above (Bailey *et al.* 1974; Brown 1973; De Villiers and de Villiers 1973; Dulay and Burt 1974b).



As we will shortly see in further detail, Vainikka *et al.* (2017) explored this order in low-educated L2 adult immigrants and found that the route of acquisition is uniform and not affected by the levels of literacy.

2.6.6 Counter-arguments against OG

In addition to supporting evidence for the Full Transfer/Full Access hypothesis, criticisms have been made of the numerous OG assumptions. The following section outlines the crucial evaluations (Epstein *et al.* 1996; Haznedar 1997; Lardiere 1998a; Schwartz and Sprouse 1996; White 2003; Yuan 2004).

White (2003) doubts the Minimal Trees Hypothesis on empirical, methodological, and conceptual grounds. On a conceptual level, White explains why and how functional categories should be missing from the initial state when the L1 grammar remains stable and that Universal Grammar includes an inventory of functional projections. Furthermore, White presents research results and maintains that the Minimal Trees Hypothesis claims are inconsistent with empirical evidence obtained from learners with a variety of L1s, where a wide range of functional categories such as DP, NegP, TP, and CP exist during the initial state. White provided examples from the study of Vainikka and Young-Scholten (1994) in which functional projections appeared to be present. She also argued that the assumption that the absence of functional categories in production data indicates

the absence of abstract grammatical types is questionable. On a methodological level, she contends that designating 60% of a linguistic item's occurrence in required situations as acquired was arbitrary. As Table 2.11 shows, Chen (2017) summarised criticisms other than those of White (2003).

	Opposed claims	Evidence
Schwartz and Sprouse (1994)	No transfer of functional projection of L1 at the initial stage	There was early fronting of finite auxiliary and modal verbs in finite positions and the fronting of non-subjects elements to be assigned with the nominative case in Turkish adult's acquisition of German word order.
Epstein *et al.* (1996)	The syntactic analysis of German; Stage-like development	Experimental tasks and acquisition criteria were not reliable in eliciting and testing L2 syntactic knowledge; the German syntax could be analysed differently; stage-like development was unconvincing, as there was evidence of IP stage production within the VP stage and CP production within the ArgP stage.
Haznedar (1997) and Haznedar (2003)	Morpheme-syntax coupling; VP-IP-CP development; Missing surface morphemes as missing functional projections	CP prior to IP in a turkish-speaking child's acquisition of English and the absence of morphological forms in interlanguage data mirrored the realisation problem of surface morphology other than the deficiency in functional projections.
Lardiere (2000) and Lardiere (1998a)	Implicational development; IP>CP	A chinese-speaking learner of English produced the most difficult types of questions but had difficulties in making past tense.
Yuan (2001, 2004)	L1 VP headless transfer	There was no instance of head-final VP transfer from German into Mandarin; rather, German-speaking learners at different proficiency levels used the head-medial VPs in their negative sentences.

Table 2.11: Counter-arguments against OG (ibid.: 32)

Given the above, as well as Vainikka and Young-Scholten (2011), it is determined that the following data points may be sufficient to reject OG: L1 transfer of functional projection occurs at the initial state, but it also occurs at other stages, which contradicts OG's claims about incremental development stages.

Previous research has suggested that the presence of yes-no questions and *wh*-questions with subject-verb inversion and embedded clauses introduced by syntax-morphology cou-

plings such as complementisers and infinitive clauses at an early stage of L2 acquisition of Germanic languages is a sign of the suppliance of CP functional projections. In other words, each occurrence would indicate that learners developed a full syntactic tree from the outset. The primary motivation for learners to have an entire tree with functional projections at the initial state is that questions are comprehended from the start.

As question formations are beyond the scope of this thesis, it is vital to briefly mention the lack of CP layer of functional projections in empirical studies from the OG perspective, which involves Mobaraki (2007) and Vainikka and Young-Scholten (1994, 1996a,b, 2011).

Vainikka and Young-Scholten (1994) found only occasional production of CP projections in their study of the German acquisition of Korean and Turkish speakers. Six out of seventeen participants produced *wh*-questions and yes/no questions during the AgrP stage (the first two stages are VP and FP). According to Vainikka and Young-Scholten (2011), it is questionable whether the participants have acquired the CP layer of functional projections, as their production demonstrates head-initial projections as opposed to the functional head-final suppliance of native German speakers. Furthermore, Vainikka and Young-Scholten (1996b) note that *wh*-questions and related constructions may indicate 'an emerging CP projection' during the FP stage of German learning by Spanish and Italian speakers. The following example cannot be considered to have a complete CP projection since the subject *wh*-phrase is in its subject position:

(2.8) *wo* *kenn?*
Where meet-*1SG
'Where (did you) meet (him)?'
(Wo hast du ihm kennengelernt?)
(ibid.: 171)

Vainikka and Young-Scholten argue that the lack of an overt subject as in (2.9a) or subject-in-spec VP position, as in (2.9b), in the data collected during the second and third sessions of their study of L2 German acquisition by three young adults who speak L1 English are examples of adjunction of the *wh*-phrase to VP. The AgrP for the three participants surfaced during the fifth data collection session, and their full CP projections emerged around the seventh.

(2.9) a. warum sprechen Deutsch?
Why speak*-INF German
(Warum spricht man Deutsh?)
'Why does one speak German?'

b. Wer sprechen Deutsch?
Who speak German
(Wer spricht Deutsh?)
'Who speaks German?'

(Vainikka and Young-Scholten 2011: 289)

Mobaraki (2007) argues, in support of the developmental stages of OG, that CP projection is absent in the VP and IP stages due to the absence of subject-auxiliary inversion in both yes-no and *wh*-questions. Hence, the evidence from Vainikka and Young-Scholten (1994, 1996a,b) and Mobaraki (2007) indicates an incremental development from IP to CP functional projections.

A further important argument advanced by Vainikka and Young-Scholten (1996a) against those who claimed to have evidence of functional categories in the L2 grammar was that most of the participants in those studies, as well as in their own study, progressed beyond the initial state.

For example, Grondin and White (1996a) gathered data on the transfer of VP-headedness from learners who may have progressed beyond the non-functional projection stage. Furthermore, Vainikka and Young-Scholten (1996a) also argue that Grondin and White's (1996a) assertion was pertinent to early German studies, in which researchers concluded that Turkish L2 German learners started with SVO. Vainikka and Young-Scholten claim that these learners were already relatively advanced and were raising verbs.

2.7 The Modulated Structure Building Approach

A similar approach to OG is Hawkins's (2001) Modulated Structure-Building Approach (MSBA), which is based on the reanalysis of early data by Stauble (1984) regarding the acquisition of English as a second language by learners with different L1 backgrounds, focusing on verbal morphology. Hawkins proposed a predictable order of acquisition for inflectional morphology that supports Structure Building and Organic Grammar.

Initially, a bare VP (first in L1 and then in L2) is projected. The acquisition of verbal-related inflectional morphology then follows a uniform pattern, regardless of the learners' age, type of exposure, and L1 influence. The sequence begins with the copula 'be', followed by the auxiliary 'be', the regular past tense *-ed*, and the third-person singular *-s*. Similarly to OG, MSBA assumes that sentential negation is acquired before inflectional morphology is established. When applying the Split-INFL hypothesis (Pollock 1989) to the data Hawkins discusses, the following general order of acquisition for L2 English emerges: VP⟶ AspP ⟶ TP⟶ AgrP ⟶ CP

Unlike OG with respect to the role of L2 beyond the initial state, MSBA predicts L1 transfer for functional morphology. According to MSBA, although L1 functional projections are not transferred at the initial state, they emerge when L2 functional projections are developed at later stages; yet, MSBA coincides with incremental development based on OG's predictions.

Hawkins adopts a similar approach to that of Vainikka and Young-Scholten's (1998) in terms of triggers. Based on Stauble's (1984) comparison, Hawkins found that beginning and low intermediate English learners, both Spanish and Japanese speakers, used a large number of 'no+zero copula' constructions (e.g., 'he no old') in contexts where English typically has an overt copula, and a high proportion of 'no + thematic verb' constructions (e.g., 'she no write'). Furthermore, Hawkins observed that the acquisition of IP led to an increased use of 'don't' with thematic verbs.[2] Based on this, Hawkins (2001) argues that IP is responsible for the acquisition of the partial 'n't/not' and explains this in the following way:

> On the assumption that Universal Grammar is involved in the construction of L2 mental grammars, learners will have access to the universal selectional properties of I, which optionally selects NegP as its complement. This means that when a copula construction is negated, be will move from the VP to Neg and then to I.

Learners in the early stage of acquisition may treat the sentential negator 'no' as the head of NegP (e.g., 'he no play'), but when the copula 'be' is acquired, it acts as a trigger for IP projection, and this, in turn, drives the restructuring of NegP; the sentential negator 'no' no longer functions as the head of NegP. Otherwise, it would be carried along when 'be' moves from Neg to I, resulting in ungrammatical syntactic structures such as

[2]The idea that L1 has an effect on the word order of 'no+thematic verbs' in early English of Spanish speakers is examined and refuted.

'She no is old'. The acquisition of 'n't/not' structures, located in the specifier position of NegP, leads to the gradual disappearance of the sentential negator 'no' from the learner's mental grammar. The establishment of I also creates a position in which the auxiliary verb 'do' can be inserted when thematic verbs are negated. Consequently, this leads to an increase in the appearance of the construction 'be + not' and a corresponding increase in the use of 'don't'.

To support this assumption, Hawkins observed that data from three intermediate-level learners, Josie, Keiko, and Kumi, show a meaningful relationship between the acquisition of the copula 'be' and the use of 'n't/not'. Along with the development of the pattern 'be + negator', the use of 'n't/not' as a negator has increased, replacing 'no'. Compared to low intermediate learners, there is also a significant increase in the use of unanalysed 'don't' (and even analysed 'don't' for the learner: Kumi). At the advanced level, the participants seemed to acquire target-like properties of negation. The use of 'doesn't' as a sentential negator for tense and agreement increased, while the use of unanalysed 'don't' declined. The development of the 'be + negator' pattern and the use of 'n't/not' as the negator were almost target-like. This is consistent with the statistical analysis performed for this thesis. The increased use of the copula 'be' coincides with the emergence of NegP.

Reviewing the early studies mentioned above, such as the morpheme order studies, and the developmental stages, including the development of negation, question formation, and ZIZA research, Hawkins (2001) concludes that learners' L1, age, educational background, and type of exposure are all independent of these stages. This finding is particularly relevant in the context of immigrant adults with little or no education or literacy, as previously indicated in Chapter 1. Research on the role of literacy in L2 acquisition is still marginal within the field of second language acquisition and has been primarily conducted by LESLLA researchers (Bigelow and Tarone 2004; Mocciaro 2019b; Tarone and Bigelow 2005; Tarone *et al.* 2009; Young-Scholten and Strom 2005). However, the primary focus of LESLLA studies is on the development of literacy skills and corresponding pedagogical challenges, not on the acquisition of L2 linguistic competence (Mocciaro 2019b; Young-Scholten 2013). Young-Scholten (forthcoming) notes that only 6.71% of the 418 presentations at the LESLLA annual symposia from 2005 to 2017 were devoted to language acquisition. Let us now turn to the main focus of this thesis: exploring studies on the role of literacy in the development of morphosyntax.

2.8 Literacy and the acquisition of morphosyntax

Research on non-/low-educated or non-/low-literate learners has been formulated under two distinct approaches. The first is the cognitive approach, which draws from experimental psycholinguistic research on phonological awareness and working memory. Tarone and Bigelow (2005) and Tarone *et al.* (2009) argue that literacy significantly affects L2 morphosyntax. The second is the generative approach, as illustrated in Organic Grammar (Vainikka and Young-Scholten 1994, 1996a,b, 2011), which focuses particularly on the concept of triggers in both L1 and L2 morphosyntax (Vainikka and Young-Scholten 1998).

A key distinction between these approaches is their emphasis on explicit knowledge in adult L2 learning. The cognitive perspective views acquisition as a mostly conscious process that requires explicit attention to linguistic forms, aligning with Schmidt's (1994) Noticing hypothesis. In contrast, the generative approach considers acquisition a subconscious operation driven by Universal Grammar and triggered by simple exposure to input. Before delving into relevant research on the role of literacy in L2 morphosyntax development, it is crucial to underscore the importance of including non-/low-educated learners in SLA research.

2.8.1 The importance of including non-/low-educated learners in SLA research

Although the primary goal of second language research is to identify universal assumptions about how all second-language learners (L2ers) acquire their second language, research focusing on L2ers with no or low literacy skills is scarce. Most SLA research on post-puberty learners focuses on highly educated individuals such as college students or graduates in foreign language classes (Tarone *et al.* 2009; Van de Craats, Kurvers, and Young-Scholten 2006; Young-Scholten 2013). Typically, these learners have high levels of L2 literacy and engage with various academic forms of literacy, including media and digital literacy, providing them additional sources of input unlike non-/low-literate learners.

If a theory is to be empirically robust, it must explain all major findings in the field (M. H. Long 1990). However, if these findings are based predominantly on high-literate groups, then SLA theoretical assumptions might only apply to such groups (Tarone *et al.*

2009). To verify whether our theories extend to non-/low-literate learners, researchers must include these learners in their studies. Currently, there is a significant gap in SLA research, with only a few major findings based on data from non-/low-literate groups (Tarone *et al.* 2009; Young-Scholten 2015). This omission also limits the utility and practical applicability of SLA research. Hill (1970) reported that it was common for uneducated individuals from remote areas to learn L2s in the 1970s. When these adult learners migrated to US towns, ESL teachers seeking guidance from SLA research found little information on illiterate learners because there was no substantial data on their acquisition patterns.

Including illiterate and low-literate L2ers in research would not only help build universal theories of SLA but also allow us to examine how our own cultural presuppositions influence our research perceptions. Investigating an understudied population can significantly broaden our understanding of SLA's nature and human language learning potential (Lantolf 2000; Preston 2000). Moreover, it reveals new issues that have received little attention, such as the use of non-target single/multi words in L2 acquisition (Young-Scholten 2013).

2.8.2 Literacy and the acquisition of morphosyntax under a cognitive approach

As just mentioned, the work of Bigelow *et al.* (2006), Bigelow and Tarone (2004), and Tarone *et al.* (2009) is influenced by the idea that illiterate learners approach language differently from literate learners. Therefore, the following section will first review oral processing and literacy, followed by the role of interactions in second language acquisition (L2A), where learners are assumed to be literate to notice L2 forms. This includes Tarone *et al.*'s (2009) work on the relationship between literacy and the acquisition of morphosyntax.

2.8.2.1 Oral processing and child L1 literacy

Researchers of monolingual language development claim that L1 literacy skills alter how learners think about and process language in all contexts (Olson 2002; Ong 1988). According to Olson, preliterate children think that written signs indicate events and meanings rather than words or sentences describing those events. For example, when asked to write 'two cats', a preliterate child wrote two scribbles; when asked to write 'no cats',

the same child refused to write anything. Olson maintains that preliterate children lack the meta-linguistic concept of words. He observes that a significant discovery made by children is that their own and others' more or less continuous speech can be conceived of as a bundle of lexical units or words. Olson further claims that literacy facilitates meta-linguistic awareness.

> Writing is what introduces our speech to us, revealing our speech as having a particular structure. Children do not know that they speak words, that is, that the flow of speech can be thought of as a string of lexical items. But children in an alphabetic society do come to think about language, minds, and world in terms of the category systems used in writing (p. 164).

One of Olson's (2002) most intriguing claims, in light of the current interest in L2 processing of corrective feedback, is that the ability to correct grammatical errors is inextricably linked to literacy skills. Literacy standards influence our concepts of prescriptive grammar, which we employ to correct ourselves and others. 'With writing, editing becomes inevitable' (p. 162).

There is evidence from L1 acquisition research that children must first become literate before acquiring the more complex syntactic patterns associated with written language. Ravid and Tolchinsky (2002) coined the term 'linguistic literacy' in reference to this more nuanced kind of native language competency. Ravid and Tolchinsky define linguistic literacy as a component of language knowledge characterised by the availability of different linguistic resources, as well as the ability to consciously access one's own linguistic knowledge to view language from different angles. The critical key to this ability is the realisation that written language is fundamentally different from spoken language, and the perception and growing command of the representational system used in the written modality. Native language syntactic patterns developed through the acquisition of literacy require high-register Latinate vocabulary and connectors, heavy generic noun phrases, passivisation for distancing reasons, non-finite subordination, and complex auxiliary constructions. Although native-speaking children acquire simpler syntactic structures before becoming literate, scholars claim that a set of complex structures may not be developed until several years of literacy experience in one's native language. Of course, it is unclear whether this late acquisition of more complex syntax stems from simple maturation or is linked to the acquisition of literacy in native language acquisition. Ravid and Tolchinsky make a compelling argument that literacy is the source of this complex acquisition. Research on older children and adolescents also suggests that 'simple maturation' does

not always result in sophisticated syntactic structures (see, e.g., Karmiloff-Smith 1986; Romaine 1984; Yule 1997).

2.8.2.2 Oral processing and non-literate adult learners

Experimental studies in cognitive psychology and neuropsychology with adults corroborate the transformative effect of literacy (e.g., see Adrian, Alegría, and Morais 1995; Loureiro, Braga, Souza, Filho, Queiroz, and & Dellatolas 2004; Morais, Bertelson, Cary, and Alegría 1986; Read, Yun-Fei, Hong-Yin, and Bao-Qing 1986). These studies reveal that non-literate adults consistently perform differently from literate adults on oral cognitive tasks that require knowledge of linguistic units.

These studies employed oral cognitive processing tasks to evaluate skills such as phonetic discrimination (e.g., '[ta-sa]: same or different?'), rhyme awareness (e.g., 'mepu/pepu: rhyme or not?'), syllable detection (e.g., 'Is the syllable [pa] included in the word [pati]? For example, what do you have if you omit [pa] from [pati]?'), phoneme detection (e.g., 'Do these words begin with the same sound? [kar], [kus]'), syllable deletion (e.g., 'If you remove [de] from the word [kade], what do you have?'), phoneme deletion (e.g., 'If you remove [t] from the syllable [tal], what do you have?'), word reversal (e.g., 'Say zanahoria rota backwards'), and syllable reversal (e.g., 'Say [taro] backwards').

While non-literate participants perform equally well on tasks such as rhyming and phonetic discrimination, they perform significantly worse than literate groups on all tasks requiring conscious awareness of individual phonemes (i.e., matching, monitoring, deletion, and reversal of phonemes) and on some tasks requiring syllable manipulation.

According to Reis and Castro-Caldas (1997), non-literate participants employ only semantic strategies and do not engage in phonological analysis. In contrast, individuals with alphabetic literacy skills can utilise both semantic and phonological strategies, significantly improving their short-term memory and accuracy in these tasks. Reis and Castro-Caldas also found that literate individuals develop a mechanism to link visual-graphic meaning to units smaller than words, which lack semantic meaning. This strategy enables them to systematically incorporate these units into a working memory system through visual experience.

> Learning to match graphemes and phonemes is learning an operation in which units of auditory verbal information heard in temporal sequence are matched to units of visual-verbal information which is spatially arranged. ... Learning to read and write introduces into the system qualitatively new strategies for dealing with oral language; that is, conscious phonological processing, visual formal lexical representation, and all the associations that these strategies allow (Reis and Castro-Caldas 1997: 445)

These findings imply that alphabetic literacy improves awareness of speech segmentation in oral language processing and has implications for research in the field of SLA. For instance, Schmidt's (1990) hypothesis asserts that an L2 learner must consciously notice the linguistic forms of L2 to acquire them. Tarone *et al.* (2009) wondered if only literate learners can acquire L2s in this way. The research in cognitive psychology cited earlier suggests that L2 learners who are not literate in an alphabetic script may not be as capable of mentally manipulating the L2 in terms of phonemes, syllables, and possibly words. However, it is known that they do successfully acquire L2s (Bigelow and Tarone 2004).

2.8.2.3 Interaction and negative evidence in second language acquisition

To date, SLA research indicates that the processes used by educated, literate learners when interacting in another language may boost language acquisition. For instance, interaction allows learners to test hypotheses about the L2, receive negative and positive evidence, identify syntactic or lexical gaps in the L2, and negotiate both form and meaning with native or nonnative speakers (e.g., Chun, Day, Chenoweth, and Luppescu 1982). Researchers and practitioners are particularly interested in how learners exploit corrective feedback, such as recasts, in oral interaction (M. H. Long, Inagaki, and Ortega 1998; Lyster 1998; Lyster and Ranta 1997; Mackey and Philp 1998; Ortega and M. H. Long 1997). Recasts provide an immediate correct reformulation of a learner's incorrect grammar and might redirect learners' attention to non-target components of their utterances while they focus on meaning. This may be a highly useful strategy to help learners make form-meaning associations that facilitate their acquisition of L2. Nonetheless, a learner may not notice or might misunderstand the purpose of a recast (Lyster 1998; Mackey, Gass, and McDonough 2000). Misunderstandings may arise due to their proficiency levels, meta-linguistic skills, or recast qualities (e.g., the length of the recast or the number of lexical or syntactic changes made). The role of recasts in SLA has been extensively researched with literate learners, the majority of whom are adults, but not with illiterate

learners in any alphabetic scripts who may lack the meta-linguistic skills necessary to utilise recasts effectively.

As a result, Tarone *et al.* (2009) took a step to include adults who are low- or non-literate in the growing body of research on the effect of corrective feedback on second language acquisition (SLA). Tarone *et al.*'s point of departure was to partially replicate Philp's (2003) study, examining the effect of recasts on L2 question formation. The authors used various tasks to collect data, including corrective feedback, elicited imitation, and oral narratives. In the oral narrative task, the author aimed to investigate morphosyntactic elements that learners produce while focusing on meaning, which may therefore be deemed acquired. In this part, Tarone *et al.* aimed to examine the production of bare verbs and nouns (verbs and nouns that lack inflection).

The results of this research show that although both groups used bare verbs, those with moderate alphabetic print literacy produced on average more inflected verbs than those with lower literacy levels. The group with low literacy levels used bare verbs 64 percent of the time (205/321), while the group with moderate literacy levels used them 50 percent of the time (230/458). Individuals in the two groups varied in their production of bare verbs. The low-literacy group produced bare verbs at a rate of 54 percent (50/93) to 77 percent (41/53), whereas the moderate-literacy group produced bare verbs at a rate of 58 percent (16/106) to 38 percent (45/119). Both groups supplied a sufficient number of verbs in the storytelling, though the group with higher alphabetic literacy supplied more: 458 compared to 321 for the low-literacy group. Given the great degree of variability in their findings, the association between alphabetic print literacy and verbal morphology use needs, as the authors suggest, further thorough investigation. According to Tarone *et al.*, this distribution may indeed be connected to the 'simplifying' vs. 'standard' orientation first reported in the ZISA data by Clahsen *et al.* (1983).

As mentioned in previous research, Tarone *et al.* based their study on the belief that the difficulties of oral processing in child L1 literacy and adult illiteracy are related to alphabetic literacy. They linked variabilities in producing target morphology between literate and non-literate learners to their oral processing abilities, which in turn are associated with a lack of literacy. However, they observed that, despite the absence of phonological awareness, the learners were able to memorise and recite lengthy narratives in Somali. This suggests that their working memory relied on something other than

phonemic segmentation, possibly prosodic units such as rhythm, which are unaffected by literacy.

Reviewing the ZISA research presented in Section 2.4, Tarone *et al.* (2009) highlight the lack of an objective measure of literacy levels, raising questions about the connection between literacy and learners' variation in acquiring grammatical morphemes. Although ZISA collected self-reported educational levels, Ksmids *et al.* (2004) found that these levels do not influence phonological awareness in the same way as literacy does. Notably, individuals can be less educated yet literate, or have multiple years of schooling but still remain illiterate. Understanding whether literacy levels correlate with success in producing or omitting grammatical morphemes is crucial. Additionally, in relation to the developmental stages described in Processability Theory, also discussed in Section 2.4, which suggests a universal developmental sequence, Tarone *et al.* suggested that the universality assumption be specifically tested on L2 learners with low or no literacy levels.

Before moving on to discuss the role of literacy from the perspective of the generative approach in L2 acquisition, it is important to address a body of research that focuses on literacy and L1 acquisition. Notably, this research does not adopt a generative approach. Therefore, the following section will present this research, offering insights into how literacy could influence the development of L1 acquisition.

2.8.3 Literacy and L1 acquisition

The potential relationship between literacy and L1 acquisition appears in Dąbrowska (2012). In her article, Dąbrowska refutes the belief that all speakers converge towards the same grammar while acquiring their first language. She considers this belief a myth of second language acquisition research, particularly within generative studies, which will be discussed below. Instead, variations exist depending on the type of input to which learners are exposed and, consequently, the type of education received (low-educated vs. high-educated groups). She contends that this variation can result in different grammars, i.e., different linguistic competencies as opposed to linguistic performances.

Dąbrowska examined various studies on adult L1 speakers, as well as one study including L2 learners. One of these studies is Dąbrowska and Street (2006), in which researchers examined the production of various morphological and syntactic phenomena, including the Polish genitive singular masculine inflection, the dative inflection, and cer-

tain complex English constructions and passive constructions. The statistical analysis of production data for such phenomena revealed a positive relationship with the type of education. Dąbrowska claims that the effect of education is most likely due to asymmetries in the proportion of dative nouns in spoken and written language and differences in exposure to written material. The dative case is employed in spoken texts to denote semantic functions (e.g., recipient, beneficiary, addressee, experiencer). These uses, nonetheless, are very literary, if not archaic, in nature and hence mostly limited to written materials. As a result, the dative appears in a considerably wider variety of constructions in written language, and hence with a much wider variety of nouns. Because more educated speakers are more familiar with formal written language, they encounter a wider variety of noun types in the dative; and because increased productivity results from increased exposure to a greater variety of noun types with a given inflection, more educated speakers are more likely to use the dative inflections productively.

According to Dąbrowska, the relationship between education level and differences in grammatical competence arises from the more extensive experience that educated learners have with written material. In some cases, such as the study described by Dąbrowska and Street (2006), this may result from a combination of written input and exposure to specific language constructions in classroom settings. This study analysed L1 speakers' and L2 learners' comprehension of active and passive constructions. The participants were divided into four groups: highly educated L1 speakers (British postgraduate students), less-educated L1 speakers (with no more than secondary school experience), and highly and less-educated non-native speakers. Dąbrowska and Street observed that although both highly educated groups performed better on average, non-native non-graduates excelled in specific tasks compared to less-educated native speakers. This suggests that low education and exposure interact with other variables: Dąbrowska and Street argues that less-educated speakers' difficulty with passive constructions stems from their lack of experience with these structures, as passive sentences are rare in spoken language, resulting in an underdeveloped passive schema. However, exposure alone cannot fully explain the observed differences, as it does not account for why less-educated non-native speakers performed better than less-educated native speakers, who presumably had greater exposure to passive constructions. Other factors, such as the type of experience (e.g., explicit instruction, exposure to a relatively large number of exemplars over a relatively short

period), also play a role. Additionally, meta-linguistic awareness likely enhances the acquisition of the non-native group through explicit L2 instruction, alongside variations in motivation and ability (Dąbrowska 2012: 248).

This line of thought suggests that the type of input—including what learners might receive in the classroom and through interactions with written materials—can significantly influence the frequency of particular language phenomena (see similar observations in Cintrón-Valentín and N. C. Ellis 2016). Mocciaro (2019b) argues that it is essential to isolate factors to prevent overlaps in the roles of input, level of education, and explicit instruction. Although these factors interplay in language acquisition, they must be analysed independently.

Mishra, Singh, Pandey, and Huettig (2012) conducted a study on university students and low-educated adults who were all native speakers of Hindi in Uttar Pradesh. They measured eye gaze in response to pictures depicting oral declaratives with gender-marked adjectives and particles, where participants listened to a simple sentence that included the target word (e.g., 'door') and simultaneously looked at four objects (the target word 'door' plus three pictures used as distractors). The results showed that high-literate learners started to look more at the target object than at unrelated distractors. The following section presents Tarone *et al.*'s (2009) work on the approach of interaction in second language acquisition, which includes low-educated L2 learners acquiring English as an L2.

This line of research (either in L1 or L2 acquisition) has important implications within the generative paradigm, where cognitive factors, such as literacy, are considered irrelevant to the acquisition of morphosyntax. Vainikka and Young-Scholten (2007) took a step in investigating literacy under their Organic Grammar framework, which adopts a generative approach. Therefore, the next section will first present the stance of the generative approach versus non-linguistic or cognitive factors, and then discuss the role of literacy by addressing the acquisition of morphosyntax by low-literate learners within the generative approach.

2.8.4 The generative approach and cognitive factors

The role of literacy in the development of L2 morphosyntax has long been neglected by those who work in the generative paradigm or by those who work in non-generative

paradigms, such as usage-based approaches. This is because those who work within the generative paradigm assume modularity of mind in both L1 and L2 (e.g., Schwartz 1993 in how the modularity of mind assumption operates in adult L2 acquisition, where language-specific mechanisms remain active). This is the conventional view of how children acquire linguistic competence in their first language, as well as the stance advocated by L2 scholars, who maintain that language-specific mechanisms underlie L2A in both children and adults. A fundamental tenet of this paradigm is the modularity of the human mind (J. A. Fodor 1983), the prominent feature of which is the assignment of unique psychological modules to certain cognitive abilities. These are intrinsically defined, task-specific devices for processing specific types of input. J. A. Fodor makes a seminal contribution to this debate by proposing a modularised mental architecture. He proposes modules for each of the five sensory modalities and a core module for language (known as the language faculty). According to J. A. Fodor, a module is informationally encapsulated and impenetrable by other cognitive domains (Pylyshyn 1980). It only deals with input from a specific domain, which it maps into a symbolic (syntactic) form that passes to a central processing component. The latter combines input from multiple modules and infers on the symbolic structures it receives. Although a module cannot access information outside of its domain, the central system uses sophisticated learning and logic procedures to derive content from a number of sources.

Children acquire their linguistic competence—most particularly syntactic competence—independently of general cognition. This perspective is supported by several sources, including the observation that typical children acquire an intricate system whose complexity cannot be fully explained by the input alone; this phenomenon is referred to as the Logical Problem of Language Acquisition (Hornstein and Lightfoot 1981). A related issue, often associated with the Poverty of the Stimulus, is the absence of negative evidence. In the case of a child's first language, minimal exposure to positive evidence (random samples of language encountered in daily life) appears to be sufficient for them to successfully acquire their native language. Young children do not receive explicit instruction on the properties of their native language. Although parents/caregivers, in principle, do not correct their children when they produce non-target-like utterances, studies show that parents usually do not correct them, or if they do, it has little effect on their linguistic development (for a summary of key findings, see, e.g., Guasti 2002).

Data from children with cognitive deficits who nonetheless normally develop language provide additional evidence supporting the generative approach (see, e.g., Bartke and Siegmüller 2004; Bishop 2001). For example, Bellugi, Van Hoek, Lillo-Martin, and O'Grady (1993) conducted a study on children with Williams syndrome, a rare metabolic disorder associated with mental impairment and other physical issues. They demonstrated that individuals with extremely slow overall mental development, who remain cognitively below the level of a seven-year-old, can exhibit sophisticated knowledge of syntax and an adult-like vocabulary. Smith and Tsimpli (1995; see also Smith, Tsimpli, Morgan, and Woll 2010) examined the extraordinary case of Christopher, a brain-damaged man hospitalised due to his inability to care for himself. Despite these challenges, he is able to read, write, and speak in a variety of languages. The most striking feature is a significant mismatch between his verbal and non-verbal abilities, supported by long-term test results and various other examinations. The broadest generalisation is that he has a low performance IQ combined with an average or even above-average verbal IQ (Smith and Tsimpli 1995). Conversely, children who are cognitively 'normal' may have severe language impairments. Developmental Language Disorder (DLD) is defined as 'persistent difficulties with the productive rules of word-formation, the morphosyntactic prerequisites of feature agreement, and the construction of complex phonological units' (Lorenzo and Longa 2003: 645). A study of an English-speaking family revealed that 16 of 30 members in the last three generations suffered from SLI, suggesting that it is an inherited disorder and that at least some aspects of language may be determined by genetics (Gopnik and Crago 1991). The connection I just sketched between the brain and language is oversimplified (for a more comprehensive discussion, see, e.g., Friederici 2017; Jenkins 2000; Lorenzo and Longa 2003). According to these data, specific parts of the brain are associated with particular aspects of language. Recent developments in brain imaging methods have also strengthened this notion. However, the picture becomes more complicated as the tools become more advanced (Carter 1998; Dornyei 2009; Paradis 2004). Universalists have used this evidence to argue that there must be an innate language faculty that is physiologically activated to explain why language develops in children.

Schwartz (1993) employed J. A. Fodor's (1983) criteria for a mental module to explain how the modularity of mind assumption operates in adult L2A when language-specific

mechanisms continue to function and the learner maintains access to UG (see, e.g., White 2003). Schwartz added that, similar to vision (another module believed to have specific rather than general mechanisms), the input relevant to the computation of knowledge within a given module is constrained. The only meaningful input for the language module is utterances in the learner's surroundings (i.e., primarily linguistic data). In general, instructed second language learners develop an additional type of knowledge—'learned linguistic knowledge' (LLK), comparable to Krashen's (1985) 'learned system'—which develops through the use of general cognitive mechanisms in response to explicit instruction and error corrections in classroom settings. However, identifying the source of utterances in L2 learners is challenging. LLK as a source may be disregarded for young children, as they possess minimal meta-linguistic knowledge and their meta-linguistic awareness is basic and not under volitional control (see Gombert 1992). For most adult learners, however, LLK is likely the primary source of their linguistic behaviour—the language they produce. Jordens (1996), on the other hand, noted that while adult learners may use general cognitive mechanisms to develop the meta-linguistic knowledge necessary for L2 utterances, this does not imply that they acquire linguistic competence in L2A.

R. Ellis (1990) reviewed a body of empirical evidence supporting the modularity of the mind and Jordens' observations, concluding that instruction has no effect on learners' developmental routes. As a result, language-specific mechanisms must operate to some extent, regardless of the context. R. Ellis further noted that while instruction can influence the rate of development—supported by previous research showing that classroom learners progress faster and achieve more than uninstructed learners—it remains unclear why or how instruction affects the rate and ultimate attainment but not the developmental route. It is often suggested that this function is linked to the impact of meta-linguistic knowledge on language acquisition. However, it may be argued that many studies exploring the influence of instructional elements do not significantly challenge modularity, as instruction is viewed as simply providing more primary linguistic data. Despite numerous studies, evidence showing a direct impact of learned linguistic knowledge on linguistic competence remains limited.

Let us now examine two (near) facts about second language acquisition. The first is the idea that a learner's first language (L1) influences their second language (L2) development, and the second is that L2 acquisition exhibits inter- and intra-learner variations.

Both of R. Ellis's conclusions regarding the effects of instruction and the modularity of mind/Universal Grammar access in adult L2 acquisition accommodate variations in terms of rate and ultimate attainment or end state, but not route.

The acquisition of morphology is central to the debate between those who believe that second language acquisition is driven by language-specific mechanisms and those who argue that general cognitive mechanisms drive language acquisition. Morphology also plays a key role in the ongoing debate about whether all second language acquisition, regardless of the age at first exposure, is driven by language-specific mechanisms. In the context of acquiring verbal inflectional morphology, for instance, there is considerable variation among learners, which fuels this debate. If it is assumed that general cognitive mechanisms govern second language acquisition beyond puberty, then there is no reason to dismiss the significance of non-linguistic factors such as literacy in the development of morphosyntax. As indicated above, Vainikka and Young-Scholten (2007) investigated the role of literacy (both the rate and the route of acquisition) in the development of morphosyntax from an Organic Grammar perspective. As discussed previously, studies of morpheme order, L2 negation, and question formation are typical examples of systematic developmental studies that have been undertaken in various target languages since the 1970s. Research from the 1980s onward (e.g., Clahsen 1983; Vainikka and Young-Scholten 1994, 1996a,b) supports the existence of systematic developmental stages. This evidence has often been interpreted to indicate that learners follow a consistent path in at least certain aspects of L2 syntax, even though the pace at which they progress varies.

2.8.4.1 Acquisition of morphosyntax by low-literate learners under the generative approach

Vainikka and Young-Scholten (2007) argue that the relationship between phonological awareness and second language acquisition cannot be taken for granted, contrary to the claims made by Tarone *et al.* (2009). Instead, they suggest that variation in learners' morphosyntax is linked to the difficulty of L2 phonology. As discussed in Section 2.6.2.1, this idea has been further developed within their theory of Organic Grammar. Detailed in Section 2.6, Organic Grammar posits that, similar to first language learners, second language learners gradually acquire syntactic structures. Initially, they primarily posit lexical projections and then incrementally add the relevant functional projections. All

learners, whether L1 or L2, are expected to acquire the language successfully. However, L2 research reveals that adult L2 learners' acquisition takes longer than that of child L1 learners. To explain this variation, Vainikka and Young-Scholten (1998) suggest that L1 and L2 acquisitions differ in terms of the triggers that facilitate acquisition. In reviewing Young-Scholten and Strom's (2005) small-scale study presented in Section 2.6.4, Vainikka and Young-Scholten (2007) link the slow progress of non-literate learners to the concept of triggers, as Young-Scholten and Strom's findings showed that phonological awareness increases when the learner becomes literate, but word awareness and prosodic awareness appear to be independent of reading skills.

Young-Scholten and Strom's findings support the correlation between low-/non-literacy and morphosyntax development, though they caution against drawing causative conclusions. Other factors, such as phonological competence and language input, which will be discussed in the following section in more detail, might still be at play in influencing morphosyntax development. As shown in Table 2.12, neither the reading level nor the OG stages are linked to the amount of ESL instruction or the length of residency in the USA. Despite varying levels of ESL instruction (ranging from 4 months to 12 years) and differing lengths of residence (between three-quarters of a year and 12 years), six of the eight uneducated (or unschooled) learners were placed at OG stage 1 or 2. Significantly, S3 reached level 4 in English reading (despite not being able to read in Somali) and OG stage 5, with only two weeks of ESL instruction and two years of residence in the USA at the time of testing.

Learner	Sex	Age	NL school	ESL	In USA	Reading level	Syntax stage
V6	f	70	0	1 yr	2 $\frac{1}{2}$ yrs	1	1
S2	f	47	0	2 yrs	5 yrs	1	1
S10	f	66	0	1 $\frac{1}{2}$ yrs	3 yrs	1	2
V1	f	51	0	1 yr	20 yrs	1	2
S9	f	54	0	1 yr	4 yrs	1	2
S8	f	31	0	4 mns	9 yrs	1	2
S4	f	38	0	3 yrs	9 yrs	2	2
V2	f	64	2 yrs	2 yrs	8 yrs	3	1
V5	m	34	1; 4 yrs	$\frac{1}{2}$ yr	$\frac{3}{4}$ yr	3	1
V7	m	53	5 yrs	$\frac{1}{2}$ yr	3 yrs	3	1
V4	f	43	3 yrs	$\frac{1}{2}$ yr	13 yrs	3	2
V3	f	31	3 yrs	4 yrs	12 yrs	4	5
S6	f	24	2 yrs	1 yr	2 yrs	2	2
S5	f	32	2 yrs	1 yr	2 yrs	3	1
S7	f	30	5 yrs	1 $\frac{1}{2}$ yrs	9 yrs	3	3
S3	m	30	0	2 wks	2 yrs	4	5
S1	m	26	4 yrs	0	1 yr	5	5

Table 2.12: morphosyntax level and reading level of L2 adults with little or no schooling (Young-Scholten and Strom 2005: 52)

Investigating the route of acquisition under Organic Grammar, Vainikka *et al.* (2017) explored the developmental route of 14 learners from three linguistic backgrounds (Arabic, Somali, and Urdu) with varying lengths of residence, amounts of native language schooling, and literacy levels. Data were collected orally through picture description tasks to investigate the acquisition of the VP, NegP, TP, and AgrP. Literacy was measured using Ehri's (1994) criteria. The conclusion drawn by Vainikka *et al.* was that (1) the route of acquisition is uniform and compatible with research on L1 and L2 morphosyntax development as presented in Chapter 2, and (2) it is unclear whether the low levels of morphosyntax were due to literacy or the nature of the instruction the learners had received. Therefore, a lack of phonological competence may negatively impact the analysis of L2 sub-foot constituents. The difficulty in distinguishing heads may be exacerbated for less-literate learners because they have less visual reinforcement of linguistic forms (Vainikka *et al.* 2017). The following section will explore the role of input, which is crucial to consider because it is difficult to draw conclusions about the relationship between literacy and morphosyntax without understanding the input these low-literate learners receive, as input is a significant factor in language acquisition.

2.9 The role of input

L2 learners receive language input in classrooms through two approaches: traditional and communicative. This means that some learners may receive less explicit grammar instruction than others. Learners who receive input not in the form of primary linguistic data may have the potential to interact with the operation of linguistic mechanisms differently. The receipt of input for first language acquisition is likely to differ from second language acquisition. This complicates the task of researchers comparing the acquisition of L1 and L2 or child L2 and adult L2 learners. Many researchers collect data from younger and older classroom learners. Although the benefit is quick availability and easy control of some variables, it is unavoidable that some input will not be in the form of primary linguistic data.

In the 1980s, some researchers turned to collecting data from naturalistic adult learners to strengthen the validity of such comparisons, because the type of input in second language acquisition is expected to vary from first language acquisition. However, this poses another issue: when it comes to adults, individuals who do not take advantage of classroom instruction and are available to researchers in post-industrialised countries tend to be low-skilled immigrants with only secondary education at most. Therefore, there are fewer opportunities for these individuals to interact with native speakers of the target language, as they are unlikely to integrate into mainstream society. If these non-classroom learners receive little primary linguistic data, they are more likely to make little progress beyond the earliest stages of language acquisition.

From a variety of perspectives, the input (both quantity and quality) appears to be the reason for these learners' slow acquisition and, perhaps, for the differences typically observed between the development of child L1 learners and younger and older naturalistic learners (see Vainikka and Young-Scholten 2011).

In the generative approach, input remains under-examined. Kempen (1998) called for a closer examination of input characteristics before drawing any conclusions about the causes of differences in the development of syntax between child L1 learners, child L2 learners, and adult L2 learners. However, this call was largely ignored. In this section, I will present a brief but serious discussion of two types of input (both quantity and quality). This may help us understand why some immigrants stabilise at the bare VP stage despite

their length of residence in the UK and the duration of their ESOL instruction.

2.9.1 Input quantity

The estimated amount of active primary linguistic data, or input, that children receive has been long reported. Sharwood Smith (1994) estimates that by the age of five, children will have received around 9,000 hours of input. But is this same amount of input required for L2 learners to converge on the target language? This remains uncertain. Aside from L2 learners who receive input solely in classroom settings, little is known about the total amount of input that L2 learners encounter (see Carroll 1999, 2001). Learners in immersion settings will likely have more frequent interactions with native speakers than those in classroom environments. However, studies examining age of initial exposure and ultimate attainment in immersion contexts often rely on an inaccurate proxy for exposure: length of residence.

2.9.2 Input quality

Adult immigrants who interact mainly with other non-native speakers will receive a substantial amount of non-native pronunciation, ungrammatical forms, or interlanguage input, which results in slower or even different language development (Piske, MacKay, and Flege 2001). One variable that could affect input for L2 adults is literacy. For some low-literate learners, the end state of acquisition is the VP; this appears to be the case for low socio-economic stratum/low-educated immigrant learners. As previously mentioned, low-literate learners are more likely to be found on the fringes of mainstream society. A lack of literacy would result in even less input for the least educated of such learners, as they would not be able to read books, newspapers, magazines, or blogs.

As Van de Craats (2011) and Tarone *et al.* (2009) point out, it is crucial to consider the effect that lack of education and hence illiteracy may have on acquiring a second language. Van de Craats contends that the developmental patterns observed in her LESLLA adult corpus vary from those reported in younger learners and instructed learners. Tarone *et al.* present a detailed argument revolving around differences in educated and unschooled learners in terms of the development of morphosyntax. They argue that illiterate L2 learners' acquisition is fundamentally different from that of literate L2 learners due to brain alterations resulting from learning to read and write. As previously discussed,

Tarone *et al.* argue that alphabetic literacy influences how learners process input, resulting in slower progress. Tarone *et al.* acknowledge that it is still unknown how literacy affects learners' noticing of these linguistic forms, although much attention has been devoted to how instructed L2 learners deal with functional morphology in the input (see overviews in R. Ellis 1999; Robinson 1996; VanPatten, Jessica Williams, and Mark 1996). Some forms may be simpler to read than to hear. Awareness of phonemes develops only in tandem with alphabetic script literacy development. This conclusion is supported by extensive research in three categories: (1) pre-school versus school-age children (Goswami and Bryant 1990); (2) adult illiterate monolinguals (Morais *et al.* 1986); and (3) readers of a logographic writing system (Chinese) with some exposure to alphabetic orthography (Pinyin) versus those without exposure (Read *et al.* 1986).

As previously discussed, Tarone *et al.* (2009) considered the Multi-Dimensional Model to account for differences in language development among learners in the ZISA study (Clahsen *et al.* 1983), noting that the study measured only the amount of education but not literacy. They assume that where learners differ in the extent to which they accurately supply morphosyntactic features, literacy—not just native language schooling—might be the underlying cause. First, there is variation in child L1 acquisition that might be compared to that in adult L2 acquisition according to the Multi-Dimensional Model (e.g., d'Avis and Gretsch 2011). This variation is further illustrated by longitudinal data of three *ab initio* post-puberty American exchange students presented in Section 2.6.3, which indicate that acquisition varies among literate learners. Vainikka and Young-Scholten note that the amount of written input differs among these learners, which may explain why Joan behaves similarly to George. Although their self-reported reading amounts are equal, Joan's self-reported writing output surpasses George's. In other words, differences between oral language and written language as sources of input for literate learners may also contribute to these variations. Miller (2002) points out that oral languages are considered relatively simple, with complexity primarily developed in written language through education. These differences may also be linked to cognitive mechanisms associated with alphabetical literacy.

The effect of written input on language processing is not limited to adult L2 learners. For example, Kail (2002) proposes that the differences in processing observed in older French children may be associated with increased literacy, which involves exposure to

written scripts that provide clearer and more regular morphological cues than what is provided in aural primary linguistic data. In another study of two Farsi-speaking siblings presented in Section 2.6.4, Mobaraki (2007) reports similar results. The child who read more in Farsi and then English showed more rapid development in his production of the regular past *-ed* and third-person singular *-s* than his sister. A higher frequency of exposure to these suffixes could well have been responsible. However, he performed better on tasks aimed at measuring working memory and processing, which may have resulted in better reading, making it difficult to determine the actual cause of the rapid development.

It is unclear whether data from socially marginalised low-literate learners provide the best evidence for literacy's role in L2 acquisition. Ideally, researchers should closely examine the acquisition of low-literate learners who receive sufficient aural input. For example, adult L2 learners in multilingual sub-Saharan Africa can fit this description (Vainikka and Young-Scholten 2011); see also Hill (1970) for the remarkable phonological achievement of adults in illiterate cultures. It is also unclear how input from printed texts influences the acquisition of morphosyntax. Second language learners who read or are read to (by parents or instructors) have access to an additional source of input in which the salient elements, in terms of their position, might vary from those in spoken language, and the frequency of complex constructions can be much greater (see, e.g., Biber 1991; Halliday 1994). Researchers should also look at other non-linguistic factors that may interplay when investigating the role of literacy, such as trauma, ongoing stress, as well as motivation in this learner population. These factors will be discussed in the following section.

2.10 Non-linguistic factors

This section presents non-linguistic factors that may have an indirect negative effect on the process of second language learning (or acquisition). These factors are trauma, ongoing stress in the new county, and motivation. In the realm of second language acquisition, trauma and ongoing stress in the new country have been neglected, either in Usage-based or generative linguistic approaches. Motivation, on the other hand, has been thoroughly investigated, but not in the context of LESLLA learners.

2.10.1 The psychological stress the learners face: Traumatic stress and migration-related stress

Recognising that refugees are a heterogeneous group and do not necessarily face the same issues, 'all refugees are trauma survivors' (Turner 2018: 332). The mental health of refugees could be negatively affected as a result of their traumatic life experiences (Mollica, Wyshak, and Lavelle 1987). Based on the Canadian Refugee Resettlement Project, Beiser (2018) argues that 'refugees, by definition, have suffered pre-migration stress.' By the time a refugee adult learner starts attending language classes in their settled country, they will already have experienced the stress related to relocating to a new country, including leaving behind their friends, community, and even family members.

A considerable proportion of refugees experience psychological distress due to their stay in refugee settlements or detention facilities, both before and often after arriving in a host country (Beiser 2009; Global *et al.* 2018). Pre-migration stress encompasses trauma caused by involuntary separations and the scarcity of essential resources such as food and water, as well as exposure to or personal encounters with violence, cruelty, and torture in the refugee's place of origin.

Unfortunately, the difficulties encountered by refugees do not end with their relocation to a host country. The stress associated with job hunting, experiencing discrimination, and adapting to a new culture, including the need to learn the language and customs of the host country, could potentially have adverse impacts on mental well-being (Adkins *et al.* 1999; Beiser 2009; Beiser, Noh, Hou, Kaspar, and Rummens 2001). Nou (2006) emphasises the simultaneous presence of pre-migration stressors, trauma-related stressors, and continuing stressors in the host country. The author states that the majority of Cambodians face the challenge of reconciling unpleasant historical memories with the challenges of adapting to a new social setting. Nicholl and A. Thompson (2004) highlight the significance of recognizing several elements that impact the well-being of refugee populations, such as previous encounters with trauma, socio-economic disadvantages, a lack of social support, and the isolation resulting from living in exile. Hinton, Rasmussen, Nou, Pollack, and Good (2009) identified three distinct sources of post-migration stress faced by refugees in their study on family-directed anger among traumatised Cambodian refugees: acculturative stress, financial stress, and stress resulting from residing in neighborhoods characterised by frequent violence and gang activity.

Stewart and Martin (2018: 25) use the term 'triple trauma effect' to describe what arises for refugees when they are exposed to stressful events in their home countries, as well as the stress that comes with moving and adjusting to a new culture. Mental health professionals agreed that current constraints and past trauma are mutually reinforcing factors. The immigrants with whom they collaborated encountered considerable obstacles, including the search for affordable housing, supporting their families through low-paying jobs, and overcoming communication barriers.

In addition to this triple trauma, there is a feeling of grief and powerlessness. Even if, through fortuitous circumstances, a refugee has successfully evaded the triple ordeal of relocation, they are unable to return to their place of origin as long as the threat of violence or persecution persists (Beiser, Dion, Gotowiec, Hyman, and Vu 1995). The refugee experience may threaten their very sense of identity and agency (Adkins *et al.* 1999), and the traumatic events that characterise the experience could potentially result in them losing their 'sense of control, connection, and meaning' (Herman 1992: 33). Given how the refugee experience alters or disturbs everything the refugees used to believe about themselves, their role in society appears to capture the essence of the refugee experience when he refers to it as a 'social earthquake' (as cited in Ward, Furnham, and Bochner 2005: 221).

2.10.1.1 Traumatic stress and language learning

The impacts of trauma may make language acquisition extremely challenging since language learning requires the ability to maintain control, develop connections, and construct meaning—all of which are impeded by trauma (CCVT as cited in Finn 2010). Individuals with a history of trauma may exhibit manifestations such as low self-esteem and confidence, an all-or-nothing mindset, dissociation (also known as 'spacing out'), impaired concentration, flashbacks of prior trauma, and challenges in developing trust with others. The resource guide for LINC instructors, designed for teaching learners who have a history of trauma (LISTN 2016), includes a list of behaviors, problems, and negative emotional responses that may be used as indicators that a learner has undergone trauma.

Stevens's (2001) investigation of Cambodian people living in Australia revealed that 91% of the participants had both post-traumatic stress disorder (PTSD) and symptoms of depression. A significant number of study participants had symptoms that had an

immediate impact on their learning and cognitive abilities, including headaches (69%), insufficient concentration (58%), and memory impairment (57%). According to a study by Bremner *et al.* (1993), Vietnam veterans with PTSD from the war had significant impairments in both short-term recall and long-term memory storage.

Anecdotal and instructor accounts have investigated the effects of trauma on learning. Horsman's (2000) seminal study focuses on the effects of violence on adult learners. Although Horsman's research centres on the prevalence of domestic and sexual violence against women, it does not examine language acquisition explicitly. In addition, almost all of the study is devoted to native English speakers.

Few studies have thoroughly investigated the specific impacts of trauma on adult ESL learners. Ying (2001) conducted a case study with a Vietnamese man who experienced severe headaches, recurring nightmares, and feelings of anxiety. Part of the physician's treatment advice was based on the man's admission that his lack of focus was the reason for his difficulties in attending his ESL class. During his sessions with Ying, the man revealed that he had endured thirteen years of imprisonment in a 'reeducation camp' where he was subjected to cruelty. Santoro (1997) reports on the challenges faced by adult refugee learners who have experienced trauma, including poor motivation, high levels of anxiety, and reluctance to engage in verbal participation. According to McDonald (2000), trauma-related symptoms such as difficulty concentrating, dissociation, and memory loss significantly influence the cognitive function and learning ability of adult refugees in Canada.

The most detailed research on the effect of stress on language acquisition is Söndergaard and Theorell's (2004) nine-month longitudinal study of Iraqi refugees in Sweden. The authors demonstrated that migrants exhibiting more symptoms of PTSD acquired a second language at a significantly slower pace. The correlation between the pace of language acquisition and the severity of PTSD was much stronger ($p = .001$) compared to the duration of the language classes attended ($p = .178$) or the level of education in the home country ($p = .116$). The authors note that prior research has focused mainly on examining the cognitive effects of trauma among individuals who are native speakers and have suffered battle trauma, rather than refugees. Consequently, the potential impact of trauma on language learning in these refugee populations may have been significantly overestimated. They suggest that promptly identifying and treating PTSD in refugee

populations might significantly reduce post-migration stress and enhance the process of integrating refugees into the host community.

It is unclear whether trauma and ongoing stress influence the acquisition of morphosyntax. On the other hand, Eamer, Fernando, and King's (2017) study reveals some evidence that trauma influences acculturation and, ultimately, the acquisition of second languages. Acculturation increases the likelihood of interaction with native speakers, resulting in input. As stated above, generative linguistics regards input as an important aspect of morphosyntax acquisition.

2.10.1.2 Theoretical perspectives: the impact of trauma on second language acquisition

Notwithstanding the critical nature of early refugee acculturation and the significance of second language acquisition for refugee integration, the field of second language acquisition has yet to formulate a theoretical framework regarding the influence of trauma on second language acquisition or provide guidelines for instructing trauma survivors. In contrast, the acquisition of a second language serves as evidence that traumatic experiences significantly affect cognitive processing, attention, and memory, all of which are critical to the acquisition of a second language (Gordon 2011).

According to Schmidt (1990), learners can only use what they consciously notice in the input as intake for acquiring a second language. Ellis (2006) asserts that explicit learning and consciousness are crucial in second language acquisition. This includes awareness of negative evidence and attention to language forms in various contexts. Research in cognitive psychology indicates that traumatic events can modify neural circuitry in the brain, potentially having long-lasting effects on learning. The human brain is equipped to detect, analyse, and retain information about potentially dangerous events. The 'fight or flight response' engages all regions of the brain and body, allowing humans to act swiftly in the face of threats. Due to the involvement of multiple brain areas, traumatic experiences have a broad and profound impact on the brain (Perry 1999).

For those suffering from PTSD, this vital reaction hinders their ability to digest information effectively. When exposed to a particular sound (such as fireworks) or situation (such as a crowded room), these people may have a terror reaction. Consequently, the brain reacts to the stimulus by eliciting a widespread fear reaction, as if it were happening for the first time. This flashback, or reactivation of the fear reaction, triggers a

heightened state of vigilance in the brain. However, to effectively handle, retain, recall, and react to new information, the brain must be calm and attentive. During a condition of panic, the brain of a trauma victim, particularly the regions of the cerebral cortex responsible for language acquisition, becomes engaged in the fight or flight response and is unable to comprehend new information (Perry 2006).

The emotional filter theory proposed by Krashen (1982, 1985) presents potential avenues for investigating language learners who have experienced trauma. Krashen theorised that emotional elements, such as anxiety and self-confidence, have a significant impact on the process of acquiring a second language. Learners' low self-esteem and severe anxiety work together to increase their emotional filter, which acts as a mental barrier that hinders the reception of comprehensible information by the language learning device. Although Krashen's concept has been criticised for its lack of empirical support, later neurological investigations show that anxiety levels among learners affect the transmission of information and storage. According to neuroimaging and neurotransmitter research, the flow of information into the brain's areas responsible for higher cognitive memory consolidation and storage is hampered during stressful situations. Simply put, when stress activates the brain's emotional filters, it blocks the flow of information to higher cognitive networks, resulting in a halt in the learning process (Willis 2007).

The study of the neuroscience of learning and the role of the affective filter is significant for those who have experienced trauma since they often suffer from anxiety and poor self-esteem. The functions of anxiety and attention in language learning, as described by theories such as the emotional filter and cognitive processing, have not been studied in the context of learners who have experienced trauma.

2.10.2 Motivation in LESLLA learners

Larsen-Freeman and Long (1991) conducted a comprehensive analysis of the literature to identify many factors that influence an individual's language acquisition abilities beyond their first language. One of these factors is the learner's motivation, which has been extensively investigated in the field of second language acquisition. For example, those with integrative motivations, such as immigrants or those entering into a marriage with someone from a foreign culture, tend to achieve more success compared to those driven by instrumental incentives, such as seeking job promotion. These types of motivation do

not fit the context of LESLLA learners, who are viewed as refugees forced to immigrate to a new country.

In their ongoing discussion of motivation and LESLLA learners, Dalepro and Grinden (2022) examined whether Dörnyei's L2 Motivational Self-System (L2MSS) can be applied to LESLLA learners. The idea of the future self is central to this theory: the gap between an individual's current self and their envisioned or expected self in the future is seen as the basis of what drives learners. The ideal self refers to the learners' envisioned picture of themselves that they want to become, whereas the ought-to self is influenced by the social context and pertains to the expectations learners feel they should conform to. In other words, Dörnyei posits that learners are motivated by their goals or their perceived obligations.

Dörnyei's concept is mainly based on the results of a thorough survey he conducted with large groups of students. The groups varied in terms of their backgrounds (Dörnyei and Taguchi, 2010); however, there was not much difference in terms of their educational levels, as observed by Dalepro and Grinden. These researchers noted that completing the questionnaire requires advanced reading skills. Consequently, they are quite certain that LESLLA learners were not included in Dörnyei's research.

The researchers considered the timing of applying L2MSS to LESLLA learners and whether these learners could imagine their future selves in an unfamiliar country, particularly in the early stages of language acquisition. They also questioned, but were unable to explain, how cultural background and social status would influence this ability to imagine one's future self.

Currently, policymakers in many European countries are increasing the requirements for civic integration (Rocca *et al.* 2020). In the UK, a B1 or higher level is required for permanent residency or citizenship. However, most LESLLA learners are below A1, with exceptions for refugee status and humanitarian protection. In Norway, this increase in requirements is framed as motivational. Dalepro and Grinden, however, point out that the theory of Ryan and Deci (2000) raises doubts about this approach, at least for LESLLA learners. According to Ryan and Deci, learners need competence to maintain motivation, which implies that they must believe that the task at hand is doable. If the expectations are too high or the tasks too difficult, learners may withdraw (Schalge and Soga 2008).

As discussed, refugees in general face traumatic events and migration-related stress. Many psychosocial impacts of trauma, such as dissociation, attention difficulties, missing classes, or withdrawal from courses, can be interpreted as a lack of motivation (Kerka 2002). Individuals who struggle with unpleasant memories and adapting to life in the United States may have difficulty maintaining their commitment to learning English and regularly attending classes. Therefore, the effect of traumatic events and ongoing stress in the new country should be included when exploring motivation in LESLLA learners. Furthermore, the timing of when motivation is activated should be considered in relation to persistent trauma and ongoing stress. If trauma continues to persist, the lack of motivation persists.

2.11 Summary

In this chapter, I have reviewed the early studies on L1-L2 functional morpheme orders (Bailey *et al.* 1974; Brown 1973; De Villiers and de Villiers 1973) and developmental stages of negation (e.g., Cancino *et al.* 1978; Ravem 1974, followed by influential studies on the acquisition of morphosyntax by low-educated adult L2 learners (see, e.g., According to the results of these studies, the path of language development is consistent and unaffected by general cognitive variables. For example, Hawkins (2001) reviewed these studies and concluded that the developmental path of L2 morphosyntax is fixed, regardless of learners' L1, type of exposure, or education. Using Schmidt's (1994) Noticing, Tarone and Bigelow (2005) and Tarone *et al.* (2009) contested this view and argued that literacy has an undeniable effect on the acquisition of morphosyntax.

On the other hand, generative studies (Vainikka and Young-Scholten 2007; Vainikka *et al.* 2017; Young-Scholten and Strom 2005) looked at the acquisition of morphosyntax by learners with little or no education in their home language. These studies were undertaken under Organic Grammar adopted in this thesis. Again, the analysis of these studies demonstrates that the path of L2 morphosyntax is fixed and unaffected by literacy. Now we will move on to the next chapter, in which I will review the literature on overgeneralisation in L1 and L2 acquisition. As noted above, the second aim of this thesis is to explore overgeneralisation by low-educated learners. Again, overgeneralisation will be investigated under Organic Grammar.

In this chapter, I have also discussed other variables (e.g., language input, which is

an important factor when it comes to the acquisition of morphosyntax, as well as other non-linguistic factors that could potentially impede the process of acquiring a second language, such as trauma, ongoing stress in the new country, and motivation). In the field of second language acquisition, particularly in the generative approach, language input remains under-examined and requires closer attention, while the impact of trauma and ongoing stress experienced in a new country has been overlooked in both usage-based and generative linguistic approaches. On the contrary, motivation has been extensively examined, although not within the specific context of LESLLA learners.

Chapter 3. Overgeneralisation in the acquisition of morphosyntax

3.1 Introduction

It has long been recognised that most non-native (second language or L2) adult learners struggle with the acquisition of verbal morphology (Clahsen 1983; Vainikka and Young-Scholten 1994, 1996a,b). The absence and variability of verbal morphology on thematic verbs characterise this acquisition process. For researchers such as Clahsen (1984) and Meisel (1997), these persistent morphological issues provide evidence against the operation of Universal Grammar in L2 acquisition. On the other hand, many of these issues are related to the status of functional categories at the initial state of non-native grammars (Eubank 1992, 1993, 1996; Grondin and White 1996b; Ionin and Wexler 2001; Lakshmanan 1993/94; Schwartz and Sprouse 1994, 1996; Vainikka and Young-Scholten 1994, 1996b; White 1996, among others). As discussed earlier, according to some researchers, functional projections such as IP and CP are initially absent in learners' grammars and are built incrementally as a result of the interaction between UG and L2 input (Eubank 1993/94, 1996; Vainikka and Young-Scholten 1994, 1996a). The lack of target-like verbal morphology, according to these researchers, indicates that L2 learners do not project the relevant syntax. In contrast to this perspective, others who advocate the Strong Continuity hypothesis believe that functional categories and their projections are available at the initial state of L2 (e.g., Grondin and White 1996b; Lakshmanan 1993/94; Lardiere 2000; Lardiere 1998b; Schwartz and Sprouse 1994, 1996; White 1996), but the absence of verbal morphology is due to other reasons, such as a problem in the realisation of surface morphology (Haznedar and Schwartz 1997; Lardiere 2000; Lardiere 1998b). As mentioned in Chapter 1, the second aim of this thesis is to investigate other aspects of verbal morphology issues in L2 learners: the use of superfluous functional elements

(non-target elements), either single function words or multi-word expressions, as overgeneralised forms during the acquisition of verbal morphology.

Overgeneralisation of bound morphemes in L1 acquisition is something researchers are familiar with. This has long been observed in children when they are in the process of acquiring affixes such as the regular past tense *-ed* or plural *-s*, which they initially apply to irregular forms. For example, they attach *-ed* to 'go' to form 'goed' (see e.g., Becker 1975). In L2 acquisition, Wagner-Gough (1975) reported a slightly different type of overgeneralisation by a young L2 English learner who overused *-ing* forms with verbs regardless of their functions. Unlike in L1 acquisition, L2 researchers observed that learners oversupply single free morphemes (e.g., the preposition 'for'; Lakshmanan 1993/94). The current literature in L2 acquisition also shows that L2 learners use *be*-forms in a non-target way. These forms are usually 'is' and 'are' and often precede thematic verbs (e.g., he 'is' open the door). This sort of overgeneralisation often relates to verb movement under the Minimalist Program (e.g., Chomsky 1981). Some researchers analysed the use of non-target 'be' as a topic marker (e.g., Ahn 2003, 2006; Shin 2001). As mentioned in Chapter 1, overgeneralised forms are not always single words; L1 and L2 researchers in the Usage-based approach have reported that learners use non-target multi-word utterances in the course of language development (see e.g., Clark 1974; Hakuta 1974, 1976; Myles 2004; Myles *et al.* 1998; Peters 1983). Recent Usage-based research has shifted toward a paradigm in which multi-word expressions are crucial for the acquisition of syntax (e.g., N. C. Ellis 2003; N. C. Ellis 2012; Eskildsen and Cadierno 2007; Myles 2004; Myles *et al.* 1998). This perspective was contested by Bardovi-Harlig and Stringer (2017), who argued that multi-word expressions did not serve as a catalyst for learners' acquisition of morphosyntax, but rather the opposite was the case.

In the context of LESLLA learners, there is emerging research on the use of non-target elements in L2, which has been framed within various theoretical backgrounds. Under the theory of Organic Grammar, Vainikka and Young-Scholten (2019) and Vainikka *et al.* (2017) examined this phenomenon in L2 German and L2 English. In L2 Dutch studies, researchers such as Blom and De Korte (2011), Julien *et al.* (2016), and Van de Craats and Van Hout (2010) investigate this usage under the Economy Principle (Chomsky 1995). Mocciaro (2019a,b) follow these non-target elements along the path of development of L2 Italian, which is rooted in the theory of the Basic Variety.

This chapter is organised as follows. Section 3.2 briefly covers overgeneralisation in L1. Section 3.3 reviews overgeneralisation in L2 and the relevant accounts, including evidence from several studies. This section also includes overgeneralisation of multi-word strings. Section 3.4 presents overgeneralisation by L2 adult immigrants, which has been investigated under three different theoretical frameworks: the Minimalist Program (Chomsky 1995), Organic Grammar (Vainikka and Young-Scholten 1994, 1996a,b, 2010), and the Basic Variety (W. Klein and Perdue 1997). Section 3.5 provides a summary of this chapter.

3.2 Overgeneralisation in L1A

Overgeneralisation (or overregularisation errors) is probably the most thoroughly investigated topic in child language, and several articles and books have been inspired by it. Overgeneralisation has been observed when children add a regular inflection to a verb or noun that should be represented in an irregular form in English (see e.g., Becker 1975; Berko 1958). What are we referring to when we say this? Numerous inflectional paradigms have exceptions—verbs or nouns that do not accept the regular inflectional endings. The English noun plural morpheme *-s* is a classic example. To make regular nouns plural, the morpheme *-s* is added to the end of the noun. So, the noun 'car' becomes 'cars'. This regular rule, however, does not extend to all nouns. For instance, the word 'foot' becomes 'feet', not 'foots', when making a plural noun, and the word 'man' becomes 'men', not 'mans'. The same applies to the English past tense. The regular rule is to add the past tense morpheme *-ed* to the end of the verb, so the verb 'start' becomes 'started', and the verb 'jump' becomes 'jumped'. However, there are some irregular forms. The verb 'run' becomes 'ran', not 'runned', and the verb 'fly' becomes 'flew', not 'flied'.

When children acquire inflections, they frequently make errors by overusing regular inflections in English rather than producing correct irregular verbs. For instance, at the age of 22 months, Lara started to oversupply *-ed*, as in 'Lara doed it' rather than 'Lara did it'. When she was about three years old, instead of saying 'I don't like men', she said 'I don't like mans'.

Overgeneralisation does not occur at the beginning of the acquisition of inflections. Instead, it follows a stage where children appear to use inflections appropriately, followed

by a final stage where overgeneralisation errors disappear. Cazden (1968) neatly identified this developmental pattern by analyzing the speech of three of the most prominent children in the literature of child language acquisition: Adam, Eve, and Sarah (Brown 1973). As shown in Figure 3.1, Cazden suggested four stages in the development of the plural morpheme *-s*. Stage A shows a total absence of inflection. For example, in response to the question 'What are those?', children may respond 'toy' rather than 'toys'. Stage B was 'defined by occasional production with no errors or overgeneralisations' (Cazden 1968: 436). In other words, children began to use both regular and irregular plurals appropriately, though neither were abundant. Then, in stage C, children began to supply a greater number of plurals but also made overgeneralisation errors by incorrectly applying the plural *-s* rule. Finally, in stage D, children started to use the plural ending correctly nearly all the time. This pattern of development—correct usage (stage B), followed by overgeneralisation errors (stage C), and then correct use again (stage D)—is referred to as U-shaped development, also known as U-shaped learning (Berko 1958), because if the data were shown on a graph with time as the X-axis and the proportion of correctly inflected words as the Y-axis, the corresponding curve would be U-shaped.

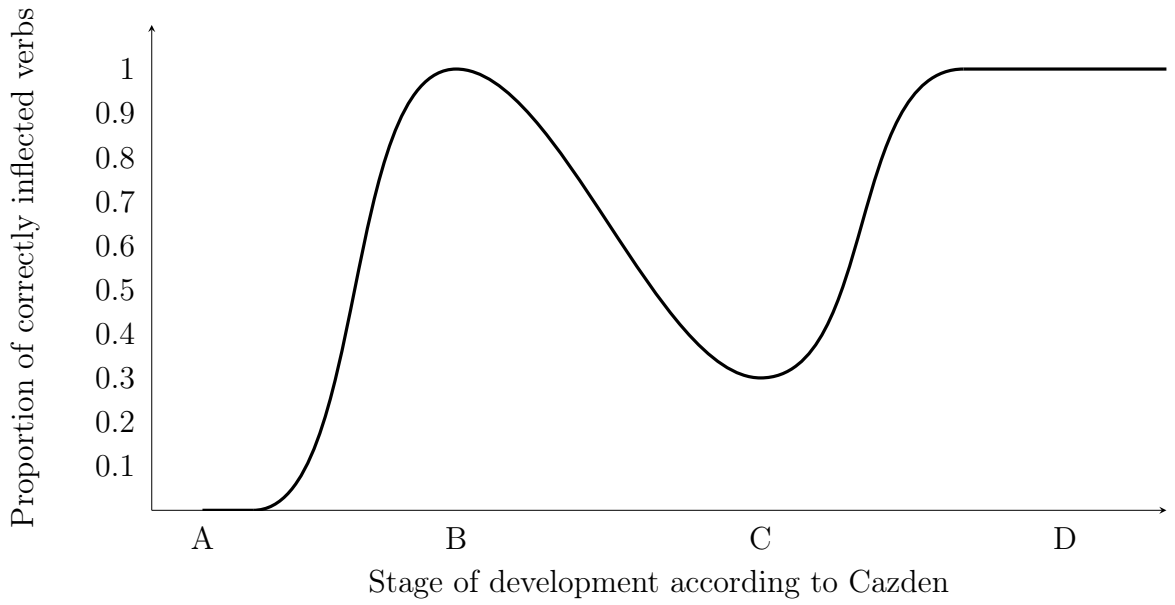


Figure 3.1: Proportion of correctly inflected verbs; redrawn from Cazden (1968)

The U-shaped development has been one of the arguments taken as strong evidence supporting Chomsky's claim that language acquisition is internally driven rather than environmentally driven.

3.3 Overgeneralisation in L2A

Wagner-Gough (1975) reported a slightly different type of overgeneralisation in L2 English by a six-year-old Iranian boy, Homer. She found variation in his form-function mapping of *-ing* with verbs regardless of their functions, as shown in (3.1).

(3.1) a. I'm going and found them. (Immediate intention)
b. I am tomorrow going /in/ beach. (Distant future)
c. I'm playing with that Mark. (Past)
d. O.K. Sitting down like that. (Imperative)

Now, we will move on to overgeneralisation in L2 acquisition. Apart from the bound morpheme *-ing*, L2 learners also use free morphemes as overgeneralised forms instead of bound morphemes, such as the past tense *-ed* and third-person singular or plural *-s*. Examples of overgeneralised forms in L2 include the preposition 'for', multi-word strings, personal pronouns, or *be*-forms.

3.3.1 Overgeneralisation of other types in L2

L2 learners also use other types of overgeneralised forms in their oral production. The use of words such as 'for', 'with', and the conjunction 'and' has long been observed by Lakshmanan (1993/94), who investigated longitudinal data from a Spanish-speaking child named Marta, first reported by Cazden *et al.* (1975). The data consist of a total of 15 samples, obtained roughly once every other week over a period of eight months. At the onset of the study, Marta had been in the USA for approximately one month and had been attending an English-only nursery school, meaning that the first sample of her data was collected after she had been exposed to English for one month. Marta's age range during the study was 4;6 to 5;1;14 (years; months; days). The examples listed in (3.2) illustrate the insertion of the preposition 'for', which appeared in the earliest samples (i.e., Samples 1 and 2) of spontaneous speech. According to Lakshmanan, these might represent syntactic functions in Marta's interlanguage grammar.

(3.2) a. Carolina is for English and Espanol. (S1) [Carolina speaks English and Spanish.]
b. *For* the baby. (S2) [in response to: What are you doing?]

c. *For* hello. (S1) [=Say hello]

d. *For* the lamp. (S1) [In response to: What did I do? (said while turning the light off and on)]

e. *For* you. Or *For* mommy. (S2) [In response to: What do you hear?]

f. For the head the little girl. (S3) [In response to: What's cookie monster going to do?]

g. I going *For*, *For* little chair. (S3) [— I'm going to get a/the little chair.]

(Lakshmanan 2014)

Similar to the *for*-constructions seen in (3.2), the preposition 'for' also emerged in the context of a picture description task in which the researcher introduced Marta to pictures containing an action (one at a time). Each picture depicted a character engaging in an action, and Marta's task was to tell the researcher what the character in the picture was doing. The descriptions of Marta's utterances in Sample 2 are shown in (3.3).

(3.3) a. This is the boy *for* the cookies. [Picture of boy eating cookies]

b. This is the girl *for* (shaking her hands) tamboron. [Picture of girl playing the tambourine]

c. This is the girl *for* the baby. [Picture of girl giving a baby/doll a bottle]

d. This is the boy *for* the milk. [picture of boy pouring milk into a glass]

e. This is the girl and the boy *for* the boot. [picture of a girl and boy putting on their boots]

(ibid.)

Lakshmanan noticed that the utterances in the examples above lacked main verbs and that the preposition 'for' preceded the object noun phrases (NPs). She also observed that the insertion of 'for' had no semantic function but rather a syntactic function—specifically, as a Case assigner. That is, instead of the verb assigning case, 'for' may assign case to NPs. To simplify, the DP (the cookies) in the ungrammatical sentence 'the boy for the cookies' was taken as the complement of an empty (implicit) verb, 'eat'. The preposition 'for', like the infinitival particle 'to' in adult grammar, is claimed to be in INFL (as discussed in Haznedar 1997: 129). According to Lakshmanan, this participant 'knows and obeys the Case Filter requirement', as shown in (3.4). This was because Marta did not violate the Case Filter requirement proposed in UG in any of her utterances.

(3.4) Case Filter: Every phonetically realised NP must be assigned (abstract) a Case.
(Chomsky 1986: 74)

Verbless utterances are not unique to Marta's L2 grammar; they have also been observed in the acquisition of English by other young L2 learners. For instance, as shown in (3.5), Cheo (a native Spanish speaker) produced verbless utterances during the early stages of his L2 development.

(3.5) a. The boy *with* the milk. [describing picture of a boy pouring milk into a glass]
b. The boy *and* the cookie. [describing picture of boy eating cookies]

Xi-Xi, a five-year-old native Chinese speaker who acquired English as a second language, demonstrates a similar pattern. In his initial state of L2 acquisition, Xi-Xi often omitted lexical verbs. Although the transitive verb is omitted from the instances in (3.6), the linkage 'and' and the verb 'is' are overtly present in his utterances. As with Cheo's early L2 grammar, the coordinator in this Chinese-speaking child's verbless sentences seems to act as a Case assigner for the object NP.

(3.6) a. This cat *and* car. [= The cat is driving a car.]
b. *and* house. [=The pig is building a house.]

Huang (1971; cited in Lakshmanan 2014) reported on the acquisition of L2 English by a young Chinese child, Paul (5 years old). In the early stages, Paul produced constructions that required the copula (e.g., this is X). However, unlike Marta, Paul generally omitted the copula. Huang further stated that Paul usually paused between 'this' and the next category (usually a noun). Perhaps, as Lakshmanan suggests, the pause acts as a placeholder for Infl features.

3.3.2 Be-production

The use of non-target *be*-forms has been observed by some L2 researchers and analysed through two different theoretical accounts: (1) verb movement under the Minimalist Program and (2) the Topic Marker Hypothesis. Both accounts will be presented in the following two sections, along with supporting evidence from studies. Consistent with the current literature, I will use these terms (non-target *be*-forms, non-target *be*-insertion, non-target 'be') to refer to the overgeneralisation of 'be' (e.g., the boy 'is' play).

3.3.2.1 Verb movement

Researchers such as Fleta (2003), García Mayo *et al.* (2005), Ionin and Wexler (2001), and S. Yang (2014) associate the use of non-target *be*-forms with verb movement in English, which states that English thematic verbs (i.e., base verbs) do not undertake verb movements from the base position in the syntax (Déprez and Pierce, 1993; Pollock 1989 Pollock 1989; Chomsky 1993; Paradis *et al.* 2008). That is, finite lexical verbs do not rise out of the VP domain to check formal features such as tense and agreement. They remain in the VP domain, in contrast to finite auxiliaries (i.e., modals and *be*-forms), which undergo movement to T, thereby checking their features. For example, finite auxiliaries appear directly below the T layer and precede the negation marker 'not'. In English interrogative formations, finite auxiliaries undergo subject–auxiliary inversion (sometimes called subject–auxiliary operation). As an illustration, *be*-forms raise to the C layer, changing places with the subject in matrix questions; therefore, the word order becomes auxiliary–subject, in contrast to the order in English declaratives, which is subject-verb. Thematic verbs in yes/no questions and wh- questions, on the other hand, do not undergo these movements; they take a non-finite form, and the do auxiliary must be inserted. Under Minimalism, all finite verbs must be raised to the functional layer for feature checking. In terms of English, a finite auxiliary overtly raises, whereas main verbs covertly raise. As the basis for this phenomenon, Guasti and Rizzi (2001) suggested that overt movements are more easily expressed in the morphology than covert movements; therefore, L2 learners might overgeneralise *be*-forms. The rest of the section presents studies that have observed the use of non-target *be*-forms in English.

3.3.2.2 Supporting evidence

The use of non-target *be*-forms was first reported by Bickerton (1975), who studied Navajo learners of English and found that *be*-forms are often used before the thematic verb, as in (3.7), to mark the past tense. Bickerton demonstrated that both 'was' and 'were' were employed as preverbal tense markers. In all of the examples in this chapter, the non-target elements (overgeneralised forms) are presented in italics.

(3.7) Ron *was* drive to work.
(ibid.)

Similar use of *be*-forms has also been observed in a more recent study examining how Arabic learners encode the past tense in their L2 English. Moursi (2013), for example, investigated the writing of Egyptian Arabic L2 English learners at the low to intermediate levels. He observed seven non-target strategies for marking the past tense. For example, the learners used past *be*-forms before lexical verbs. However, unlike the data on Navajo learners, Egyptian Arabic learners followed *be*-insertion with different types of verbs, such as simple past verbs, past participles, or gerunds. Importantly, these forms often disagreed with the subject.

Muftah (2016) found comparable use of *be*-forms as past tense morphemes in the oral production data of Yemeni-Arabic English learners, elicited from picture description tasks. The examples in (3.8) demonstrate that the use of *be*-forms can be followed by regular or irregular past verbs (e.g., 'kicked' and 'spent') or gerunds (e.g., 'playing'). 'Was'-production after the plural subject 'the two boys' indicates that *be*-forms do not always agree with the subjects.

(3.8) a. They were playing football, then one of boy *was* kicked the ball over a wall.
b. And the two boys *was* spent time in the party.
(ibid.)

However, Muftah argues that the use of *be*-forms is an agreement morpheme rather than a past tense morpheme. In the Arabic past perfective tense, the verb must include an agreement suffix, and the tense is marked covertly (Benmamoun 2000). Reflecting the transfer of split tense/agreement marking in the L1, Arabic learners may use *be*-forms to encode agreement in past clauses in their L2 English. The lack of agreement between the plural subject 'the two boys' and 'was' as an overgeneralised form, as in (3.8b), suggests that the target-like morphology has not been acquired, and not that the used *be*-forms lack the agreement feature.

On the other hand, García Mayo *et al.* (2005) found that Basque-Spanish children, who were learning English as an L2, often used 'is' before the main verb, as in (3.9), which was also used as a default form, as in (3.9b). The context was in the present tense, and regardless of the number of subjects, 'is' was used. Moreover, the researchers found frequent uses of the personal pronouns 'he', as in (3.9c), and 'they', as in (3.9d), both of which again appeared before the main verbs.

(3.9) a. The little boy *is* want the frog.
b. David and the dog *is* see two frog.
c. The wolf *he* opened the door.
d. The father and the woman *they* love.
(García Mayo *et al.* 2005)

When the participants were grouped into three proficiency levels, *is*-production occurred more often at Level 1, whereas *he*-production was more frequent at Level 2. García Mayo *et al.* argue that *is*-/*he*-production represents a syntactic structure of Basque and Spanish in which an overt morpheme fills T. Assuming that agreement morphemes in these L1s are weak pronouns and are hence equivalent to English pronouns such as 'I', 'he', 'they', and the like (see, e.g., Kato 1999). García Mayo *et al.* explicitly argue that the use of 'is' and 'he' acts as placeholders for T, which respectively carry T and D features.

However, in other studies, using *be*-forms as functional morphemes has been interpreted as an L2 strategy rather than an L1 transfer. In longitudinal and cross-sectional data from Spanish learners of English, Fleta (2003) observed that although code-switched Spanish verbs are often preceded by *be*-forms, as in (3.10), because these verbs are infinitive, the data indicate that *be*-forms were projected in I, whereas thematic verbs stayed in their base position.

(3.10) a. *Is* probar a little.
try.INF
'Goldilocks tries a little porridge.'
b. To one friend of mine *is* picar one.
sting.INF
'A jelly fish stung a friend of mine.'
(Fleta 2003)

Furthermore, *be*-forms are used in negatives and interrogatives, as shown in (3.11). In example (3.11a), the use of 'is' precedes the negation 'don't', suggesting that 'is' precedes the negative. The inserted *be*-forms may precede or follow the subject in interrogatives, as shown in (3.11b) and (3.11c). However, the insertion of *be*-forms in the higher position,

which appears to involve subject-be inversion, emerged only in the data of more advanced learners.

(3.11) a. Yolanda *is* don't going to come.
b. The frog *is* have head?
c. *Is* he can jump?
(Fleta 2003)

According to these observations, Fleta asserted that *be*-insertion is similar to *do*-insertion and that *be*-forms may be projected in either I or C.

This hypothesis is convincing in that *be*-insertion is commonly used before the acquisition of inflectional affixes and gradually disappears after these affixes are acquired (Fleta 2003; Ionin and Wexler 2001). Furthermore, in Ionin and Wexler's (2001) investigation of Russian learners' L2 production, the inserted *be*-forms often agree with the subjects, as in (3.12). The thematic verbs after *be*-forms were rarely inflected, with the examples in (3.12b) and (3.12c) being the only exceptions in their data.

(3.12) a. They *are* help people when people in trouble.
b. I *'m* never opened this one.
c. He *is* goes to elementary school.
(ibid.)

S. Yang's (2014) analysis of written production data also supports this hypothesis. S. Yang's large-scale, cross-sectional data showed that L1 Chinese learners of English as a second language rarely used *be*-forms in non-finite positions, as in (3.13a). Almost 50% of *be*-forms in past contexts were in the past/past participle forms, as in (3.13b) and (3.13c).

(3.13) a. Mrs. Forestier seemed not to *be* recognize Matilda.
b. The robbery *were* die.
c. One man *was* looked after the people.
(ibid.)

These findings, according to S. Yang, indicate that the inserted *be*-forms were in I and included inflectional features (e.g., T). In contrast to Ionin and Wexler's (2001) analysis,

S. Yang found that around 20% of thematic verbs that followed *be*-forms were inflected, as in (3.13c).

To summarise this section, although the primary purpose of these studies is to examine how tense and agreement are projected in interlanguage, non-target *be*-forms have been frequently observed. According to these researchers, *be*-forms fill in the heads of functional projections, mainly depending on their sentential position. Based at least in part on elicitation methods, most *be*-forms in some data were in present tense contexts, whereas most *be*-forms in other data were in past tense contexts. In some data, only the default *be*-forms were reported, but in others, *be*-forms tended to be inflected for tense and/or agreement. Furthermore, the data differ on whether the following thematic verbs can be inflected. Although these studies have different results, they all support the idea that *be*-forms are functional morphemes that replace inflectional affixes in the target language.

3.3.2.3 Topic Marker Hypothesis

Several researchers regard the use of non-target *be*-forms before the main verbs as a topic marker transferred from the L1 (Ahn 2003, 2006; Shin 2001). According to Chafe (1976), a topic is defined as a particular referent in the minds of the interlocutors in given contexts. The function of a topic is to limit the domain to which the main predication can apply. In contrast to the subject, which is a sentence-internal concept and determined by the verb, the topic is a discourse concept determined by extra-sentential elements (Li and S. A. Thompson 1976). Proponents of the Topic Marker Hypothesis claim that L2 learners can confuse subjects and topics, with non-target *be*-forms that follow the subject being interpreted as topic markers that follow topics. This is because subjects are often topics at the same time.

Since topics are not limited to subjects, *be*-forms may be added after phrases in a wider variety of categories, provided they are topics. Importantly, the main idea behind the Topic Marker Hypothesis is that there is a specific topic morpheme in some L1s, and only learners from these L1s can use *be*-forms as topic markers owing to L1-topic markers transfer.

3.3.2.4 Supporting evidence

A few researchers support the Topic Marker Hypothesis. Shibata (2006), for example, analysed the writings of Japanese university students and found that *be*-forms are often used after non-subject but topic phrases, such as in (3.14). Both 'Nago' and 'that' are not subjects, but *be*-forms follow them as they are topics.

(3.14) a. Nago *is* still Typhoon stay...
b. That *is* my mother made.
(ibid.)

Furthermore, the added *be*-forms are sometimes followed by a clause with a subject and a verb, as in (3.15). Even if the initial phrase and *be*-forms are omitted, the sentences that follow are complete (e.g., 'my friend come from Kobe'; 'everything went well'). The pronoun 'she' in (3.15d) looks like a subject, but unlike the other phrases preceded by a 'be-form', it is not governed by the verb 'went'. These sentences have a topic-comment structure, which is the most unmarked form in languages where topics are important in their grammar system (i.e., topic-prominent languages, Li and S. A. Thompson 1976).

(3.15) a. Today *is* my friend come from Kobe.
b. Today *was* few customers came at store.
c. Today *was* I watched movie.
d. She *was* everything went well while.
e. My summer vacation is best of memory *is* I join to Kin town immigrants experiment program.
(Shibata 2006)

In support of this analysis, these sentences become more target-like when the *be*-forms and their preceding phrases are replaced by 'as for XP', an English topic-marking expression, as in (3.16).

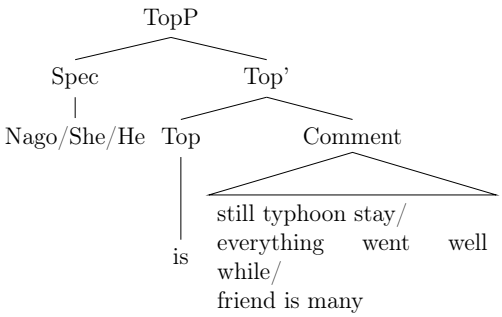
(3.16) a. As for today, my friend come from Kobe.
b. As for her, everything went well ...
c. As for my summer vacation, best of memory is I join to ...
(ibid.)

Similar utterances have been reported in L2 English production data from Korean learners. Shin (2001) attested to the double insertion of *be*-forms, as in (3.17a), and Kim (2011) observed that the learners' sentences contain the construction 'Topic–be–(thematic subject)–V', as in (3.17b).

(3.17) a. He *is* friend is many.
(Shin 2001)
b. Japan *is* (I) went.
(Kim 2011)

The structure of topic-comment sentences with an added *be*-forms can be shown as below in (3.18). In the structure, topics are established in the [Spec, TopP] position, and *be*-forms occupy the head of the TopP, carrying the [+Topic] feature. The comment can be sentential, with arguments that might be omitted.

(3.18)



3.3.3 Overgeneralisation of multi-word expressions in L2

3.3.3.1 Overview of formulaic language

Multi-word expressions are traditional expressions that must be memorised in the same manner as single words. When we speak a language we are familiar with, we rely on these expressions and employ them as easily as single words from our mental lexicons. Formulaic language was once believed to constitute only a small portion of native speakers' normal production because the emphasis was on the power of syntax to generate an infinite number of sentences. The current view is that native language is permeated with recurrent formulaic sequences (e.g., Erman and Warren 2000; Nattinger and DeCarrico 1992, among others), and the prevalence of formulaic sequences in language production was reported early in L1 studies under the framework of Phraseology (e.g., Altenberg 1998; P. Foster 2001; Howarth 1998; Pawley and Syder 1983; Sorhus 1977).

Such expressions lead learners to creative use (Young-Scholten and Naeb 2020). For example, Clark (1974) found that when her three-year-old son began to say 'wait for it to cool' at mealtimes, he later used 'wait for it' with other phrases, such as 'wait for it to dry'. However, such expressions are likely to be overused in contexts where they are not appropriate. Another similar example, this time in L2 acquisition, was observed in a young boy, Homer, studied by Wagner-Gough (1975). For instance, he used 'is it' to form questions like 'Is it bicycle is Judy' and used 'where's' to produce declaratives such as 'Where's Mark is school', where the target utterance was 'Mark is at school'. Haznedar (2001) reported a similar instance in L1 Turkish learner Erdem's L2 English production data. Erdem used 'It's a' before a negative sentence, as in 'It's a I don't know', in response to 'What are these?'.

The overgeneralisation of multi-word sequences has also been observed in L2s other than English. Myles (2004) identified three common chunks in L2 French that are frequently overgeneralised: *J'aime*, *j'adore*, and *j'habite* ('I like', 'I adore', 'I live', respectively). For example, learners produced sentences such as *Monique j'aime* ('Monique I like') with the target utterance 'Monique likes', and *La garçon j'aime le cricket?* ('The boy I like cricket') with the target utterance 'Does the boy like cricket?'. The following subsection presents the conflicting views on the role of multi-word expressions in the development of morphosyntax, followed by experimental studies (Eskildsen 2009; Hall 2010; Myles, Mitchell, and Hooper 1999) designed within the Usage-based framework.

3.3.3.2 The role of formulaic language in L2 acquisition under a Usage-based approach

In the 1970s, the role of multi-word expressions in language acquisition was a controversial topic, and it has been revived in more recent studies. From the start, a much-debated question was whether formulaic utterances serve as the database for the acquisition of syntax. Two different perspectives arose, the broad contours of which are readily apparent in modern applied linguistics.

Researchers such as Clark (1974), Hakuta (1974, 1976), and Peters (1983) claim that these expressions (e.g., chunks) serve as the main input for the acquisition of syntax. The general agreement was that in the early stages of first and second language acquisition, children identify formulaic sequences in input, adopt them for use, and subsequently segment and analyse them. The analysis may occur later, in part due to neurological

development and the resulting development of cognitive analytic skills. This argument was considered valid for both L1 and L2 acquisition. According to Clark, her son built his acquisition on 'wait for it' in 'wait for it to cool' to produce 'wait for it to dry'. Clark argued that 'many, though not necessarily all, the productive rules originated as invariable routines, which were in use for some time with the original lexical items before new lexical items were inserted' (p. 4) and concluded that child language 'becomes creative through the gradual analysis of the internal structure of sequences which begin as prepackaged routines' (p. 6). Hakuta (1974; see also Hakuta 1976) analysed longitudinal data over a thirteen-month period from a five-year-old Japanese girl, Uguisu, an untutored/naturalistic L2 English learner. After the first five months of exposure to English, Hakuta observed that Uguisu used the pattern 'do you' in *how*-embeddings (e.g., 'I know how do you you write it' instead of 'I know how to write it').

In the context of child L2 acquisition, Wong-Fillmore (1976) adopted and expanded on the formulas-syntax link. She studied five Spanish-speaking children acquiring L2 English. Not only did she observe the use of acquisitional formulas, but she also noted the conventional nature of many everyday linguistic interactions in this environment (e.g., 'finish your milk'; 'it's time to clean up'). She postulated that children acquire formulas for forming and sustaining social relationships and then use them as the seed material for their grammar acquisition.

On the other hand, several researchers claimed that, despite the presence of memorised chunks used for communicative reasons, the acquisition of syntax seems to be independent of formulaic language. In Brown and Hanlon's (1970) account of the L1 acquisition of *wh*-questions by Adam, Eve, and Sarah, they argue that early memorised routines are not tools for acquisition. The *wh*-chunks 'what's that/what doing' preceded the development of the *wh*-structure; afterwards, *wh*-questions emerged in preposed form (e.g., 'what he wants?'). However, the first *wh*-chunks were not rebuilt in accordance with emerging grammar; they remained chunks. The forms that would be expected if such strings followed well-documented stages of question formation (e.g., 'what that is?' or 'what you are doing?') were not attested, casting doubt on the notion that chunks serve as databases for the acquisition of syntax.

In child L2 acquisition, Hatch (1972) and Wagner-Gough (1975) reached a similar conclusion: formulaic and analytic speech are two distinct processes that coexist as al-

ternative communication strategies. Perhaps the most influential and frequently cited article in this debate cycle was the extensive review by Krashen and Scarcella (1978). They summarised their stance after analysing data from acquisition studies as well as research on verbal aphasia (agrammatism) and left hemispherectomy:

> routines appear to be immune to rules at first. This clearly implies that routines are part of a system that is separate from the process generating rule-governed, propositional language. It is also evidence that automatic speech does not 'turn into' creative constructions. Rather, the creative construction process evolves independently (p. 286).

Formulaic sequences are not always target-like forms in interlanguage grammar. In a subsequent study, Scarcella (1976) classified 55 (of 450) responses as partially acquired routines, including 'watch up' (=watch out), 'shut off' (=shut up), and 'time is off' (=time is up). Scarcella concluded that this implies the learner may treat routines as analysable units rather than unanalysable wholes. Similar observations have been reported in pragmatics. Although the main aim was not to investigate non-target formulaic sequences, Osborne (2008) found instances of pluralised adjectives in formulaic sequences produced by advanced learners (e.g., 'bathrooms fittings', 'diets ads', 'adults smokers') and other formulaic units ('in others words', 'the good olds times', 'basics rights'). These non-target forms seem to be similar to the overgeneralisation of the plural *-s* presented earlier in this chapter and refer to creativity. That is, these learners extend the *-s* to adjectives.

Since the beginning of the twenty-first century, the formulaic approach to acquisition has resurfaced as a powerful stream of thought in applied linguistics (see, e.g., N. C. Ellis 2002, 2003; Eskildsen and Cadierno 2007; Eskildsen 2009, 2012; Myles 2004; Wray 2002, 2008). In the Usage-based theory, the general agreement is that the underlying units of languages are constructions and that language primarily consists of formulaic sequences. In construction grammar, constructions range from the smallest unit of a word or simple morphemes (e.g., the grammatical morpheme *-ed*) to complex and abstract syntactic frames such as Subject-Verb-Object-Object (e.g., 'John gave Sam a book'). Therefore, the past tense morpheme *-ed* or the present tense agreement *-s* attached to the main verb ('V+ed' and 'V+s' respectively) are considered constructions.

N. C. Ellis (2002, 2003), in favour of the claim that formulas play a role in L2 grammar and that acquisition progresses from large to small, propose that L2 acquisition proceeds

through the following putative development sequence: formula, via low-scope pattern, to creative construction. According to N. C. Ellis (2002), a formula is initially taken from input and then holistically retained in long-term memory as a form-meaning pair. The formula is not analysed by the learner during the intake stage. Determining the degree to which the formula affects the acquisition process as a whole is complicated by the fact that some formulaic sequences are not necessarily acquired in their entirety by adults but are readily learnt due to their high frequency and extreme functionality. Hasselgren (1994) reported that L2 learners, including advanced learners, overuse high-frequency basic words (e.g., 'very') and avoid less frequent words (e.g., 'exceedingly'). These high-frequency basic words are known as lexical teddy bears. Similarly, high-frequency phrases (e.g., 'nice to meet you') or functional phrases (e.g., 'put it on the table') are more readily learnt and used than less frequent phrases. N. C. Ellis (2002) refer to these as phrasal teddy bears. These phrases are more likely to be candidates as seeds in construction. In contrast, other formulaic sequences, such as idioms, are not as readily learnable because they are less frequent and often non-transparent in the input (N. C. Ellis 2008; N. C. Ellis, Simpson-Vlach, and Maynard 2008; Pawley and Syder 1983).

One of the most cited studies into the role of formulaic language in L2 grammar is the large-scale longitudinal study by Myles (2004; see also Myles *et al.* 1999) on 16 beginning and 60 intermediate learners. Myles found that not only did the young learners use formulaic sequences, but they also segmented them to meet their more complex communication demands throughout the course of the two-year research project. Initially, learners were able to use unanalysed sequences to communicate simply, but then they began to break the sequences and use some parts in different ways as their routine classroom communication needed to be developed beyond simple communication of personal information, including discussion of third-person activities and characteristics. When the requirement for third-person communication increased, the process of segmentation started and subsequently accelerated. Myles categorised these early grammars as being composed of only lexical projections and formulaic sequences, without evidence of functional categories (a claim made in Organic Grammar):

> Chunks do not become discarded; they remain grammatically advanced until the grammar catches up, and it is this process of resolving the tension between these grammatically advanced chunks and the current grammar which drives the learning process forward (Myles 2004: 125).

Additionally, data analysis demonstrated a strong correlation between the use of formulaic sequences and general language development. For example, learners who failed to retain formulaic sequences showed slower progress, whereas those who retained these sequences moved forward.

Eskildsen and Cadierno (2007; see also Eskildsen 2012) examined exemplar-based learning as proposed in the Usage-based theory. Specifically, Eskildsen and Cadierno studied the trajectory of L2 negation constructions by two adult English learners (pseudonyms: Carlos and Valerio) involved in the Multimedia Adult English Learner Corpus, which consists of audio-visual recordings of classroom interaction in an English as a Second Language (ESL) context in the United States. The data showed that Carlos's inventory of L2 negation began with 'I don't know', followed by a gradual expansion to include other verbs and pronouns (Subject+don't+Verb).

Eskildsen (2009) analysed longitudinal oral L2 classroom interaction for the use of the auxiliary 'can' by one student, Carlo. The word 'can' first appeared in the data in the formula 'I can write'. However, Eskildsen noted how formulas are interactionally and locally contextualised, which means that they may be transitory in nature, their deployment over time being occasioned by specific recurring usage events.

Bardovi-Harlig and Stringer (2017) offered a generative account of formulaic language acquisition as an alternative to current Usage-based accounts. They analysed conventional expressions (e.g., 'I agree with you'; 'I'm sorry I'm late'; 'I really appreciate it') produced by L2 learners of English with different linguistic backgrounds (Indo-European, East-Asian, Semitic, and Central-Asian languages). The data were drawn from 271 learners who participated in three previous studies (Bardovi-Harlig 2009, 2014; Bardovi-Harlig and Vellenga 2012), including Bardovi-Harlig and Stringer's (2017) study. All studies examined the same collection of conventional expressions through oral and written production tasks. The learners were grouped into four different levels as part of a seven-level intensive English programme. The combined results show that conventional expressions are analysed and exhibit interlanguage structures, non-target use of tense, hesitations and repairs, and non-use of intensifiers. For example, learners used 'I am agree with you' rather than 'I agree with you'. The results also demonstrated pauses and repairs, which indicate that the learners were 'engaged in syntactic processing' (ibid.: 77). Bardovi-Harlig and Stringer argue that formulaic language (at least conventional expressions)

does not drive the development of syntax. Instead, the development of syntax as an independent process causes changes in the production of conventional expressions.

3.4 Overgeneralisation by LESLLA Learners

As mentioned previously, overgeneralisation in the context of LESLLA learners has been investigated under three frameworks: Organic Grammar, the Minimalist Program, and the Basic Variety. Researchers of L2 Dutch, such as Julien *et al.* (2016) and Van de Craats and Van Hout (2010), investigated overgeneralisation under the Minimalist Program (Chomsky 1995). For L2 English and L2 German, Vainikka and Young-Scholten (2019) and Vainikka *et al.* (2017) explored overgeneralisation under the structure-building approach of Organic Grammar (Vainikka and Young-Scholten 1994, 1996a,b, 2011). Mocciaro (2019a,b) studied this phenomenon in the developmental path of L2 Italian based on the Basic Variety of W. Klein and Perdue (1997). These researchers have argued for the status of overgeneralisations as dummy auxiliaries (Julien *et al.* 2016; Mocciaro 2019a,b; Van de Craats and Van Hout 2010; Verhagen 2013) and as placeholders (Vainikka and Young-Scholten 1994, 1996a,b, 2011). All these approaches lead to the same conclusion: Overgeneralised forms represent syntactic functions. In this section, I will present this work, including the underlying approaches.

3.4.1 Dummy auxiliaries in L2 Dutch

Dummy auxiliaries (i.e., overgeneralised forms or placeholders) have been reported in both L1 and L2 Dutch. Researchers investigating dummy auxiliaries have interpreted them through various theoretical approaches, including the generative syntactic approach, the semantic-pragmatic approach, and a more theory-neutral approach. Starren (2001) conducted a longitudinal study based on oral production data from two Turkish and two Moroccan learners of Dutch as a second language and claimed that the non-target insertion of 'is' serves an aspect-marking function. Starren found that, at the outset of acquisition, learners used adverbs to mark tense and aspect in sentences, as shown in (3.19a). At more advanced stages of acquisition, these learners replaced adverbs with 'is', as in (3.19b). She concluded that 'is' functions similarly to adverbs by marking durative or perfective aspect, depending on its position in the sentence.

(3.19) a. Ik altijd ongeluk maken.
I always accident make.INF
Target: Ik maak altijd ongelukken.
'I always make accidents'.
(Starren 2001: 151)

b. Hij is terug lopen.
he be.3SG back walk.INF
Target: The boy cleans the room.
'He is walking back'.
(ibid.: 185)

On the other hand, the use of dummy auxiliaries is often associated with verbal morphology; for example, see Chomsky's (1995) Economy Principle. Dummy auxiliaries are thought to be easier to access than thematic (main) verbs because they are stored inflected in the lexicon, where they can be retrieved and inserted directly into the slot/position of a functional head. In contrast, producing lexical verbs in the same slot requires movement, which is considered a more costly operation. Van Kampen (1997) and Zuckerman (2001) reported that monolingual Dutch children use dummy auxiliaries in main clauses that require verb movement but not in subordinate clauses that do not require verb movement. This finding has inspired researchers to investigate this phenomenon in the acquisition of L2 Dutch. Therefore, the rest of this section reviews a cluster of studies exploring the use of dummy auxiliaries during the acquisition of L2 Dutch verbal morphology. The primary aim of these studies was to determine whether dummy auxiliaries carry syntactic meaning.

Van de Craats and Van Hout (2010) investigated the use of dummy auxiliaries by LESLLA learners[1]. The study was based on longitudinal data from six Moroccan women over approximately 15–18 months. Mina, Hayat, and Najat were Berbers who spoke Moroccan Arabic fluently and were all instructed in L2 Dutch using a communicative approach. The data came from a large corpus of semi-spontaneous speech and experi-

[1]The participants joined a literacy class where they learnt the Latin script; they were not illiterate in the Arabic script; however, two of them, Mina and Nezha, had never attended primary school. They learnt to read through relatives. At the outset of the study, they were all below level A1 of the Common European Framework (i.e., they were all beginners), though some of them had been in the Netherlands for many years (12 years) when they started the course. After 0.5–2 years of schooling, they had all mastered a basic level of vocabulary. At the end of data collection, Mina and Zohra had achieved a level A2 of proficiency

mental data. The researchers examined the emergence of dummy auxiliaries (specifically 'is'-patterns or *ga* 'at') in connection with verb movement and inflectional morphology. The results suggested that 'is' patterns have a purely structural function and appear to be a precursor to the acquisition of finiteness. The researchers claim that Moroccan learners of Dutch use dummy auxiliaries as an intermediate stage between L1-based sentences without finiteness marking and target-like sentences with inflected and raised thematic verbs. Moreover, they observed that learners in the early stages of acquisition do not use dummy auxiliaries but rather use thematic verbs that seem finite but are, in fact, default forms or forms that do not specify tense, person, or number (see, e.g., Prévost and White 2000). At a later stage of development, learners use dummy auxiliaries, first in default forms such as 'Hij is komt' ('He is comes'; target: 'Hij komt' 'He comes'). They subsequently acquire infinitival forms and merge dummy auxiliaries with infinitives in similar elicitation contexts (e.g., 'Hij is komen' 'He is come'; target: 'He comes'). This second step, according to the researchers, is important because it reveals that learners have distinguished between finite and nonfinite verb forms. Van de Craats and Van Hout (2010) examined two dummy auxiliaries: 'is' and *ga* (or 'gaat'). They claim that learners use dummy auxiliaries because they are bootstrapped from the copula *zijn* and *ga* 'at' from the Dutch thematic verb *gaan* (go). The authors suggest that 'is' and *ga* 'at' play the same role in the acquisition of finiteness: both emerge during an in-between stage when finiteness marking is absent, and both help learners acquire the distinction between finite and nonfinite verb forms. Notably, Van de Craats and Van Hout emphasise that 'is' and *ga* 'at' lack a semantic role. In terms of 'is', Van de Craats and Van Hout (ibid.: 491) argue that 'there is nothing that points to any lexical content whatsoever'.

Blom and De Korte (2011) also analysed data from Moroccan and Turkish L2 Dutch learners, both children and adults. Their key inquiry was whether these learners' dummy auxiliaries were used as a strategy to avoid verb movement. To test this, they compared the use of dummy auxiliaries in main and subordinate clauses. In Dutch, finite thematic verbs and dummy auxiliaries must be raised in the main clause of the sentence but not in the subordinate clause. Therefore, if movement plays a role, learners are predicted to use dummy auxiliaries in the main clause but not in the subordinate clause, or at least not as frequently. The results from sentence completion tasks demonstrated that dummy auxiliaries were more common in main clauses than in subordinate clauses, in-

dicating that the use of dummy auxiliaries is a strategy to avoid verb movement or a sign of a learner's failure to move verbs in less advanced learners. Blom and De Korte's (2011) investigation compressed some semantically vacuous auxiliaries; thus, it is unclear whether their conclusions apply to all other auxiliaries to the same extent. Furthermore, the researchers do not address whether dummy auxiliaries have any semantic meaning.

Verhagen (2013) broadened the investigation into the use of dummy auxiliaries by incorporating data from two studies: production and comprehension data (Verhagen 2009). The first study focused on production data from a relatively large sample of Moroccan adult learners (n. 55) of Dutch. The aim of this study was to investigate the assumption that dummy auxiliaries are used as a structural strategy when learners have difficulty producing finite verbs (Blom and De Korte 2011; Van de Craats and Van Hout 2010). The results of the production data demonstrate that there is a negative correlation between the use of dummy auxiliaries and finiteness. The use of *zijn*, in particular, was negatively correlated with the use of verb raising. The more dummy auxiliaries learners used, the fewer thematic verbs they raised in their utterances. This corroborates Blom and De Korte's (2011) and Van de Craats and Van Hout's (2010) findings that learners use 'is' if they have problems placing verbs in the second position. In contrast, the use of the auxiliary *hebben* was strongly linked to both subject-verb agreement and verb raising, supporting prior findings that the use of *hebben* is required for the development of finiteness in Germanic languages (Dimroth 2008; Schimke 2009; Vainikka and Young-Scholten 1996a, 2011) for L2 German; (Jordens and Dimroth 2006; Verhagen 2011) for L2 Dutch. According to the findings of Study 1, there is an early stage in Moroccan-Dutch interlanguage in which learners use 'is' with both infinitival and finite verb forms. At a higher stage, learners use it with infinitives and past participles instead of finite forms, consistent with Van de Craats and Van Hout's (2010) findings. The second study was based on longitudinal comprehension data from eight Moroccan adult learners of Dutch. The aim of this study was to determine whether learners associate *is* with an aspectual meaning, a claim made by Starren (2001). The results revealed that the learners did not link *is* to the perfective aspect. They predominantly matched dummy verb sentences (e.g., *Hij is de auto wassen* = 'He is the car wash'; 'He is washing the car') to pictures of an individual washing a car, as opposed to pictures of a completed action (e.g., a washed car).

Recently, as with the aforementioned studies, Julien *et al.* (2016) explored whether adult learners of L2 Dutch employ dummy auxiliaries as a structural syntactic step towards the acquisition of finiteness. The researchers contributed to the study of dummy auxiliaries in three key ways. First, they expanded the participant pool to include not only L1 Arabic speakers from Turkey and Morocco but also L1 Berber Tarifiyt speakers. Second, they investigated not only the production but also the comprehension of (dummy) auxiliaries in Dutch, aiming to determine whether the use of dummy auxiliaries is devoid of semantic meaning. Third, they employed a variety of elicitation tasks to stimulate dummy use in various contexts and investigate factors that may hinder or increase the occurrence of dummy auxiliaries. As reported by the participants' teachers, all participants had not yet fully mastered functional morphology and word placement in Dutch. In light of this, learners with higher levels of Dutch proficiency produced more dummy auxiliaries than those with lower proficiency levels. Julien *et al.* divided their participants into three groups based on their proficiency levels, and all three groups employed dummy auxiliaries extensively.

Similar to what was observed by Verhagen (2013), the results from the comprehension task demonstrated that all learners matched the picture depicting the ongoing action to the construction *zijn + INF* and often also to the construction *gaan + INF*. While Verhagen views this as evidence that learners ascribe an imperfective reading to those constructions, Julien *et al.* interprets this as evidence that these two constructions have no semantic meaning in the early stages of language acquisition. According to the researchers, this perspective is supported by the fact that all participants in their study used both *gaan + INF* and *zijn + INF* arbitrarily to describe all three conditions (prospective, imperfective, and perfect) and often added lexical items to express aspect. Aspect is still represented by lexical rather than grammatical means.

Additional evidence that the use of dummy auxiliaries is semantically vacuous is supported by the absence of any significant effect of the lexical aspect of the verbs tested in the experiments. Dummy auxiliaries were used arbitrarily with all verb classes, regardless of their semantic constraints. The participants in the study did not adhere to the semantic constraints that statives in Dutch do not permit continuous constructions or forms like *zijn + aan het + INF* and *zitten/staan/liggen + te + INF*, and that most transitive and intransitive verbs in Dutch require the use of the auxiliary *hebben* 'have'.

Julien *et al.* indicate that the developmental route these learners follow is the same as that observed in previous investigations. They also suggest that the use of dummy auxiliaries should be split into two stages: an initial stage where both dummy *zijn* and *gaan* are used, and a subsequent stage where the occurrence of dummy *zijn* begins to disappear while dummy *gaan* continues to be used. They conclude that L2 language proficiency affects both the use and the choice of dummy auxiliaries.

3.4.2 Overgeneralisation under Organic Grammar

Building on the findings presented in Section 2.8.4.1, Vainikka *et al.* (2017; see also Vainikka and Young-Scholten 2019) observed overgeneralisation of single-function words and multi-word expressions by non-/low-literate learners during the acquisition of morphosyntax. Data analysis indicated that one learner, Zabila (a Panjabi speaker), was in the process of projecting AgrP, as she supplied the third-person singular *-s* in 5 of 9 instances where the suffix *-s* was obligatory. However, this process of acquisition, as Vainikka *et al.* suggest, is misleading because Zabila's use of the suffix *-s* was not always correct. She adopted the strategy of adding it to verbs regardless of whether the subjects were third-person singular or the words were verbs. Her data indicated that the suffix *-s* was overgeneralised to several content words. The researchers also observed that learners used non-target elements (i.e., additional lexical items). For example, in response to the negation task, Sultani (a Dari speaker) used 'is don't' instead of 'not'. Zabila produced no auxiliary verbs and only a few tense or copula 'be' forms for the task that aimed to collect data on TP projection. However, she overgeneralised the suffix *-s*, not only to main verbs in all persons, singular and plural, but also to other content words, as in (3.20). The first utterance was produced in response to a task where she had to depict a picture of a girl reading a book. The second shows two girls stirring a single pot. Another instance is the insertion of 'I'm', as in (3.21), in a task where single individuals and two or more individuals are displayed engaging in various activities.

(3.20) a. reads books ladies.
b. girls this cookens. . . foods this cookens.

(3.21) a. two guys I'm reading message.
b. three guys I'm washing . . . washing in dishes.

Another example is by Mohammad S, who overused items that are not from the same word class (e.g., 'in the'). In example (3.22a), the stimulus was a picture of a lady reading a newspaper, and in example (3.22b), it was a picture of a boy setting the table. With respect to AgrP-level placeholders, Vainikka and Young-Scholten observed that Rawdha, who is more advanced than the learners above, appeared to use the nominative pronoun 'he' as a placeholder for agreement. The words in brackets were provided as stimulus:

(3.22) a. [afterwards] this woman in the book in the writing.
b. [every day, this boy] in the cook in the spoon in the fish.

(3.23) a. the woman he shower the dog.
b. every morning the boy he get up and drink.
(Vainikka and Young-Scholten 2019: 213)

According to Vainikka *et al.*, these overgeneralisations are predictable: 'Learners are in the process of figuring out which forms mark singular and which mark plural and how auxiliary (do) vs. auxiliary (be) function' (p. 230). In line with Myles (2004), Vainikka *et al.* observed that the learners use holistic or unanalysed chunks (see Table 3.1).

Learner	L1 lit	L2 lit	Place holder	Task	Responses
Zabila VP	0	Lowest	n/a	All tasks	Overgeneralisation of –s to nearly all verbs
Amro NegP	0	Lowest	You need I am / I'm	Habitual action in 3rd singular	You need is smoking; I am read; I'm cook; I am is clean; this girl I'm go; this man I'm go
Tazeem NegP			I'm + V-ing	Progressive in 3rd singular and plural	two guys I'm reading; three guys I'm washing
	ok	Some	is go is go to	Negation	(boy) is go to don't drink; is go to no wash; is go to no play; go to no painting; go to no play
			go to is go to	Habitual action in 3rd singular	is go to read; is go to wash; is go to food cooking
			is go; like go to	Progressive in 3rd singular and plural	(singular) is go to eat; (plural) every three like go to cleaning
MohS TP	0	Lowest	in the	Habitual action in 3rd singular	in the drink; in the writing; in the coming
Sultani TP			in in the	Progressive in 3rd singular and plural	in writing; in the eat; all plural; in the cooking; in the no cooking; in writing; in the wash
	ok	Lowest	don't don't like	Negation	is don't open door; don't like; is don't like painting; don't like drive
			for	Habitual action in 3rd singular	think for cornflakes; is reading for a book
			for in	Progressive in 3rd singular and plural	(sg) eat for; (sg) laugh for; (sg) is like for; (sg) is laugh for; (pl) is in cooking for; (pl) is wash for
MohM TP	0	Lowest	I don't	Negation	I don't + subject-verb (object/IO/object) subject + I don't + object I don't + subject-auxiliary-verb
			the	Habitual action in 3rd singular	the smoking; the have
			the	Progressive in 3rd singular and plural	(sg) the play; (pl) the write; (pl) the walk
Naz AgrP	ok	Good	dislike	Negation	dislike washing; dislike driving; dislike to open

Table 3.1: Placeholders in acquisition of TP and AgrP

What is interesting about their data is that even when these sequences come from categories other than expected, they are still closed-class elements. These words and sequences are not randomly selected from the input; learners do not simply choose content words that frequently appear in the input (e.g., table, book, or bus). This provides strong evidence that they subconsciously recognise and employ function words after identifying them in the L2 input they encounter.

These single words or sequences produced by learners seem to convey a syntactic function. Vainikka *et al.* hypothesise that learners are in the process of acquiring TP and AgrP projections and, as a result of their continuing access to Universal Grammar, they

know from the syntax of human languages that every phrase requires a head (T for TP and Agr for AgrP). In other words, UG leads them to fill the head, but because they are unsure of what exactly fills that head, they instead use functional elements other than the target elements.

As reviewed in Section 3.3.3, much recent work within the Usage-based framework has shifted toward an approach in which multi-word expressions are vital for the acquisition of syntax. According to this approach, multi-word expressions constitute the relevant input for the development of syntax because accurate storage of constructions allows for statistical induction of abstract grammatical categories and, hence, target-grammar behaviour through general cognitive principles. Vainikka and Young-Scholten (2019; see also Young-Scholten and Naeb 2020) included the Usage-based approach's debate in their work on Organic Grammar and find Bardovi-Harlig and Stringer's (2017) claim that multi-word expressions are not the main input for the development of morphosyntax convincing. At the same time, they do not dismiss the possibility that multi-word expressions contribute to acquisition.

Individual differences exist in how learners use placeholders at the time data was gathered. To begin with, (1) there are some differences in the words and sequences recruited. This may be due to the projections that learners are acquiring (e.g., for MohH, DP (determiner phrase) in his use of 'the + verb') or functional elements that have been the focus of classroom instruction. (2) Exposure to instruction varies depending on the teachers with whom learners interact and their attendance. (3) Not all learners make use of placeholders. (4) Placeholders were utilised by (a) learners who have progressed beyond the VP and NegP stages of Organic Grammar but have not yet reached the CP stage; and (b) non-literate learners in the sample, who tend to employ placeholders that are not directly related to the verbal head (e.g., 'the' or 'in'). According to Vainikka *et al.* (2017: 26), 'this may be due to greater reliance on auditory as compared to visual memory.'

3.4.2.1 Predictions of placeholders under Organic Grammar

As presented in Chapter 2, analysis of previous studies under Organic Grammar (Vainikka and Young-Scholten 1994, 1996a,b, 2010) on adult L2 speakers of English, Korean, Italian, Portuguese, Spanish, and Turkish who were acquiring German in naturalistic contexts

shows that early syntax acquisition follows a similar pattern to that of L1 acquisition. Adult L2 learners, assuming full access to Universal Grammar (UG) (e.g., White 1989), use the same mechanisms as children for positing functional projections. Recent investigations on instructed learners of two other L2s confirm this: Chen's (2017) two-year study of English-speaking adults' production of L2 Mandarin and Kahoul's (2014) cross-sectional study of Arabic and Mandarin adults' production, perception, and processing of TP and AgrP in English.

As previously stated in Chapter 2, functional projections are classified by UG into three groups: the VP-group (verb phrase), the IP-group (inflectional phrase), and the CP-group (Vainikka and Young-Scholten 2010). Under Organic Grammar, UG provides a set of potential syntactic heads to L1 or L2 learners (presumably as an ordered list), and each head is subconsciously considered by learners as they determine whether the language input provides evidence for that head. This is evident in the incremental acquisition of functional projections from the bottom up, ultimately resulting in the mature tree structure. Within the structure-building approach of Organic Grammar, learners' successful production of the relevant head in the target language indicates that the corresponding structure is being built. Conversely, failure to identify the relevant heads in the ambient input to posit the appropriate functional projection may result in the use of non-target production of functional elements, i.e., placeholders.

Recall from Chapter 2 that in the 1990s, Vainikka and Young-Scholten analysed data from the study of naturalistic adults who participated in the cross-sectional Heidelberg project. Recently, Vainikka and Young-Scholten (1996b) revisited the same learners, focusing more specifically on the use of non-target elements in the low-intermediate stage (TP and AgrP) of development, during which learners begin to posit functional projections. The findings indicated that learners steadily and gradually projected INFL-level syntax and recruited non-target elements. For example, Tomá, who had been in Germany for less than four years, marked tense with a modal and participle instead of using an *-n* form, *habe*, or a raised main verb.

(3.24) a. Ich muss gesehen (= yo lo he visto – 'I have seen it').
I must see-past
[Ich habe das/es gesehen./Ich sah das/es.]
'I saw that/it
(Vainikka and Young-Scholten 2019)

While Becker (1975) argue that such use was related to the frequency of the word *muss* on the manufacturing floor where Tomá and the other participants worked, Vainikka and Young-Scholten (2019) provide a more rational explanation, suggesting that it functioned as a placeholder for the head of TP as Tomá searched for target-like material to fill the head.

Vainikka and Young-Scholten (ibid.) also examined the data of Italian, Portuguese, and Spanish speakers in the ZISA study, where the researchers proposed a uniform path of acquisition for all learners, who were divided into two groups: those who produced inflectional morphology and those who omitted it. This indicates variation in the acquisition rate, possibly due to the level of education or other external factors affecting the learner's search for a verbal head in the input. Vainikka and Young-Scholten (ibid.) observed that there were three verb forms in the data of Romance language learners that did not mark agreement: *-n*, the non-finite form in German, and *-ed*, and bare verbs. There were almost no auxiliaries or modals developed in the early stages of acquisition for L1 Italian (Salvatore and Bongiovanni) and L1 Spanish (Jose and Rosalinda) learners of German. However, Rosalinda supplied utterances of *wolle = will* 'want' (wollen nonfinite), which Vainikka and Young-Scholten refer to as a placeholder because she was tentatively projecting TP. Furthermore, she produced five instances of the copula *ist*. Other instances of the use of non-target function words were observed in Jose, whose longitudinal data reveal very rapid acquisition. Jose produced 39 instances of *ist* and a single instance of *bin* 'am' in the first recordings. Vainikka and Young-Scholten claim that the overused *ist* can be a placeholder for the head of TP or AgrP. Based on this analysis and the analysis of L2 English, Vainikka and Young-Scholten proposed predictions of placeholders for L2 German and L2 English. Table 3.2 presents only the predictions of placeholders for L2 English. For L2 German, see Vainikka and Young-Scholten (ibid.).

Phrase and head	Examples	Head identification	Placeholders predicted
AspP	Progressive aspect (*–ing* suffix) Is the action on-going?	Easy: *–ing* straightforward to identify as head as it is a syllable and varies little (this phrase excludes forms of auxiliary 'be').	No
NegP	The morpheme *not* Did the action take place or not?	More difficult: requires forms of 'do'.	Yes
TP	Past tense (*–ed* suffix) When did the action take place?	A challenge: existence of regular and irregular past tense morphology.	Yes
AgrP	Subject-verb agreement; the suffix *–s* in 'he walks' Who did something?	The greatest challenge: weak paradigm + confusion about what *–s* marks (plural, possessive, agreement).	Yes

Table 3.2: Placeholders in L2 English

3.4.3 Overgeneralisation in L2 Italian

Overgeneralised forms have been attested in L2 Italian data. These forms are referred to as interlanguage constructions, non-target constructions, or non-target analytical constructions, as they involve overgeneralised uses of *essere* 'to be' and 'fare' 'to do'.

Investigation of these forms has been conducted within the continuum of L2 Italian varieties, which is based on the Basic Variety theory of W. Klein and Perdue (1997), presented in the preceding chapter along with subsequent studies. These studies emphasise the importance of incorporating a post-Basic Variety stage, characterised by more complex morphosyntactic elements. For example, L2 Italian researchers have described several aspects of learners' morphosyntax. Table 3.3 provides a summary of the key findings on the acquisition of verbs and utterance organisation in L2 Italian, as reported by Banfi and Bernini (2003) and Andorno, Bernini, and Ramat (2003), based on decades of L2 Italian research (for a detailed description, see Mocciaro 2019a,b).

	Variety		
	Pre-basic	*Basic*	*Post-basic*
Gramatical categories	None	Verb/Arguments	Verb/Arguments
Morphology Organisation of the utterance	None Pragmatic	Basic forms (uninflected) Semantic-syntactic	Inflected nouns and verbs Syntactic
Dependency from the context	High	<.....>	Low

Table 3.3: The continuum of the L2 Italian varieties, adapted from Mocciaro (2019b: 108)

Bernini (1989) reported several analytical constructions in which the overgeneralised forms of *essere* 'to be' and, to a lesser extent, *avere* 'to have' occurred in combination with unanalysed lexical forms, so that the grammatical meaning and the lexical meaning were encoded separately. In other words, *essere* and *avere* acted as auxiliaries that carried grammatical meaning instead of the lexical verb, which conveyed non-linguistic thoughts. These non-target constructions have been viewed as temporary solutions for expressing functions that learners appear to be aware of, even if they lack the target inflectional morphology necessary to encode them. The use of non-target constructions, as Benazzo and Starren (2007) suggests, constitutes a distinct interlanguage stage (analytical stage), which precedes and paves the way for the development of finiteness (i.e., it creates a provisional analytical stage). The non-target constructions are predicted to disappear as learners progress along the route of L2 morphosyntax development. The presence of interlanguage constructions that still belong to the basic variety is, instead, only occasional.

Mocciaro (2019b) investigated these forms through a longitudinal study of 20 low-educated learners with various L1s. She observed that the use of *fare* 'to do' occurred as soon as the learners entered the post-basic variety, while 'essere' 'to be' emerged at a slightly more advanced stage when the learners were in the process of acquiring the copula 'be'. In line with Bernini (1989), Mocciaro argued that the use of overgeneralised forms was a temporary strategy before the acquisition of morphosyntax. Mocciaro linked the persistence of non-target forms to the slow development of morphosyntax, consistent with Bernini's (ibid.) claim. However, the data in her study did not allow for observations of more advanced levels of morphosyntax due to the learners' limited morphosyntactic development.

Based on Vainikka *et al.*'s (2017) claim that the use of overgeneralised forms (particularly multi-word utterances) was more frequent among non-literate learners, Mocciaro investigated whether the use of *essere* and *fare* was related to the degree of literacy. Mocciaro argued that non-/low-literate learners predominantly used these forms, whereas literate learners used them more erratically and transiently. Additionally, in line with Vainikka and Young-Scholten's (2019) observation that learners selected only function words (not content words) as overgeneralised forms, Mocciaro classified these forms as functional elements (e.g., copula, auxiliary verbs, and light verbs such as 'to do'). Moc-

ciaro noted that the learners used these elements to convey grammatical functions in their interlanguage.

3.5 Summary

Overgeneralisation in L1 is a well-known phenomenon that occurs when children extend regular grammatical patterns to irregular words, resulting in overgeneralisation (also known as over-regularisation errors). For example, children might add the suffix *-s* to create 'tooths' instead of 'teeth' or the suffix *-ed* to create 'goed' instead of 'went'. These errors indicate that children are applying rules and developing the grammar system in their native language. This chapter has also explored overgeneralisation in child L2 acquisition. An example of overgeneralisation in child L2 acquisition is the use of non-target *be*-forms and multi-word utterances.

More relevant to the current thesis, L2 researchers such as Julien *et al.* (2016), Mocciaro (2019a,b), Vainikka and Young-Scholten (2019), Vainikka *et al.* (2017), and Van de Craats and Van Hout (2010) have observed that low-literate/low-educated learners use overgeneralised free morphemes (i.e., function words or multi-word utterances) while acquiring morphosyntax. These observations have been investigated under various theoretical frameworks. These researchers use three different terms: placeholders in L2 English (Vainikka and Young-Scholten 2019; Vainikka *et al.* 2017), dummy auxiliaries in L2 Dutch (Julien *et al.* 2016; Van de Craats and Van Hout 2010; Verhagen 2013), and interlanguage constructions (Bernini 1989; Mocciaro 2019b). All of these different frameworks lead to the same conclusion. These lexical items have syntactic functions, and learners use them as a strategy while developing their L2 morphosyntax. For example, Julien *et al.* (2016) argue that the use of dummy auxiliaries is a structural step from non-finiteness to finiteness (i.e., inflected verbs). Vainikka and Young-Scholten (2019) and Vainikka *et al.* (2017) maintain that L2 learners use these items as placeholders for the heads of functional projections as a result of the interaction between UG and the input. In L2 Italian, applying the Basic Variety of W. Klein and Perdue (1997), Mocciaro (2019b) claims that L2 learners use these items as soon as they transition from the basic variety to the post variety (a stage with more complex morpho-syntactic elements). In this thesis, I will further examine low-literate/low-educated learners' overgeneralisation of functional morphemes (the use of non-target function words or multi-word utterances)

through the structure building of Organic Grammar and confirm specific Organic Grammar stages at which L2 learners engage in such overgeneralisation. Now we will move on to the methodology chapter of this thesis, which includes the procedure of data collection, production tasks (picture descriptions), comprehension tasks (computerised quizzes), and the counting procedure, including the statistical analysis adopted in this thesis.

Chapter 4. Methodology

4.1 Introduction

This chapter will present the methodology used in this thesis, including the procedure for participant recruitment, a description of the participants, and a detailed explanation of all instrumentation (i.e., tests and tasks). As previously mentioned, the aim of this thesis is twofold: The first aim is to investigate the route of acquisition by low-educated adult L2 immigrants using the theory of Organic Grammar advanced by Vainikka and Young-Scholten (1994, 1996a,b, 2011). The second aim is to explore overgeneralised forms by low-educated adult L2 immigrants within the framework of Organic Grammar. This chapter will also present the counting procedure and the types of tests used for statistical analysis in this thesis.

4.2 Participants

As shown in Table 4.1, the data were collected from 60 learners (33 males; 27 females) who were low- and moderately educated (i.e., they did not finish their compulsory education). All of them were Arabic speakers, the majority of whom came from Syria. The learners' ages varied, ranging from 19 to 60 years. Their length of residence in the UK also varied, ranging from 8 months to 9 years. Their ESOL instruction ranged from none to 5 years, and their native schooling ranged from 2 to 11 years.

Variable	N	%	Mean	Min.	Max.
Age			35	19	60
Sex					
...M	33	55%			
...F	27	45%			
Length of UK residence				8 mns	9 yrs
ESOL instruction				0	5 yrs
Native Schooling				2 yrs	11 yrs

Table 4.1: Demographic and educational background of participants

Based on the questionnaire designed for this study, learners were placed at different levels of English for Speakers of Other Languages (ESOL). To maintain participants' anonymity, I followed the practice of many second-language researchers (e.g., Haznedar 2001; Schwartz and Sprouse 1996; Vainikka and Young-Scholten 1994) by using memorable pseudonyms. Before discussing data collection, the following section will provide an overview of Adult ESOL schools and their corresponding ESOL levels.

4.2.1 Adult ESOL schools in England and ESOL levels

The adult ESOL scheme in the United Kingdom is part of the government's adult literacy and numeracy strategy. ESOL designates language courses for people whose first language is not English and who need to use English to communicate in everyday life situations. It prepares learners for the skills required to obtain employment, citizenship, and further study in graduate programmes or certificates (C. Roberts, Cooke, Baynham, and Simpson 2007). The population of ESOL learners in the UK is very diverse, comprising asylum seekers, refugees, nationals from the European Union (most recently from countries that have recently joined the European Union), established communities, and newly arrived spouses. It also includes those who join families, those with work permits, and those who were born in the UK but spent their childhood abroad (Baynham 2006; Cooke 2006). ESOL learners and other Skills for Life learners have varied educational backgrounds and literacy levels, which is one of the main sources of variation. Although some institutions ensure that people with literacy needs receive targeted tuition, it is still common to find people with little schooling and basic literacy needs being taught in the same class as others with more advanced educational backgrounds.

ESOL courses are provided by a wide variety of organisations, including tertiary

education colleges, schools (for the growing 16–18 ESOL population), government welfare-to-work programmes, the military, prisoners' institutions, and employers. These courses run at five different levels: entry levels 1–3 are introductory or primary-level courses, while levels 1 and 2 are equivalent to the General Certificate of Secondary Education (GCSE). A pre-entry level is available for learners with very little or no English literacy, helping them to learn how to read and write letters of the alphabet and familiar words (D. Foster and Bolton 2018).

4.3 Data Collection

Data were collected through two individual sessions. I met with each participant individually via phone call (WhatsApp). Each session followed the same data collection procedures as listed in (4.1):

(4.1)
a. opening conversation to put the participant at ease;
b. reading assessment/tests designed to measure L1-L2 literacy levels;
c. picture descriptions to test the acquisition of morphosyntax;
d. audio–picture match questions to test the acquisition of morphosyntax in relation to overgenlisation of single words or multi-word strings

The reading tests were adopted from sub-tests (Young-Scholten and Strom 2005) designed to measure reading levels in L1 and L2. To investigate the acquisition of morphosyntax, learners were presented with sets of pictures to describe orally. Each of these sets was designed to test a particular morphosyntactic element. Both the reading tests and picture description tasks were piloted with five participants who had little or no education. The results from the picture description tasks led to the decision to add audio–picture match questions to further assess the acquisition of morphosyntax, particularly in relation to the overgeneralisation of single words or multi-word strings.

4.3.1 Procedure for data collection

In line with research ethics and the impact of Covid-19, I reapplied for ethical approval from Newcastle University, explaining my procedure for data collection. This includes collecting data through WhatsApp, which I believed is more accessible to my learners than any other complex technology (like Teams or Zoom) and the most effective way to collect

data during the impact of Covid-19. I also considered that my target participants are low-literate, and online forms may not be suitable for them. Therefore, in my application, I made it clear that consent would be obtained orally rather than in written form.

Having obtained this ethical approval (through email), I sought help from a coordinator who closely works with immigrants as a volunteer. The coordinator assisted in finding and contacting the required participants and explaining the purpose of my study to them. After obtaining the contact details of the first ten participants, I had friendly small talks with each participant to put them at ease (see, e.g., Tarone *et al.* 2009) and briefly explain the purpose of my study to obtain initial consent. I then asked them some questions about their past schooling experiences and their current L2 English language classes. I also asked them to help me find similar participants (i.e., through the first ten participants, I was able to find all of my participants). Additionally, I asked if I could record the sessions. Finally, I briefly explained my plan for collecting data, including the dates and times of the tests.

On the testing dates and times, each participant was individually contacted through a phone call (both of us using a speaker) via WhatsApp. Again, they were introduced to the purpose of my research. Then I asked for their oral consent since some of the participants are low-literate and online consent forms may not work for them. The consent forms were in English and explained in Arabic, as I share the same native language. I then asked them various questions based on the questionnaire designed for the current research (e.g., the length of their native schooling, L2 English classes, and UK residence; see Appendix A for the consent form and questionnaire). To assess the acquisition of morphosyntax, each participant received the information sheets one by one. All the production testing materials were captured via my phone and saved in my phone photos (in Favourite photos for easier and quicker access) and were sent (one by one) to the participant through WhatsApp. Both the learner and I could see the photo while we were talking. Based on the learner's consent, their descriptions and utterances were recorded using an external recording device.

4.3.2 Reading assessment as a measure of literacy

To determine whether the participant could read and write in Arabic, I asked them to read the first few lines of a story used for the word awareness tasks. I also asked them

to write my name and five additional words to confirm their schooling. For the reading test, I followed Young-Scholten and Strom (2005), who used various tests adapted from other assessments to evaluate a range of basic reading sub-skills, as shown in Table 4.2. This battery of tests included presenting the learners with the alphabet unordered and in different fonts (from the Woodcock Johnson Test, 1989) and survival/environmental signs adopted from the ESL BEST Form B. To assess the learners' decoding skills, they were asked to read a list of ten high- and low-frequency monosyllabic and multi-syllabic words that they might encounter in their everyday lives. These words ranged from high-frequency words, such as 'table' and 'community', to low-frequency terms. The selection of low-frequency terms was based on the assumption that these words would not be included in the sight word repertoires of low-literacy ESL students and would thus reveal whether learners have decoding skills. Due to the focus on phonological awareness and basic reading skills, reading comprehension was not tested.

Tasks in native language	Tasks in English
- read part of a story	- read 12 varied single letter identification
- write five Arabic words	- read 4 survival signs
	- read a paragraph
	- read 10 isolated words

Table 4.2: Reading tests, partly adopted from Young-Scholten and Strom (ibid.: 54)

4.3.3 The picture description

To elicit semi-spontaneous speech that would help explore word-order patterns in English and inflectional morphology, the same testing materials (a set of picture sheets and cards) used by Vainikka *et al.* (2017) were employed (see Appendix B for all these picture sheets). The picture sheets depict actions similar to those that learners might do or encounter in daily life. These sheets include six tasks, each increasing in complexity incrementally. More details will be presented below for each of the six tasks, through which the phenomenon of overgeneralisation of single function words and multi-word utterances was explored.

4.3.3.1 VP task

In this task, the learners were introduced to a series of eight picture sheets. For the first picture sheet, when I showed the picture, I said the beginning of each sentence (e.g., 'the

girl . . . '), paused, and then indicated that they should continue and finish the sentence (see Figure 4.1 for an example). Since the focus of this task was on the word order pattern, the learner was not expected to produce a specific tense (e.g., simple present, present progressive) or specific verbs, nouns, or object names for items in the picture.

Figure 4.1: VP

4.3.3.2 NegP/Negation task

To elicit data that would help investigate negation, the learners were presented with eight groups of pictures, each of which had two similar pictures, as in Figure 4.2. The first picture contains an action, but the second one does not. For example, the first picture contains an image of a girl washing a car, and the second contains an image of a girl just holding a sponge. I asked them to explain the difference between the two pictures. It might be difficult for learners to understand the intended purpose of this task. Therefore, I prompted the learners to say 'no'. Thus, the participants were expected to produce sentences such as 'The girl doesn't wash a car'. In this task, the learners again were not required to pay attention to a specific tense, but would have to avoid structures such as 'no+V' or 'no+V+ing'. For example, if they say 'no+V', I ask them to say a full sentence.

Figure 4.2: NegP

4.3.3.3 AgrP task: Third-person agreement on auxiliaries (progressive)

The learner saw a set of pictures with progressive actions, and I asked the learner questions. For example, I asked, 'What is this girl doing?' I also implied that the action is happening now using words and phrases like 'now' or 'at the moment'. The participant's task was to supply a target syntactic form such as 'be+V–ing', as exemplified in (4.2).

(4.2) The girl is reading a book.

4.3.3.4 AgrP task: Third-person agreement on main verbs

I presented the learners with a series of pictures that depict habitual actions in the present. To coax the learners to focus on the intended syntactic form, I first told the learner that the person in the picture 'does the same thing every day'. Then, I produced the first appropriate sentence, emphasizing the word that refers to habitual actions (e.g., every day, usually). The learner was expected to supply syntactic constructions such as 'S+V-s+O', as shown in (4.3).

(4.3) Every morning, this man drinks coffee.

4.3.3.5 TP (-ed)/tense task

To investigate TP (-ed), the participants were presented with a collection of sequential pictures showing an incident where a man and a woman saw a boat hit a bridge (see Figure 4.3). The participants were then asked to re-tell the story. The learners were expected to use a construction that attached the verbal suffix *-ed* to the main verb ('V+ed'), as in (4.4a), or an irregular past simple verb, as in (4.4b).

(4.4) a. They watched a boat on the river.
b. Then the boat sank into the river.

Figure 4.3: TP (-ed)

4.3.3.6 AgrP task: Agreement on the forms of the verb 'be'

The aim of this task[1] was to elicit data related to the acquisition of the copula ('is' and 'are'). The learners were shown 8 cards (4 sets of two cards each). The first card depicts a person or two people with a profession (e.g., a picture of a doctor). To help the learners understand the purpose of this task, I asked them in Arabic to explain what was on the card in English (e.g., he is a doctor or they are doctors). The learners were expected to produce a sentence using the target copula 'be' ('is' and 'are') according to the card presented.

4.3.4 Comprehension Tasks

As indicated in Chapter 1, the comprehension tasks in this thesis contribute to existing research on Organic Grammar. The aims of these tasks are twofold. The first aim is to compare the comprehension data with the production data, providing a clearer picture of the acquisition path for low-educated learners by comparing free and bound morphemes. These morphemes are crucial in determining the stages of OG. The second aim is to test recent predictions regarding the use of overgeneralised forms under OG. As explained in Chapter 2, OG posits that learners progress through successive syntactic stages in acquiring L2 morphosyntax, influenced by the interaction between the input and Universal Grammar. The learners' task is to subconsciously search in the input for heads (e.g., bound morphemes such as the regular past simple *-ed* or the third-person singular *-s* on main verbs). Once a head is identified, UG equips learners with the tools to subconsciously build a syntactic projection. Failure to identify the relevant target functional head results in the use of non-target elements. Vainikka and Young-Scholten

[1]The research methodology was adjusted in accordance with the impact of COVID-19. The intended/proposed task was that the researcher and the learner would play a card game. The cards, some of which match, would be placed upside down on the table. Each card has a picture of a profession. The researcher and the participant would take two cards, and the researcher would talk about the profession of the first card (e.g., Doctors heal or treat patients) to familiarise the participants with the name of the professions. First, I would say, 'I am a doctor', and then I would ask the learner, 'Who are you?' The learner would be expected to produce an utterance such as 'I am a teacher' to indicate his/her acquisition of the copula *be* with agreement. For matching pictures, I would say, for example, 'I am a police officer', and 'Who are you?' The learner would be expected to say 'I am a police officer'. The researcher would then respond, 'So we are police officers'. The participant, therefore, would understand the intended purpose of the task, and he or she would say for the following card, for example, 'We are farmers'. Regarding non-matching pictures, the interviewer would ask the participant 'Who are you? And who am I?' The learner would be expected to produce a sentence such as 'You are a teacher'. To elicit data for the copula 'is', I would have placed 10 cards face up on the table and then pointed to a card and said 'He is a doctor'. Then I would have asked the learner for the remaining cards.

(2019) predict that overgeneralised forms serve as placeholders for the particle 'not', the suffix *-ed* in TP, and the suffix *-s* in AgrP, based on the difficulty of these morphemes. The difficulty of 'not' arises from its requirement of the auxiliary 'do'; *-ed* is complicated by the presence of regular and irregular past tense forms, and *-s* is challenging due to its weak paradigmatic status and ambiguous function (indicating plural, possessive, or agreement). Therefore, the comprehension tasks assess the acquisition of NegP, free morphemes (i.e., 'is' and 'are'), and bound morphemes (i.e., third-person singular *-s* and past simple regular *-ed*) in relation to the use of overgeneralised forms. Specifically, I will examine whether the difficulty of these morphemes correlates with the use of overgeneralised forms. These correlations help us determine whether these forms are placeholders for verbal bound morphemes, as Vainikka and Young-Scholten (2019) and Vainikka *et al.* (2017) claim.

4.3.4.1 Design of audio–picture match questions and procedure

The research design for the comprehension tasks followed three steps. First, I sought help from a professional illustrator to create sets of pictures reflecting the contexts/situations (see Appendix B for all these sets of pictures). Second, I asked a native speaker to record the stimuli (utterances) that target the actions. Third, using BookWidgets[2] (https://www.bookwidgets.com), I created an online auto-pictures-matches quiz/test. BookWidgets conveniently works on iPhone and Android (or similar devices), accommodating participants who might not have access to laptops or desktop computers. To instruct the learners on how to perform the test, I created a two-audio-picture-match question.

In terms of the procedure, immediately after collecting oral production data, each learner received a link to audio–picture match questions on the BookWidgets website. In this link, I explained to the learners how to take the test. As shown in Figure 4.4, for example, the learner was instructed to hear the stimulus (e.g., a book as shown in the figure), match it with the target picture by dragging, and finally type their name and submit their answers, which were sent to my email. After that, each learner received a link to a seven-audio–picture match question and was kindly asked to finish the test as

[2]The intended procedure was that each learner would be individually tested in a vacant (or quiet) room in the centre where they attend classes. The tasks would be presented to them on a laptop, using a software programme called E-prime 2.0. According to the task, each learner would be instructed to match the image with the stimulus sentence that he or she will hear through the laptop's speaker.

soon as they could. The audio–picture match questions reflect the context for each of the following morpho-syntactic elements: NegP, copula 'be', auxiliary 'be', regular past simple *-ed*, and third-person singular *-s*.

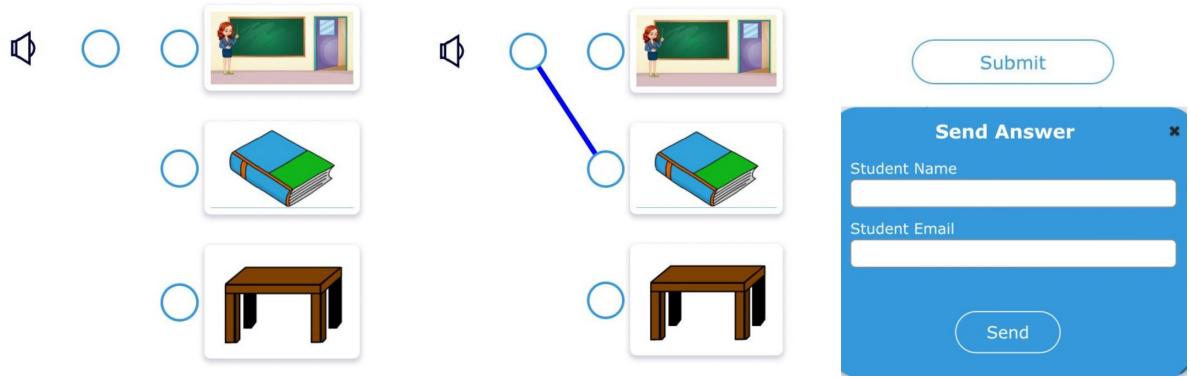


Figure 4.4: Question 2

4.3.4.2 NegP

With respect to NegP, the participants saw three sets of pictures, each with three images. For instance, the first picture showed a boy playing games on a PlayStation, the second picture showed a boy sitting next to the PlayStation with the TV turned off (indicating that the boy does not play PlayStation; see Figure 4.5), and the third picture showed a boy and a girl playing games on a PlayStation as a distraction. At the same time as the pictures were shown, the participants heard a negative sentence (e.g., 'He doesn't play PlayStation'). The learners' task was to choose the image that matched the sentence they heard.

Figure 4.5: NegP

4.3.4.3 Copula be

For the copula 'is', the learners were shown three sets of pictures, each containing three images. For example, the first picture depicted two For the copula 'is', the learners were

shown three sets of pictures, each containing three images. For example, the first picture depicted two police officers, the second picture only had one police officer, and the third picture showed two engineers as a distraction (see Figure 4.6). The purpose of this task was to test the learners' comprehension of the copula 'is' as it relates to the subject of the sentence. The learners listened to a subject-drop sentence that depicted only one police officer and began with a pause (e.g., '...is a police officer'). Their task was to choose the picture that showed one police officer. If a learner selected the picture with two police officers, it indicated that he or she did not comprehend that the copula 'is' was referring to the singular subject.

Figure 4.6: Copula 'is'

For the copula 'are', the learners were shown three sets of pictures, each including three images. For example, the first image depicted a single barber, the second depicted two barbers, and the third depicted an empty barbershop (see Figure 4.7). As these pictures were displayed, the learners listened to a subject-drop sentence beginning with a pause (e.g., '...are barbers'). The learners' task was to select the picture that showed two barbers. If they chose the image with one barber or the empty barbershop, it indicated that they did not comprehend that the copula 'are' agrees with the plural subject.

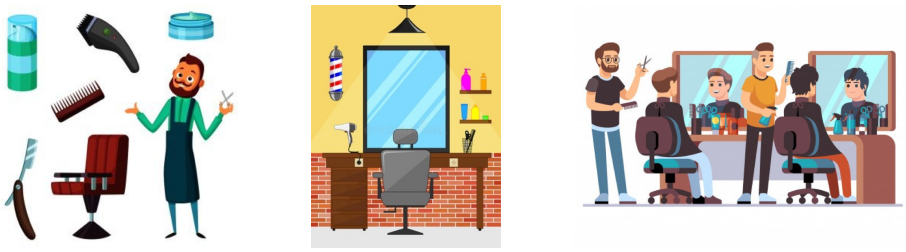


Figure 4.7: Copula 'are'

4.3.4.4 Auxiliary be

For the auxiliary 'is', the learners were shown three sets of pictures, each containing three images. For example, the first picture depicted a boy eating breakfast, the second showed a boy holding an empty dish (indicating that he had finished his breakfast), and the third depicted four girls eating their breakfast (see Figure 4.8). While these images were displayed, the learners listened to a sentence describing only one person eating breakfast that began with a pause (e.g., '...is eating breakfast'). The learners were expected to select the picture that best matched the sentence they heard.

Figure 4.8: Auxiliary 'is'

For the auxiliary 'are', the learners were shown three sets of pictures. For instance, the first picture depicted four girls washing a dog, the second picture showed one girl washing a dog, and the third picture depicted two girls not washing a dog (see Figure 4.9).

Figure 4.9: Auxiliary 'are'

4.3.4.5 Regular past simple -ed

To test comprehension of the past simple *-ed* (e.g., kicked), the learners were shown three sets of pictures, each containing three images. For instance, the first picture depicted the beginning of an action, the second showed the action in progress, and the third displayed the completed action (see Figure 4.10). As a stimulus, the learners listened to a sentence

describing the picture showing the end of the action (e.g., 'He kicked the ball'). The learners' task was to select the picture that best matched the sentence they heard. I expected the learners to choose the picture depicting the end of the action. If they did not, it would indicate that the learners did not comprehend the morpheme *-ed*.

Figure 4.10: Regular past simple *-ed*

4.3.4.6 Third-person singular -s

Similar to the tasks used for the copula and auxiliary 'be', I employed aural subject-drop sentences alongside visual sets of pictures to test comprehension of the third-person verbal agreement *-s*. For example, the learners listened to the sentence fragment '... studies in the library'. The first picture depicted a man reading in the library, the second showed two men reading in the library, and the third contained only a study room, used as a distractor (see Figure 4.11). The learners were then expected to choose the target picture based on what they heard. If a learner did not select the picture that best matched the audible sentence, it indicated that he or she did not perceive the morpheme *-s* on the main verb.

Figure 4.11: Third-person singular -s

4.4 Research questions and research hypotheses

4.4.1 Low-educated learners and the developmental path of acquisition under Organic Grammar

4.4.1.1 Research questions

1. Do adult immigrants with varying levels of native language schooling/literacy follow the same path of acquisition?

4.4.1.2 Research hypotheses

To answer the research question, three hypotheses were formulated based on Organic Grammar. As detailed in Chapter 2, Organic Grammar predicts that language acquisition occurs in successive stages, beginning with the bare L1 VP stage, followed by the bare L2 VP stage, then the NegP stage, intermediate stages (TP and AgrP), and finally the CP stage.

1. The bare VP will be projected first, followed by NegP. The bare VP can follow an L1 or L2 order pattern and lacks IP-related morphology. These VP order patterns and the absence of IP-related morphology constitute the VP stage. The NegP does not lack these elements.
2. Learners will then project TP, followed by AgrP. At the TP stage, learners demonstrate their acquisition of the NegP through consistent use of Neg constructions (*do*-forms), the copula 'be', and some auxiliary 'be'. At the AgrP stage, learners further demonstrate their acquisition of the NegP by consistently using Neg constructions (*do*-forms), the copula 'be', and regular/irregular past tense.
3. Literacy does not affect the general path of acquisition, but learners with lower levels of literacy may experience delayed progression through the stages. This is based on the observed relationship between lower levels of literacy and lower morphosyntactic levels (Tarone *et al.* 2009; Vainikka *et al.* 2017; Young-Scholten and Strom 2005).

4.4.2 Overgeneralisation by low-educated learners

4.4.2.1 Research questions

1. Do the stages of Organic Grammar influence the choice of overgeneralised forms? If so, how?
2. What is the nature of overgeneralisation by low-educated L2 learners?

4.4.2.2 Research hypotheses

1. Overgeneralised forms decrease as learners progress through the developmental path of acquisition.
2. Overgeneralised forms occur after the establishment of the bare VP.
3. Learners who use overgeneralised forms struggle with the comprehension of the regular past tense *-ed* (the head of TP) or the third-person singular *-s* (the head of AgrP).

These hypotheses are based on the idea that overgeneralised forms are placeholders for the relevant syntactic head at the functional projections (Vainikka and Young-Scholten 2019; Vainikka *et al.* 2017). They are also based on the claim that dummy auxiliaries (Julien *et al.* 2016) or interlanguage constructions (Mocciaro 2019b) are used before the acquisition of finiteness.

4.5 Data analysis and counting procedure

This section presents the research methods used in this thesis to answer the research questions. This includes the idea of implicational analysis used to determine developmental stages, followed by the counting procedure and then the statistical analysis adopted in this thesis.

4.5.1 Cross-sectional study and implicational analysis

Researchers who adopt developmental stages see that the optimal way to identify stages is through a longitudinal study. This is because longitudinal studies analyse changes that occur over time, allowing researchers to identify and define different stages in terms of acquired rules based on the systematic appearance of linguistic features. However,

longitudinal studies can present some limitations to the research findings. This is because longitudinal studies often have a limited number of participants. Cross-sectional studies, which use larger samples, typically support the research's findings. However, cross-sectional studies have more drawbacks compared to longitudinal studies. The typical approach in such studies involves investigating the linguistic performance of the learner at a certain point in time. It can be quite challenging to draw conclusions and specify stages based solely on on a single point in language development.

As briefly mentioned in Chapter 2, Rosansky (1976) questioned the validity of cross-sectional studies of the order of morphemes in L2 that support Brown's (1973) longitudinal study. Rosansky conducted a longitudinal study supported by a cross-sectional study. The purpose of the study was to investigate whether the cross-sectional and longitudinal findings are comparable. The longitudinal study included two children, two adolescents, and two adults. The study began approximately three months after their arrival in the United States, and the cross-sectional study was conducted after ten months of exposure, including five participants, as one of them returned to his home country after nine months, contradicting the intended plan. The findings indicated that the order of the morphemes appears to vary on a monthly basis.

Addressing Rosansky's (1976) criticism of cross-sectional studies, Andersen (1978) offered a more comprehensive model including implicational analysis/table, a common practice of second-language research dealing with the systematic and variability of the data. Implicational analysis is primarily employed in the field of sociolinguistics, where it involves the investigation of the relationship between specific linguistic features and individual or group speakers. This correlation implies that if a person uses a specific feature in their speech, it is likely that they possess other related attributes as well. For example, individuals who use nonstandard variations of English may also employ standard English forms in certain contexts. By studying the contexts in which these variable rules operate, researchers can explore grammatical assessments regarding the implementation of these rules. An implicational table is commonly used to illustrate the link between attributes and speakers. In this table, the presence of an attribute is represented by '1', and the absence of an attribute is represented by '0'. Speakers can be categorised or listed based on the quantity of qualities they employ. Within the realm of second language acquisition, these characteristics pertain to the assimilation of

particular rules and their accurate application in diverse linguistic contexts. Table 4.3 presents an implicational analysis for four grammatical morphemes, which were analysed in Andersen's (1978) research paper: copula 'be' (COP), auxiliary 'be' used in progressive constructions (AUX), irregular past verb form such as 'bought', 'saw' (PaI), and auxiliary 'have' in perfect verb constructions (Hv).

Level	Morphemes				Number of subjects at each level
	Cop	Aux	PaI	Hv	
1	1	1	1	1	10
2	1	1	1	0	25
3	1	1	0	0	28
4	1	0	0	0	24
4	0	0	0	0	2

Table 4.3: Implicational table for the sequence Cop, Aux, PaI, and Hv, modified table from Andersen (ibid.)

If the learners produced the morpheme 80% or more of the time, '1' represents it in this table. This indicates that the morpheme was acquired. On the other hand, it is represented by '0' if the learner produced it 79% or less of the time, indicating that the morpheme was not acquired.

Like Rosansky, Meisel, Clahsen, and Pienemann (1981) ask whether one can draw conclusions and specify stages based solely on a cross-sectional study. Similar to Rosansky's (1976) argument, the authors argue that fluctuation in longitudinal studies casts doubt on the validity of cross-sectional studies with respect to the developmental stages. Referring to Molony's (1977: 279) finding that 'a dramatic increase in verb errors from three and a half months (March) to five months, then a steady decline so that by October there are almost no errors', Meisel *et al.* clarify that cross-sectional studies cannot detect such variations because linguistic factors, such as structural contexts, and non-linguistic factors could be linked to this development. They suggested that any indicators of new stages of development should be represented on an implicational scale. This occurrence is a natural consequence of the underlying assumption in most language acquisition research, which holds that specific rules are acquired in a predetermined order, as outlined in (4.4). These rules should be tested in longitudinal studies.

(4.4) $R_n \supset R_{n-1} \supset \cdots \supset R_1$

That is, if a learner has acquired rule R_n, they have also acquired R_{n-1} and so on, until R. However, the learner who has acquired R_{n-1} may not have acquired R_n. The

argument in favour of implicational ordering is not only that it aligns with the common practice of language acquisition research, but also because, as convincingly argued by Bailey (1976, 1977), it presents a psychologically plausible hypothesis about what can be learned. In other words, if we observe that all L2 learners who have acquired rule R_3 also possess the rules R_2 and R_1, but those who have not yet acquired R_2 also do not use R_3, then we can assume that the three rules are ordered as shown in (4.5).

(4.5) $R_3 \supset R_2 \supset R_1$

Meisel *et al.* assume that each rule indicates a new developmental stage, but these should be tested through longitudinal studies. Therefore, researchers who adopt developmental stages of acquisition view language acquisition as a system of rules acquired in a specific order.

Vainikka and Young-Scholten (1994) employed this implicational ordering in their cross-sectional study, as discussed in Chapter 2. For example, the implicational ordering for the acquisition of the agreement paradigm is as follows:

(4.6) Agreement → Verb Raising → No Null Subjects → Head-final VP

That is, if a speaker has acquired the agreement paradigm, they have also acquired verb-raising, non-pro-drop, and a head-final VP. However, acquiring verb-raising does not necessarily mean that they have acquired the agreement paradigm. Verb-raising precedes the agreement paradigm, indicating that raised verbs may not always exhibit agreement markings. Through this implicational ordering, Vainikka and Young-Scholten argue that their cross-sectional data can reveal the predicted OG stages, but they also underscore the importance of including longitudinal data.

As mentioned in Chapter 2, Vainikka and Young-Scholten used a cut-off point of 60% as a criterion to determine whether a particular construction has been acquired. This 60%, according to Vainikka and Young-Scholten, fits their data sample better than higher percentages (e.g., 80%). For example, if they had used 80%, they would have had very few learners who would have acquired verb-raising in German. They also note that the 60% cut-off point distinguishes between learners who raised the verb from those who did not. Similarly, this study uses an implicational analysis method. However, instead of using percentages, I followed Chen (2017), who used three target-like productions in a longitudinal study as a criterion instead of percentages. This criterion fits my data

sample, as if I had used more target-like productions, I would have had very few learners who acquired TP. This method also distinguishes between learners who prefer to use bound morphemes (the regular past tense *-ed* or the third-person singular *-s*) and those who do not, as pointed out by Vainikka and Young-Scholten. This criterion could reveal the early stages of functional projections (TP and AgrP), as Chen pointed out.

4.5.2 Counting procedure

The dataset obtained from the picture description tasks included 60 audio tracks. The length of these files ranged from 30 to 45 minutes. Some learners exceeded the 45-minute limit. Following the methodology of some researchers (e.g., Chen 2017; Kahoul 2014; Vainikka and Young-Scholten 1994), these data were transcribed orthographically. In line with Unsworth (2005), instances where participants immediately corrected themselves were seen as reflective of their intended utterances. Therefore, these self-corrections were included in the data rather than in the original utterances or forms.

Utterances were not counted if they only contained a proper name, as in (4.7a), a single noun, as in (4.7b), or a verb describing what was happening in a picture in one word, as in (4.7c).

(4.7) a. Sam.

Target utterance: Sam cleans the room.

b. Woman

Target utterance: The women reads newspaper.

c. Draw.

Target utterance: The man draws a horse.

Error-free utterances were those that did not include morphological, syntactic, or lexical errors. Morphological errors included, as an example, plural marking, as in (4.8a). As with Vainikka *et al.* (2017), the use of the past simple marker *-ed* and the third-person singular *-s* was included regardless of the purpose of the intended task. For example, learners used the past simple *-ed* in a task that required progressive actions, as in (4.8b).

(4.8) a. Two girl are washing the dish.

Target utterance: Two girls are washing the dishes.

b. Sarah washed the car.

Target utterance: Sarah is washing the car.

Examples of syntactic errors included non-target-like word order, as shown in (4.9a), the absence of functional components, such as determiners, as shown in (4.9b), and the copula and auxiliary 'be', as illustrated in (4.9c) and (4.9d), respectively.

(4.9) a. Woman breakfast eat.[3]

Target utterance: The lady eats breakfast.

b. Girl drink coke.

Target utterance: The girl drinks coke.

c. He doctor.

Target utterance: He is a doctor.

d. The girls washing the dog.

Target utterance: The girls are washing the dog.

The instance in (4.9b) clearly shows that the VP has been established, and the aim of the task is to determine whether the bare VP has been acquired, regardless of the omission of the determiner 'the'. Lexical errors included non-target-like nouns, as in (4.10a), and non-target-like verbs, as in (4.10b).

(4.10) a. Faris and Abdul write homework.

Target utterance: Faris and Abdul write letters.

b. Girl write horse.

Target utterance: The girl draws a horse.

Notably, non-discrimination between the verb 'clean' and 'wash' is common among all participants in this thesis. The counting of verbs included only the target-like copula and auxiliary 'be'. For example, the non-target auxiliary 'be', shown in (4.11), was excluded as it does not show agreement. In contrast, lexical non-target-like verbs were counted. As mentioned above, the aim is to determine whether the bare VP has been established.

(4.11) Two boy is writing letter.

Target utterance: Two boys are writing letters.

[3]This is very rare

4.5.3 Statistical analysis

Following counting, raw production and comprehension data were entered into Microsoft Excel and then imported into SPSS for statistical analysis. The production sheets include names, OG stages, reading levels (see learners' OG placement and their reading levels in the following chapter), years of native schooling, length of UK residence in months, ESOL instruction in months, age, and gender. The sheet also includes their raw scores for each morphosyntactic element in question.

In order to investigate the relationship between OG stages and literacy levels, this study used Spearman's rank correlation, which measures the strength and direction of the association between two ordinal variables. Both OG stages and reading levels (literacy) are ranked/ordinal variables. The OG stages (VP, NegP, TP, and AgrP) were first assigned numerical values of 1, 2, 3, and 4, respectively, and the reading levels (1, 2, 3, and 4) were also assigned numerical values of 1, 2, 3, and 4, respectively. These numerical values were then entered into SPSS, which was used to calculate Spearman's rank correlation and measure the relationships between these ranks automatically.

The assumptions for Spearman's rank correlation include the variables being measured on an ordinal scale, independent observations, and a monotonic relationship between the variables. However, unlike Pearson's correlation, Spearman's is a non-parametric test, which places no strict assumptions apart from the data being ordinal or ranked (see, e.g., Kvam, Vidakovic, and Kim 2022). The assumption of independent observations has been taken into account, as each learner's data are independent of the data from other learners, and Spearman treats these data as pairs of ranks that are independent.

Additionally, ordinal regression was conducted to investigate whether other factors, as independent variables, affect the relationship between literacy levels and OG stages as a dependent variable. Ordinal regression is particularly suitable for this analysis because it allows for the investigation of multiple independent variables (whether ordinal, continuous, or categorical) on an ordinal dependent variable, such as OG stages. This method helps detect any contributions of each non-linguistic factor to the variability observed in OG stages in relation to literacy levels, providing a better understanding of the relationship.

In this test, OG stages are treated as categorical dependent variables because they represent distinct categories or levels (e.g., VP, NegP, TP, and AgrP assigned values of

1, 2, 3, and 4, respectively). The other variables, such as the amount of native language schooling, ESOL instruction, and age, are considered independent variables. The key assumption of ordinal regression, built in SPSS, was checked through the Test of Parallel Lines. The p-value for the Chi-Square test is 0.304, indicating no significant violation, making the test applicable to examine the relationship between OG stages as the dependent variable and the other independent variables.

To compare learners' performance in negation and verbal morphemes with their OG stages, I first investigated the distribution of the verbal morphemes separately from their OG stages. This provided a general sense of their performance and helped detect any consistency or variability in their comprehension of these morphemes. This initial assessment provides a better understanding when comparing their performance according to their OG stages.

This investigation was first carried out using descriptive statistics, particularly the mean and the SD, to observe the overall trends and variability by comparing the mean with the SD. The Friedman test, using RStudio as its subsequent Dunn post hoc test, was employed because Dunn is not directly built into SPSS. The Friedman test was used to compare the differences between these morphemes as an equivalent to the Kruskal-Wallis test when the data fail the assumption of independent observations (as is the case with my data). For example, each learner in my study was assessed on multiple morphemes, meaning that the observations are dependent on each other and thus violate the assumptions of the Kruskal-Wallis test. The Dunn post hoc test was used to determine which verbal morphemes differ from each other. It serves as an equivalent to the Tukey post hoc test following a one-way ANOVA, which requires the data to be normally distributed, which is not the case in my data.

To compare learners' performance in comprehension with their OG stages, I used descriptive statistics, focusing on the mean and SD, but this time in relation to learners' OG stages. I also used data visualisation (scatter plots) to observe any potential relationship between the two variables in question. Spearman's correlation was used to determine whether there is a relationship between learners' scores in comprehension and their OG stages assigned based on the production data.

In addition to the main objectives, the study also investigated the idea of triggers through production data and comprehension data. Although this was not the main

focus, it is still crucial for better understanding the relationship between literacy and the development of morphosyntax. The focus here was to first investigate the acquisition of verbal morphemes in relation to the acquisition of the copula 'be' as a potential trigger and then the acquisition of the copula 'be' in production as well as comprehension in relation to literacy. Similarly, the investigations included descriptive analysis (mean and SD) followed by a visual presentation (scatter plots) and a correlation test (again Spearman). The data here fail the main assumptions of linearity in Pearson's correlation, thus Spearman's correlation was used. Learners' scores were treated as ranked data, and the nature of the study aligns with independent observations where each learner's data are independent. As previously mentioned, Spearman's correlation treats each learner's data as independent, measuring the strength of the relationship between each pair of ranks as independent observations.

As stated in this thesis, this study aimed to investigate the relationship between learners' overgeneralisation and the acquisition of verbal morphemes through comprehension tasks. More specifically, the study aimed to determine whether learners' difficulty in comprehending bound morphemes (*-ed* and *-s*) results in using overgeneralised forms as placeholders for these morphemes during the acquisition of morphosyntax, as Vainikka and Young-Scholten (2019) claims and as expected in one of the hypotheses in this thesis. Therefore, investigations included analyzing the relationship between the mean values and the SD for the occurrence of these forms in relation to bound morphemes. Additionally, the study explored potential relationships through scatter plots followed by Spearman's correlation. Similarly, the investigations examined any potential relationship between the acquisition of free morphemes (the copula 'be' and the auxiliary 'be') to gain a better understanding of learners' use of overgeneralisations and the development of morphosyntax in general. Again, the data variables fail to meet the linearity assumptions of Pearson's correlation when plotting these variables on scatter plots, as we will see in the results chapter. Therefore, Spearman's correlation is an appropriate choice to measure the strength of these ranks, which constitute independent observations, thereby following the assumptions of independence.

4.6 Summary

This chapter is divided into four sections. Section 4.2 presents descriptive statistics for 60 learners with various external linguistic variables, including (1) the learners' age when tested, (2) the length of native schooling prior to arriving in the UK, (3) the length of L2 English instruction (i.e., ESOL classes), and (4) the length of UK residency. None of the learners finished their schooling in their native country. Section 4.3 describes the data collection procedure, which includes oral production and comprehension tasks. For the production tasks, participants were presented with a set of picture sheets and asked to describe orally what they saw. For the comprehension tasks, learners were assessed using audio-picture matching questions designed to test the claim made by Vainikka and Young-Scholten (2019) that failure to identify the relevant syntactic head results in the use of non-target, overgeneralised forms. Section 4.5 presents the data analysis methods, including the implicational analysis used to determine developmental stages, and the statistical analysis adopted in this thesis.

Chapter 5. Results

5.1 Introduction

As noted in this thesis, the first aim is to examine the role of literacy in adult immigrant learners' development of English morphosyntax using the stages of Organic Grammar (specifically, the role of literacy in the developmental path of acquisition). Therefore, Section 5.2 introduces the results of the reading levels. Section 5.3 will present the OG-stage placement based on oral production for all 60 learners. In line with my hypothesis that low-educated learners will follow the predictable path of acquisition under Organic Grammar, assigning each learner an OG stage will help us identify learners' acquisition trajectories to investigate this hypothesis. This section will also present the results of the statistical analysis using Spearman's correlation coefficient, which was used to investigate whether there is a correlation between OG stages and literacy, followed by an investigation using ordinal regression to test my hypothesis that lower literacy levels negatively affect the progression between OG stages.

To look for more evidence supporting my hypothesis that low-educated learners follow the developmental path of acquisition under OG, Section 5.4 will investigate learners' comprehension of verbal morphemes. In this section, I will first examine the difference between verbal morphemes through descriptive analysis including the mean and SD, followed by a Friedman test to determine if there is a significant difference between these morphemes. Dunn's test will be used as a post hoc test to identify which morphemes differ significantly from each other. Following this, I will investigate these morphemes in relation to learners' OG stages again through the mean and SD, including scatter plots to indicate the relationships, followed by Spearman's correlation to measure the strength of the relationships.

In Section 5.5, I will examine the idea of triggers in the development of morphosyntax. The focus will be on examining the potential relationship between the acquisition of the

copula 'be' as a potential trigger and the verbal morphemes, including the acquisition of negation, both in production and comprehension. Then, I will investigate the relationship between literacy and the copula 'be', again in both production and comprehension. This investigation will help us determine whether there is an indirect effect of literacy through the copula 'be' as a trigger on the development of morphosyntax (OG stages).

Similarly, the analysis in this section includes an investigation of the mean values of these morphemes in relation to the copula 'be' and examining the variability through SD, followed by scatter plots for depicting the relationships and then Spearman's correlation to measure the strength of these relationships. This approach will also be applied to the investigation between literacy levels and the copula 'be', but the mean values of the copula 'be' will be assessed in relation to literacy levels.

Section 5.6 will present the results of learners' overgeneralisations in oral production data, focusing on the OG-stage placement to pinpoint the exact stage of overgeneralisation. This will help investigate my two hypotheses: that overgeneralisation occurs after the establishment of the bare VP and that these forms precede bare verbs. This section will also include an investigation of the relationship between learners' overgeneralisations and verbal morphemes. This includes the investigation of each type of overgeneralisation in relation to each verbal morpheme (free or bound morphemes). Similarly, the investigations include a comparison of the mean values of each type of overgeneralisation to each morpheme, including the SD to assess the variability. Scatter plots and Spearman's correlation will also be used to visualise the relationships and assess the correlations, respectively. In this approach, I tested my hypothesis that learners who use overgeneralised forms face difficulty in the comprehension of bound morphemes.

5.2 Reading Results

Based on learners' results on the five reading sub-tests for English provided earlier in the previous chapter (particularly in Section 4.3.2, Table 4.2), I assigned each learner to one of the levels or implicational stages of reading displayed in Table 5.1. All learners performed well on varied single-letter identification; scores below 100% were attributable to 'p' and 'b' or 'v' and 'f' phonology-based non-discrimination. The survival/environmental sign task had the second-highest score, followed by paragraph reading and decoding. Learners' reading/literacy levels are presented below in Section 5.3.

Level	Varied single letter identification	Survival signs	Paragraph reading	Decoding of familiar words in isolation
1	75%+	25%+	no ability	0%
2	75%+	75%+	attempt, w/guessing	20%+
3	100%	100%	slow, sometimes accurate	20%+
4	100%	100%	halting, mostly accurate	60%+
5	100%	100%	fluent	100%

Table 5.1: Reading level scoring, based on % correct on English reading tests

Table 5.2 presents the number of learners categorised by reading levels along with their corresponding percentages. Level 1 includes 14 learners, representing 23% of the total. Level 2 consists of 13 participants, accounting for 22%. Level 3 comprises 21 participants, constituting 35%, while Level 4 includes 12 learners, making up 20% of the total group.

Reading levels	n	Proportion
1	14	23%
2	13	22%
3	21	35%
4	12	20%

Table 5.2: Reading levels: Descriptive statistics

5.3 Production data

5.3.1 Data analysis

Along with Vainikka *et al.* (2017), this thesis follows Scarborough's (1990) measure of productivity, which is signaled by the learners' supply of various morphemes (bound or free) in connection with the relevant syntax. To assign a syntactic stage to each learner, I analysed the data according to Organic Grammar. As detailed in Chapter 2, Organic Grammar posits that the development of morphosyntax occurs through a specific path of stages (VP, NegP, AgrP, and finally CP), as shown in Table 5.3.

Assigning each learner a stage of Organic Grammar helps us investigate various areas. First, it allows us to explore the path of language development proposed by Organic Grammar. Second, it assists in determining whether this path is fixed among all 60 learners. Third, it enables us to pinpoint the stage at which learners use overgeneralised

forms.

Stage	word order	Verb types	agreement/tense	pronouns	syntax
VP	L1 order, then L2 order	thematic (main) verbs	none	subject, object pronouns absent	none
NegP	resembles the L2 apart from complex syntax	thematic verbs; copula 'is'	none	pronouns forms begin to emerge	Negation; single clauses; formulaic or intonation-based Qs.
TP	resembles the L2 apart from complex syntax	thematic verbs, modals; copula forms beyond 'is'	no agreement; some tense, some aspect, but not productive	more pronoun forms, but they can still be missing	Conjoined clauses. Formulaic wh-Qs; yes/no Qs w/o inversion.
AgrP	resembles the L2 apart from complex syntax	thematic verbs, modals, copula forms beyond 'is'; auxiliaries in all forms and tenses	productive tense, aspect; some agreement, esp. forms of 'be'	pronouns obligatory, 'there' and existential 'it'	Simple subordination; wh-Qs but all Qs may lack inversion
CP	always resembles the L2 thematic	complex tense, aspect forms; passives; range of thematic verb, modal, auxiliary forms	forms usually correct, apart from newly attempted ones	use of 'there' and 'it' beyond stock phrases	Complex subordination. All Qs with inversion.

Table 5.3: Organic Grammar stages for L2 English

As indicated in Chapter 1, this study builds on Vainikka *et al.*'s (2017) small-scale study by including a larger sample and conducting a more detailed analysis. As discussed in the previous chapter, an implicational analysis approach was adopted to track the learners' trajectories. Through this analysis, I will investigate whether the learners follow the stages proposed by Organic Grammar, considering both preceding and subsequent stages. Unlike Vainikka *et al.*, who did not apply their implicational analysis as extensively as they did in their significant work on L2 German or follow specific criteria to assign OG stages to learners, it is possible that this decision was influenced by the low production rate of inflectional morphology observed in their learners, although this was not explicitly stated in their paper. Vainikka *et al.* assign stages if learners are actively working on a given phrase, attempting to figure out how English negation, tense, or agreement is marked morphologically and represented. Additionally, this study utilised Déprez and Pierce's (1993) predictions regarding word order in negative constructions in L1 English to determine whether the learners' NegP is within or outside the VP, particularly in the early stages of acquisition.

This section will discuss the order of morphemes while assigning the stages. Indeed, Organic Grammar considers the order of morphemes a significant factor. By studying the consistent patterns and variations in morpheme order, we can determine whether the morpheme order in this study aligns with the established order of morphemes documented since the 1970s.

5.3.2 The VP stage

As discussed in Chapter 2, all hypotheses about the initial state of adult second language acquisition (such as Minimal Trees or Organic Grammar, Full Transfer/Full Access, and Modulated Structure Building) assume at least an initial transfer of lexical items from L1. This suggests that if the verb follows its complement in the L1 VP, this pattern will be transferred to the L2 VP. When the learner is exposed to enough L2 input, they will shift in favor of the L2 pattern. Following Vainikka and Young-Scholten (1994, 1996a), 13 learners in my sample were placed in the VP stage based on their production of only a bare verb, which is a verb without any other functional projection (as indicated by the absence of associated verbal morphology such as copula 'be', auxiliaries, modals, past tense *-ed*, or third-person agreement *-s*). As mentioned in Chapter 4, the word order in Arabic can be subject-verb-object (SVO) or verb-subject-object (VSO); thus, Arabic has a head-initial VP, like English. The results of the current study show that there are two learners who transfer their Arabic word order (VSO) as in (5.1), meaning they produce VO rather than OV.

(5.1) a. Drink boy cola.

Target utterance: The boy drinks/is drinking cola.

b. Baby car drive.

Target utterance: The baby drives/is driving a car.

Name	Reading level	SVO	VSO
Mansoor	1	8/8	0/8
Fahad	1	7/8	0/8
Arshad	1	2/8	0/8
Hadeel	1	3/8	7/8
Gharam	1	8/8	0/8
Adeeb	1	3/8	0/8
Hoor	2	7/8	0/8
Balsam	2	6/8	0/8
Omar	1	8/8	0/8
Bilal	1	2/8	8/8
Andleeba	2	8/8	0/8
Susan	2	6/8	0/8
Andleebid	1	6/8	0/8
Nizar	2	5/8	0/8

Table 5.4: The VP stage

Five of these learners used 'no' with thematic verbs (e.g., 'Man no open door'). As predicted, the use of 'no' perceeds the verb as in (5.2). This early use of negation reflects the first stage of Cancino *et al.*'s (1978) stages of development in the acquisition of English negation as presented in Chapter 2.

(5.2) Boy no play football.

Again, this use of 'no' does not show evidence of progress beyond the VP. This OG placement aligns with Vainikka and Young-Scholten, who reported that the data from VP-stage learners (Bongiovanni and Salvatore) show evidence of negation. However, because these learners did not demonstrate that they had acquired verb raising in their L2 German, they were placed in the VP stage. This lack of verb raising is indicated by the position of the negation marker, which precedes the main verbs, while in L1 German, the negation marker follows the main verbs. Vainikka and Young-Scholten argued that such negative morphology occurs within the VP. The researchers also reported that, at this stage, Bongiovanni and Salvatore produced no models, auxiliaries, tense, or subject-verb agreement. My participants' early negation is comparable to that of L1 English data in terms of whether the negative morphology occurs in the VP or outside it. As shown in Table 5.5 in the following section, Déprez and Pierce (1993) proposed that constructions such as (S Neg V) do not reveal a V-to-I movement; that is, they occur within the VP.

Under Organic Grammar, the copula 'be' and the auxiliary 'be' are base-generated in INFL and are unexpectedly produced in the VP stage. However, Vainikka and Young-Scholten (1996b) reported similar examples from two VP-stage learners (Jose and Rosalinda) who produced many copulas (exclusively *is(t)* 'is', apart from one instance of *bin* 'am') produced by Jose. Similarly, the analysis in this thesis placed three learners (Mansoor, Arshad, and Andleeba) in the VP stage who used the copula 'be' five times in their attempts. This is because (1) the use of the copula 'be' emerged as a default form (e.g., 'they is student'), and (2) the use of the copula 'is' followed pronominal subjects, which suggests that rote learning is involved. Both reasons indicate that the copula 'is' is unanalysed. This use of the copula 'be' does not give any indication that they have fully acquired the agreement paradigm for the copula 'be' at this stage. A similar but different instance of such rote learning is seen with Nizar's data, who overused the phrase 'This is' in all tasks as in (5.3).

(5.3) a. *This is the* woman drink cola. (In the VP task) (Nizar)
Target utterance: The girl drinks coke.
b. *This is the* man no drive car. (In the NegP task)
Target utterance: The boy doesn't drive a car.
c. *This is the* girl read book. (In the progressive tense task)
Target utterance: The girl drinks coke.
d. *This is the* doctor. (In the copula 'be' task)
Target utterance: He is a Doctor.

Regardless of the presence of the phrase 'This is', these utterances demonstrate the structure of the bare VP of Organic Grammar; therefore, they were included when considering the acquisition of the bare VP. However, when considering the acquisition of the copula 'be', these utterances were excluded from the count, even though they occurred in contexts where the copula 'be' was obligatory, as in (5.3d). This exclusion is because the use of 'This is the' remains a chunk as they overgeneralised in a progressive tense task, as in (5.4a), and the copula 'be' task, as in (5.4b).

(5.4) a. *This is the* boy clean dish . (In the progressive tense task)
Target utterance: The boys are cleaning the dishes.
b. *This is the* doctor. (In the VP task)
Target utterance: They are doctors.

These are similar to what N. C. Ellis (2002) refers to as 'teddy bears'. As we shall see later, these formulaic language utterances differ from the overgeneralised forms that this study investigates.

The table hides the fact that there are some learners who sometimes produce utterances with null subjects but include thematic verbs, as in (5.5).

(5.5) Write letter.
Target utternace: The boy is writing a letter.

As stated in the previous chapter, such instances indicate a bare VP, as they contain a thematic verb and at least one other VP-internal constituent. Under Organic Grammar, there is a direct relationship between null subjects and bare verbs. The occurrence of null subjects varies depending on the verbal prompt of the task. For example, learners

produce fewer utterances with null subjects in the VP task, since the prompt contains the subject of the sentence, as in (5.6a). This is because learners can quickly recall the subject. Conversely, learners produce more utterances with a null subject if the prompt includes a subject following adverbs of time, as in (5.6b), which makes recalling the subject more challenging. This analysis does not intend to criticise the methods used in this study but rather to report on them. The methods are primarily designed to test the acquisition of verbal morphemes in relation to syntax while also limiting the use of empty subjects, which negatively correlate with agreement in the present progressive tense and copula 'be', as well as with agreements on the main verbs for the third-person. The prompts are in square brackets [...].

(5.6) a. Researcher: [The girl...] (In the VP task)
Participant: Girl drink coke.
b. Researcher: [Every morning, this lady...] (In the VP task)
Participant: Every morning drink coffee.

The data show that some learners at this stage used single words (either a verb or a complement) as code-switching elements. According to the counting procedures, these instances are treated as non-target-like and excluded from the count. However, they indicate the structure of the bare VP since they still represent the L1 word order due to L1 transfer, as in (5.7a), or the order of L2 words, as in (5.7b). Conversely, some learners use code-switching followed by self-correction, as in (5.7c). These corrected utterances were treated as target-like and included in the count.

(5.7) a. Yarsim (draw) boy horse. (In the VP task)
Target utterance: The boy draws a horse.
b. Girl tarmi (throw) shoes. (In the VP task)
Target utterance: The girl throws the shoes.
c. Baby yadif/push box. (In the VP task)
Target utterance: The baby pushes a toy box.

It appears that learners use code-switching to fill the syntactic slot when they struggle to recall the target verbs or nouns. This is often indicated by pauses or hesitations. Sometimes, immediately after code-switching, they say, 'I don't know the word', but in their own language, Arabic.

The data reveal that some learners produced single-word utterances, as in (5.8a). Such utterances could be considered remnants of a stage preceding the VP stage, as suggested by Vainikka and Young-Scholten (2007), who view the use of single-word utterances by naturalistic learners as indicative of an initial stage of development—Stage 0—similar to the child's one-word stage, which conveys little about syntax. Similarly, the data show that some learners produced two-word utterances. The data also show that some learners produced verb-less utterances, as in (5.8b), which again convey little about syntax. These utterances are similar to those reported by Myles (2004), who argues for the addition of a verb-less stage. The utterance in (5.8b) has no pause between the subject and the verb, indicating a verb-less structure.

(5.8) a. Coffee.
Target utterance: The boy drinks coffee.
b. Woman breakfast.
Target utterance: The woman eats breakfast.

The data also show that two learners (Bilal and Arshad) overgeneralise the word 'and' in verb-less utterances, as in (5.9). This use of 'and' appears to be a placeholder for the verbs, as this use of 'and' never occurred with verbs in their data (e.g., 'the boy and play football') when they used verbs. This type of overgeneralisation differs from the overgeneralisation of verbal morphemes, as will be discussed in Section 5.6. This verb-less utterance appears to be characteristic of VP-stage learners, as similar instances were not observed in learners at more advanced stages.

(5.9) a. Baby and car.
Target utterance: The baby drives a car.
b. Girl and book.
Target utterance: The girl is reading a book.

It is worth highlighting that the data from these VP-stage learners lack nominal functional elements (i.e., noun-related morphology). Examples of these elements include the definite article 'the' (e.g., 'the book'), the indefinite article 'a' (e.g., 'a book') or 'an' (e.g., 'an apple'), and the plural *-s* (e.g., 'books'). The learners at this stage tend to omit the articles in obligatory contexts and produce utterances like those in (5.10a), and they tend to omit the plural *-s* and produce utterances like those in (5.10b).

(5.10) a. Boy drive car.
Target utterance: The boy drives a car.
b. Two man doctor. (In the copula 'be' task)
Target utterance: They are doctors.

The errors made by all learners in these articles were omissions rather than confusion between indefinite, definite, and zero articles because most of these articles are obligatory in the context, as in (5.10). However, the production or omission of articles or plural *-s* does not have a decisive role in assigning an OG stage, since Organic Grammar only deals with verbal morphology. However, nominal morphemes are assumed to be acquired before verbal morphemes; therefore, the learners in the following stages are expected to produce far more nominal morphemes than those in this stage.

As stated in the previous chapter, self-corrections were included in the data rather than the original utterances or forms. Learners at this stage used self-correction when they did not use specific verbs (e.g., non-target verbs), as illustrated in (5.11).

(5.11) a. Woman write/Woman read book. (In the progressive tense task)
Target utterance: The girl is reading book.
b. Girl but/put sorry throw shows. (In the progressive tense task)
Target utterance: The girl throws the shoos.

5.3.3 The NegP stage

As we saw in Chapter 2, Vainikka and Young-Scholten (2011) adopt the view that the NegP projection immediately dominates the VP. This assumption is consistent with German word order and is supported by child development data assuming any version of structure building. Given the locus of NegP in the Master Tree, Vainikka and Young-Scholten (ibid.) predict that for naturalistic L2 acquisition, there will be evidence for a NegP projection prior to any other functional projections (IP or CP).

In order to assign an OG stage to each learner, I used Déprez and Pierce's (1993) predictions regarding word order in negative constructions in L1 English as well as Organic Grammar. Déprez and Pierce (ibid.) predict the word order in negative constructions for L1 English. Déprez and Pierce note that these constructions are not designed to define specific developmental stages in acquisition, where only one construction type might be

present at any one time. Instead, the consistent use of any construction type is considered an indication of the presence or absence of necessary grammatical mechanisms. As Table 5.5 shows, the word order in negative sentences is presumed to be in I if it includes an auxiliary. Based on these predictions, I placed eight learners, who used the construction (S+Aux+Neg+V), at the NegP stage.

SR	V-to-I		V-to-C	English
-	-	SV	-	Neg (S) V
-	-	VS	-	Neg V (S)
+	-		-	S Neg V
-	+	SV	-	Aux Neg (S) V
-	+	VS	-	Aux Neg V (S)
+	+		-	S Aux Neg V
+	+		+	Aux S Neg V

Table 5.5: Predictions with respect to word order in negative sentences. (SR = subject raising; SV = subject-verb order at D-Structure; VS = possible verb-subject order at D-Structure)

Vainikka *et al.* (2017) placed four learners from their sample of 17 in the NegP stage. One of these learners used the construction (S+Aux+Neg+V), while the other three produced forms like no+V or do, but they showed some usage of the copula and auxiliary 'be', or the occasional use of the past tense *-ed* or the third-person singular *-s*. Although these learners consistently used the copula and auxiliary 'be', they were placed at the NegP stage. Consequently, I have assigned 28 learners in my sample to the NegP stage.

Name	Reading level	SVO	NegP			Past simple (*-ed*)	AgrP				
			no(t) V	is no(t) V	do forms		Cop be		Aux be		3- sg -s
							is	are	is	are	
Sameer	1	8/8	0/8	0/8	8/8	0/8	4/4	2/4	0/4	0/4	0/8
Andleebid	1	6/8	0/8	0/8	8/8	0/8	3/4	0/4	0/4	0/4	0/8
Zaid	1	8/8	0/8	0/8	8/8	0/8	4/4	0/4	0/4	0/4	0/8
Waleed	2	8/8	0/8	0/8	8/8	0/8	4/4	4/4	2/4	0/4	0/8
Hind	4	7/8	0/8	0/8	8/8	2/8	4/4	4/4	4/4	4/4	0/8
Raneem	1	8/8	1/8	0/8	7/8	0/8	4/4	0/4	0/4	0/4	0/8
Reham	3	8/8	0/8	0/8	8/8	0/8	4/4	4/4	2/4	4/4	0/8
Maraam	1	7/8	7/8	0/8	1/8	0/8	1/4	0/4	0/4	0/4	0/8
Adil	1	7/8	1/8	4/8	4/8	0/8	4/4	0/4	0/4	0/4	0/8
Gharam	1	8/8	0/8	0/8	8/8	0/8	0/4	0/4	0/4	0/4	0/8
Batool	2	8/8	3/8	0/8	5/8	0/8	4/4	4/4	0/4	0/4	0/8
Rafal	2	8/8	0/8	0/8	8/8	0/8	0/4	4/4	0/4	4/4	0/8
Ehab	2	8/8	0/8	0/8	8/8	1/8	4/4	4/4	2/4	0/4	0/8
Shihab	2	8/8	0/8	0/8	8/8	0/8	4/4	4/4	0/4	0/4	0/8
Aroob	3	8/8	0/8	0/8	8/8	0/8	4/4	4/4	0/4	0/4	0/8
Jaber	3	8/8	0/8	4/8	4/8	0/8	4/4	4/4	0/4	1/4	0/8
Afnan	3	8/8	0/8	0/8	8/8	4/8	0/4	0/4	0/4	0/4	0/8
Talal	3	8/8	0/8	0/8	8/8	0/8	4/4	4/4	0/4	0/4	0/8
Itab	3	8/8	0/8	0/8	8/8	0/8	4/4	4/4	0/4	0/4	0/8
Hazim	3	7/8	0/8	2/8	6/8	0/8	4/4	1/4	2/4	0/4	0/8
Razaan	3	8/8	0/8	3/8	5/8	1/8	4/4	4/4	0/4	0/4	1/8
Ahlaam	4	8/8	0/8	0/8	8/8	0/8	3/4	2/4	0/4	0/4	0/8
Tameem	3	8/8	0/8	0/8	8/8	0/8	4/4	4/4	4/4	4/4	0/8
Nihad	4	8/8	0/8	0/8	8/8	0/8	4/4	4/4	0/4	4/4	0/8
Ajmal	2	8/8	1/8	0/8	7/8	0/8	4/4	4/4	0/4	2/4	0/8
Nadeen	2	8/8	0/8	0/8	8/8	0/8	4/4	4/4	4/4	4/4	0/8
Jameel	3	8/8	0/8	0/8	8/8	0/8	4/4	0/4	2/4	2/4	0/8

Table 5.6: The NegP stage

Four learners produced the construction (S+Aux+Neg+V) as shown in (5.12a), and other learners produced a bare construction but also produced IP-related elements in response to other tasks. For example, Adil produced the construction (S+Neg+V) as in (5.12b), but in another task, he produced IP-related elements as shown in (5.12c).

(5.12) a. The woman is no open the door.

Target utterance: The women is not opining the door.

b. The boy not cleaning the room. (In the negation task)

Target utterance: The boy is not cleaning the room.

c. The girl is reading book. (In the progressive tense task)

Target utterance: The boy is not cleaning the room.

The data show that the most frequent construction of English negations is with the unanalysed 'don't+V' as in (5.13a).

(5.13) a. The boy don't drink milk.

Target utterance: The boy doesn't drink milk.

b. The girl doesn't wash the car.

Figure 5.1 visually presents the percentages of learners' performance in English negation. As we see, the most frequent form used is *do*-forms. This finding indicates that the use of 'no + V' is likely to disappear, as it is much lower compared to what VP-stage learners use. Moreover, these learners only use ('is + no(t) + bare verb'). This finding does not help to identify a stage in which learners use the progressive tense ('auxiliary + not + V-ing'), as suggested in Stage 3 by Cancino *et al.* (1978), which precedes the final stage where the do-forms are analysed. This phenomenon can be attributed to the methods used in this study. As we recall from Chapter 4, the tasks were designed to test the acquisition of negative constructions in general, without focusing on the use of specific tenses, such as the progressive tense.

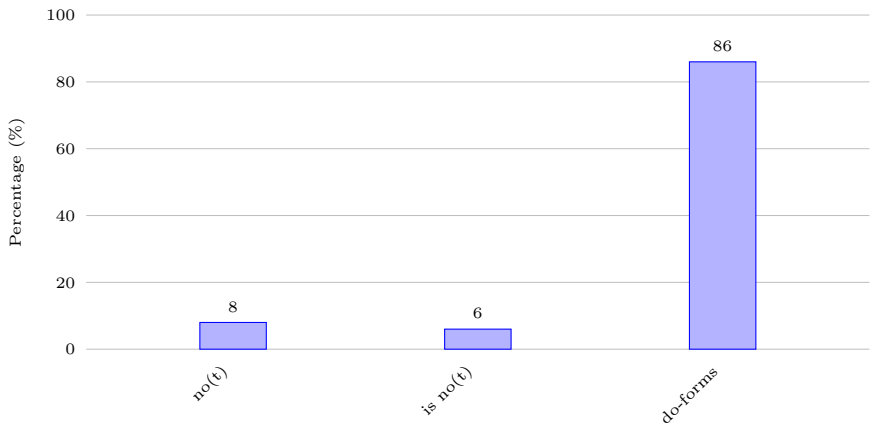


Figure 5.1: Negation by NegP-stage learners

Looking closely at the table above, we can observe that one learner, Afnan, used the simple past marker *-ed* four times. However, since she did not produce the copula 'be' as in (5.14a) or the auxiliary 'be' as in (5.14b), either in obligatory or non-obligatory contexts, I placed her in the NegP stage. These IP-related elements are vital components for the criteria of the next higher stage (the TP stage). Similarly, although Razan used the simple past marker *-ed* and the third-person singular *-s*, using them only once did not satisfy the criteria for the TP stage. In addition, she did not use the auxiliary 'be', which is again a vital component for the criteria in the TP stage.

(5.14) a. She nurse.

Target utterance: She is a nurse.

b. The girl carrying cat. (In the prgressive tense task)
Target utterance: The girl is carrying a cat.

The deviation from the anticipated morpheme order in Organic Grammar does not threaten the general morpheme order, as the remaining participants consistently appear to adhere to a fixed order.

As stated above, the data from the learners in the VP stage show the absence of using elements related to IP, apart from the inconsistent use of the copula 'be'. On the other hand, the data from NegP-stage learners show the presence of verbal morphemes. As Figure 5.2 shows, free verbal morphemes are more frequent in the NegP stage, with the copula 'be' being substantially higher than the auxiliary 'be'. The bound morphemes constitute only small proportions of the total scores. As we can see, the percentage of the simple past marker *-ed* is 2.7, and the percentage of the third-person singular *-s* is 0.4.

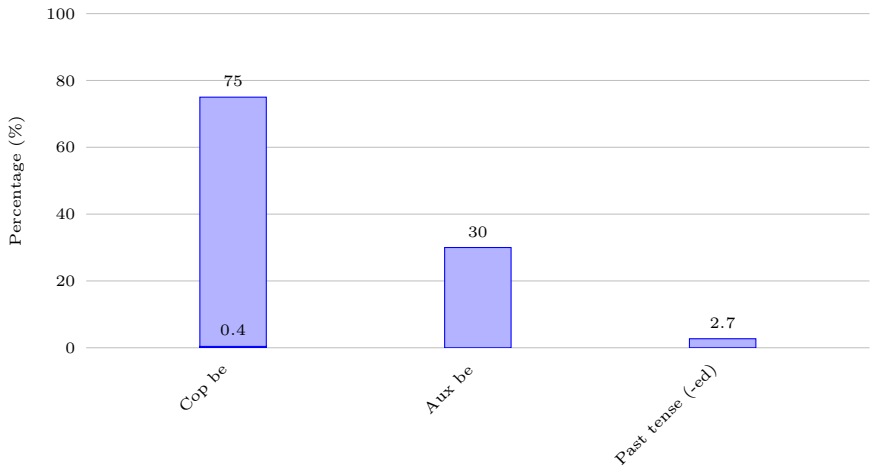


Figure 5.2: Verbal morphemes by NegP-stage learners

Like what was reported in the data from the learners in the VP stage, the learners in this stage used self-corrections. However, the occurrence of self-correction at this stage occurs in the context of negation as in (5.15a), the context of the copula 'be' as in (5.15a), and the context of the progressive tense as in (5.15b). These self-corrections highlight the low productivity of these morphemes, which are not necessarily target-like, and indicate that this stage is more complex than the VP stage in terms of morphosyntax.

(5.15) a. The woman not/the woman don't open the door (in the negation task).

Target utterance: The woman is not opining/doesnt open the door.

b. She studied/is student. (in the copula 'be' task)

Target utterance: She is a student.

5.3.4 The TP stage

As detailed in Chapter 2, Organic Grammar predicts that TP will be the first fully specified functional projection beyond NegP. According to Organic Grammar, learners will be placed at the TP stage if they productively mark tense and also show strong evidence for the previous stages (in the case of English, these are the VP and NegP stages). For the purpose of this study, the criterion for the TP stage is that learners must produce the regular or irregular past tense three times and consistently use the copula and the auxiliary 'be'. However, if they produce fewer instances of the regular or irregular past tense, they must produce one or two instances of the third-person singular *-s* at most to avoid any conflict with the criteria for the AgrP stage. These criteria are still in line with the nature of the developmental stages of OG, which states that development occurs incrementally, not abruptly.

Therefore, I assigned 13 learners to the TP stage. This is indicated by the consistent use of the correct copula and auxiliary 'be' and the emergence of the past tense *-ed*. As with the learners at the NegP stage, learners at the TP stage used the copula and auxiliary 'be', which carry agreement features. They were not placed at the AgrP stage because they do not yet show strong evidence of mastering the agreement paradigm (e.g., the agreement suffix *-s* on main verbs).

Name	Reading level	SVO	NegP			Past simple (*-ed*)	AgrP				
			no(t) V	is no(t) V	do forms		Cop be		Aux be		3- sg -s
							is	are	is	are	
Rabab	2	8/8	4/8	0/8	6/8	1/8	4/4	4/4	3/4	3/4	1/8
Jasim	4	8/8	0/8	0/8	8/8	4/8	4/4	4/4	1/4	4/4	0/8
Maher	3	8/8	0/8	0/8	8/8	3/8	4/4	4/4	0/4	1/4	0/8
Rihaab	3	8/8	0/8	0/8	8/8	3/8	4/4	4/4	4/4	0/4	0/8
Hajar	3	8/8	0/8	0/8	8/8	6/8	4/4	4/4	0/4	0/4	3/8
Yasin	4	8/8	0/8	0/8	8/8	3/8	4/4	4/4	0/4	0/4	1/8
Widad	4	8/8	0/8	0/8	8/8	1/8	4/4	4/4	1/4	4/4	1/8
Rajab	4	8/8	0/8	0/8	8/8	0/8	4/4	4/4	4/4	4/4	1/8
Moraad	3	8/8	8/8	0/8	0/8	2/8	4/4	0/4	0/4	0/4	1/8
Ameen	2	8/8	0/8	0/8	8/8	2/8	4/4	4/4	0/4	0/4	3/8
Bader	4	8/8	0/8	0/8	8/8	3/8	4/4	4/4	4/4	4/4	2/8
Nadeen	2	8/8	0/8	0/8	8/8	3/8	4/4	4/4	4/4	4/4	1/8
Hilal	4	8/8	0/8	2/8	6/8	6/8	4/4	4/4	2/4	2/4	2/8

Table 5.7: The TP stage

As shown in the table, four learners (Maher, Hajar, Moraad, and Ameen) deviated from the general order of the morphemes. These learners consistently used the copula 'be' and the simple past tense but omitted the auxiliary 'be', except Maher, who used it once. This deviation does not affect the order of the verbal morphemes predicted in Organic Grammar. As shown in Figure 5.3, the data in the graph show an order similar to that of the NegP stage, but with higher percentages. Again, learners at this stage produced more free morphemes than bound morphemes. This indicates that bound morphemes are more difficult to acquire at this stage.

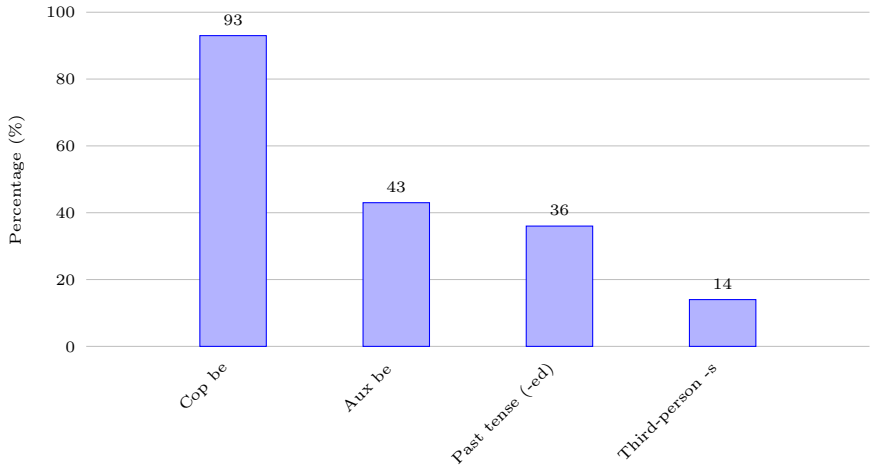


Figure 5.3: Verbal morphemes by TP-stage learners

Similaly the learners at this stage are more inclined to use *do*-forms. As we can see in Figure 5.4, the percentage of using this partical is markedly higher than the other particals.

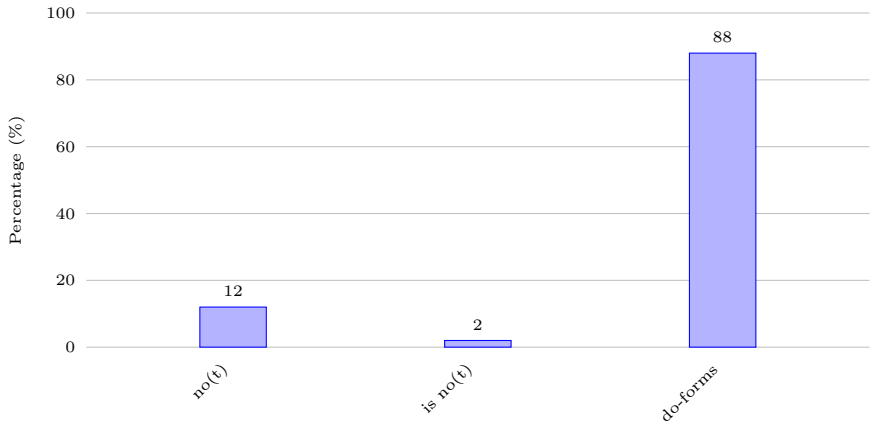


Figure 5.4: Negation by TP-stage learners

Similarly to what has been reported in the data from learners in the previous stages, the learners in this stage use self-corrections that occur in the context of the past tense, as in (5.16a), and the progressive tense (particularly are), as in (5.16b). These corrections also occur, to a limited extent, in the context of the copula 'be' (again, particularly 'are'), as in (5.16c), and in the context of negation, as in (5.16d).

(5.16) a. People see/saw boat. (In the past tense task)

b. The boy/the boys wash/the boys are wash/washing the plates. (In the past progressive tense task)

c. They doctors/They are doctors. (In the copula 'be' task)

d. The boy not east/is not eat/don't/the boy don't eat. (In the negation task)

5.3.5 The AgrP stage

Vainikka and Young-Scholten (1994) used various criteria to assign an AgrP stage to each L2 German learner. These criteria include the cessation of optional subjects, the near-obligatory status of verb raising (over 60%), and the acquisition of the agreement paradigm (modals, auxiliaries, and agreement features). Furthermore, to determine whether the agreement paradigm had been acquired, Vainikka and Young-Scholten added an additional criterion that there must be at least two correct instances of each of the four agreement suffixes. For the purpose of my study, I placed my learners at the AgrP stage if they supplied the third-person singular *-s* (three times) together with the consistent use of the copula and auxiliary 'be'. They must also mark tense (again, three times), consistently use L2 English VP (is+not+V) or *do*-forms for NegP, and cease the optionality of subjects. Consequently, I placed seven learners in the AgrP stage.

Name	Reading level	SVO	NegP			Past simple (*-ed*)	AgrP				
			no(t) V	is no(t) V	do forms		Cop be		Aux be		3- sg -s
							is	are	is	are	
Yasmeen	3	8/8	0/8	0/8	8/8	8/8	4/4	4/4	4/4	4/4	6/8
Zainab	3	8/8	0/8	0/8	8/8	4/8	4/4	4/4	2/4	0/4	3/8
Safwan	3	8/8	0/8	0/8	8/8	8/8	4/4	4/4	4/4	4/4	8/8
Nawal	4	8/8	0/8	0/8	8/8	6/8	4/4	4/4	4/4	4/4	4/8
Sajid	4	8/8	8/8	0/8	8/8	6/8	4/4	4/4	4/4	4/4	8/8
Eman	4	8/8	0/8	8/8	0/8	8/8	4/4	4/4	4/4	4/4	8/8
Nasser	4	8/8	0/8	8/8	2/8	8/8	4/4	4/4	4/4	4/4	8/8

Table 5.8: The AgrP stage

However, there was a learner (Zainab) who did not meet the criteria, but her per-

formance exceeded the criteria for the previous stage (the TP stage). She only used the copula 'are' twice and completely avoided using the auxiliary 'are' in progressive tenses, as in (5.17), and did not replace it with 'is' as was the case with the learners in the NegP or TP stage. This indicates that she knows that 'is' disagrees with the plural subject.

(5.17) The boys writing letters.

Target utterance: the boys are writing letters.

These learners are more consistent than those in the previous stages and have ceased the optionality of using empty subjects. However, they still use bare verbs, as in (5.18a), and self-corrections, which occur in the context of the past tense, as in (5.18b); the third-person singular *-s*, as in (5.18c); and to a small degree in the context of the progressive tense, as in (5.18d).

(5.18) a. Then she read newspaper.

Target utterance: Then she reads newspaper.

b. They watched/watched the boat. (In the past tense task)

c. Then he smoke a cigarette/then he smokes a cigarette.

d. The boys writing/are writing a letter. (In the progressive tense task)

In summary, the findings of the production data illustrate the gradual progression of language acquisition among learners. Each stage is marked by the presence of distinct verbal morphemes, beginning with the acquisition of the VP, followed by the NegP stage, and subsequently the intermediate stages of TP and AgrP. Moving forward, let us investigate whether there is a potential relationship or association between Organic Grammar and the development of literacy skills, as well as other external factors that may influence this connection.

5.3.6 Exploring the relationship between OG stages and non-linguistic factors

This section investigates the relationship between OG (Organic Grammar) stages and non-linguistic factors, including reading levels, years of native language schooling, duration of ESOL (English for Speakers of Other Languages) instruction, UK residence duration, and age. The analysis employed two methods. Initially, scatter plots were utilised to visualise potential relationships between OG stages and these factors, followed by Spearman correlations to quantify the associations.

Additionally, ordinal regression was performed to explore whether other factors influence the relationship between reading levels and OG stages. This method is particularly suited for this analysis because it allows for the investigation of multiple independent variables—whether ordinal, continuous, or categorical—on an ordinal dependent variable, such as OG stages. It helps uncover the unique contributions of each non-linguistic factor to the variability observed in OG stages, providing a comprehensive understanding of the relationships within the dataset.

As mentioned in the previous chapter, Spearman correlation does not impose assumptions beyond the requirement that the data be at least ordinal. Therefore, OG stages and reading levels are treated as ordinal variables (e.g., stages assigned values of 1, 2, 3, 4, and similarly for reading levels). This method is suitable for measuring the association between ordinal values and continuous data, although transforming continuous data into ranked values is preferable for this analysis. To prepare the data for analysis in SPSS, continuous variables such as years of native language schooling and ESOL instruction duration in months were transformed into sequential ranks. This transformation ensures that each value is unique, with the lowest value assigned Rank 1.

Ordinal regression is also suitable for analyzing the effects of multiple independent variables on a categorical dependent variable. In this investigation, OG stages are treated as categorical dependent variables because they represent distinct categories or levels (e.g., VP, NegP, TP, and AgrP assigned values of 1, 2, 3, 4). Variables such as years of native language schooling, ESOL instruction duration, and age are considered independent variables.

The key assumption of ordinal regression, specifically the proportional odds assumption, has been assessed using SPSS and shows no violation, as indicated by the Test of Parallel Lines. The p-value for the Chi-Square test is 0.304, which is higher than the significance level of 0.05.

5.3.6.1 OG stages and literacy

The initial investigation of the relationship between literacy and OG stages indicates a noticeable increase in the mean values. As Table 5.9 shows, learners with higher literacy levels tend to attain higher OG stages. Additionally, the SD suggests that learners with higher literacy levels are more likely to reach higher OG stages, as the SD is significantly

lower compared to those with lower literacy levels.

Literacy levels	N	Mean	SD
Literacy 1	14	1.43	0.51
Literacy 2	13	1.69	0.75
Literacy 3	21	2.57	0.75
Literacy 4	12	3.08	0.79

Table 5.9: Mean and SD for OG stages accros literecy levels

The scatter plot in Figure 5.5 provides further support for this positive relationship. It reveals a positive monotonic relationship between OG stages and literacy levels. An increase in literacy levels is accompanied by a corresponding increase in OG stages.

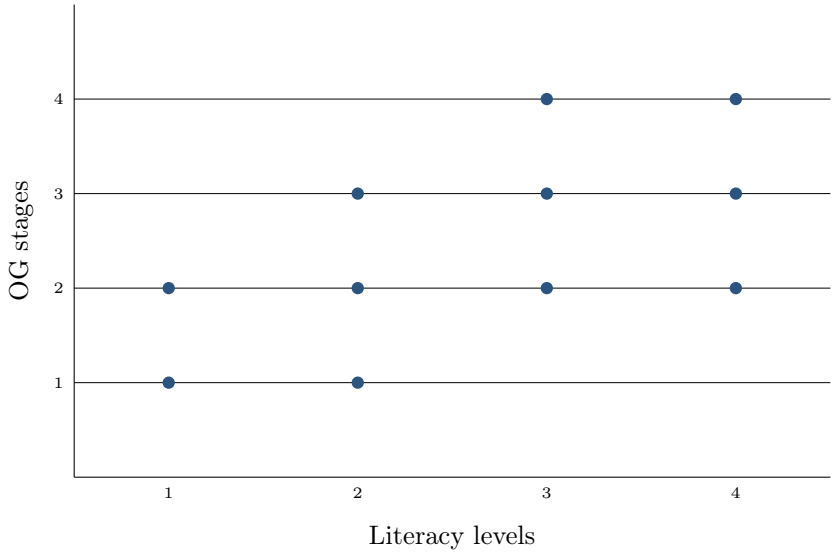


Figure 5.5: OG stages and literacy levels

As shown in Table 5.10, there is a significant positive correlation between OG stages and literacy levels. This finding suggests that learners with higher literacy levels tend to reach higher OG stages. Consequently, when learners' literacy levels are low, their progression is expected to be similarly limited. This confirms the hypothesis that lower literacy delays progression during the development of morphosyntax.

	Correlation	P-value
Literacy	0.681	0.001

Table 5.10: Correlation between OG stages and literacy

The results of the investigation into the relationship between OG stages and other non-linguistic factors (native language schooling, UK residence, ESOL instruction, and age) are presented in Figure 5.6. As we can see, the data indicate no significant connection between OG stages and these non-linguistic factors.

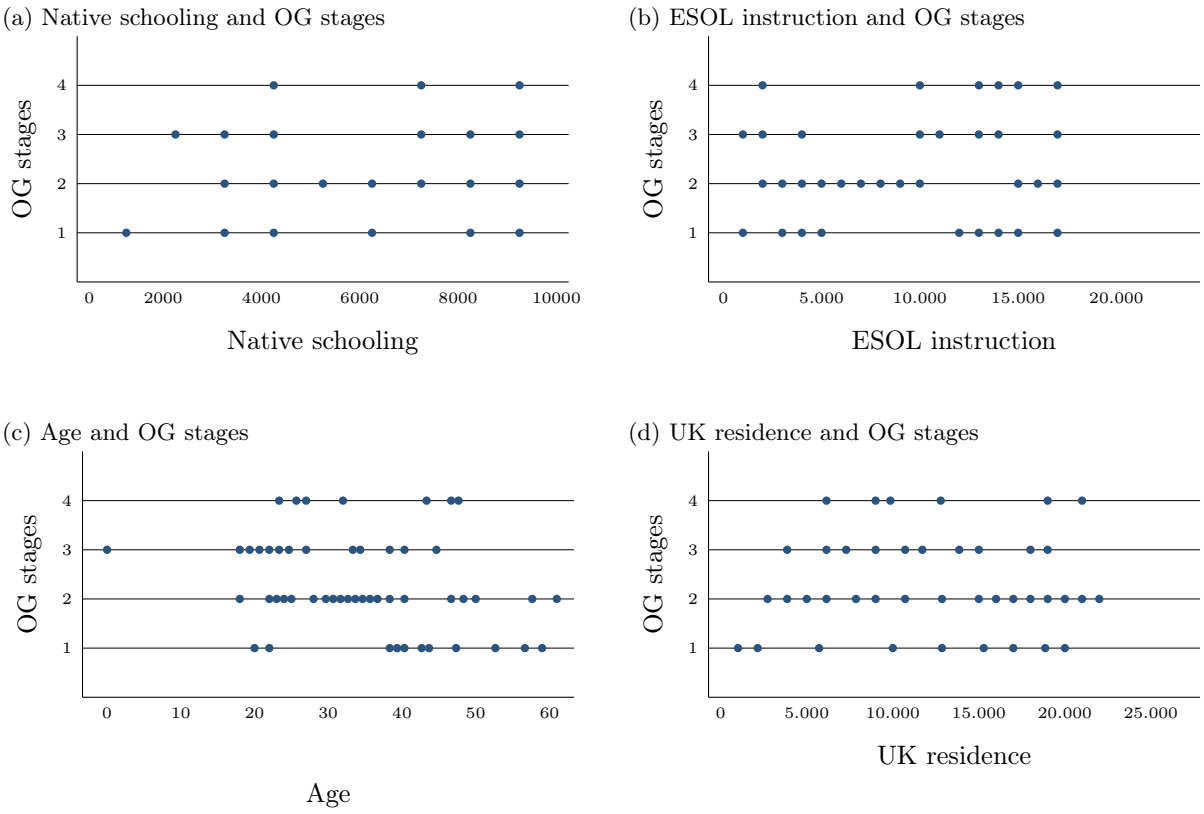


Figure 5.6: Comparison of negation in the NegP stage

However, Spearman's correlation analysis in Table 5.11 reveals a significant positive correlation between the duration of native language schooling and OG stages ($p < 0.05$). This finding suggests that as learners' native language schooling increases, their development of morphosyntax is expected to improve. However, this correlation is relatively weaker compared to the correlation observed between OG stages and reading levels, which is significant at the 0.001 level. Additionally, there is a strong negative correlation between OG stages and age, suggesting that the onset of language acquisition at older ages may negatively impact morphosyntactic development.

Variables	Correlation	P-value
Native schooling	.307	.017
ESOL instruction	.042	.748
Length of UK residence	.016	.902
Age	-.270	.037

Table 5.11: Correlation between OG stages and variables

In contrast, no significant correlation is observed between OG stages and ESOL instruction duration or the length of UK residence, as indicated by p-values greater than 0.05. This suggests that these non-linguistic factors have no impact on the development of morphosyntax.

Ordinal regression confirms that it is literacy, not other factors, that has a significant impact on the transition between OG stages. Specifically, the estimates for Reading Levels 1 and 2 indicate a notable negative effect (-4.193 and -3.610, respectively) on the transition between OG stages, with a highly significant p-value of less than 0.001. This negative effect suggests an increased likelihood of cases falling into a lower category (a lower OG stage) in the dependent variable as the predictor (reading level) decreases. In other words, the results suggest that learners with lower reading levels tend to have lower OG stages, confirming my hypothesis that lower literacy hinders the progression or transition between stages.

5.4 Comprehension data for the acquisition of verbal morphemes

This section presents the results of the comprehension study, constituting an original contribution to Vainikka *et al.*'s (2017) study. As outlined in Chapter 4, the comprehension tasks had two main objectives. The results of the first objective are presented here, aiming to investigate whether learners' performance in comprehension reflects the acquisition route discussed in Section 5.3.

To achieve this objective, the analysis was performed using two analytical approaches. First, an examination was performed to determine whether the learners in this study perform better on free morphemes than on bound morphemes, as observed in the production data in this thesis. As discussed in Chapter 2, the early acquisition of verbal free morphemes over bound morphemes aligns with the developmental stages in generative L2 acquisition theories (e.g., OG or MBSA). The analysis included both descriptive and statistical approaches. For the descriptive analysis, the study used the mean together

with the SD to determine whether the values of these verbal morphemes differed from each other and to explore the variability through the SD. For the statistical analysis, the Friedman test was performed to compare the distribution differences of these verbal morphemes. The independent variable was the verbal morphemes, and the dependent variable was the learners' scores. The Friedman test is equivalent to the Kruskal-Wallis test and is used when the data fail the main assumptions (e.g., the independence of observations). Therefore, my data did not meet this assumption. Dunn's post hoc test was used to identify precise differences between the verbal morphemes. Dunn's post hoc test is equivalent to the Tukey HSD test after ANOVA and is used following the Friedman test.

Second, an exploration was conducted to determine whether there is a relationship between OG stages and learners' scores in the comprehension of each verbal morpheme. This investigation used the mean and SD to examine whether these morphemes change in value according to the OG stages and to assess the variability through the SD. The investigation also used scatter plots to visualise potential relationships, followed by the Spearman correlation test to assess correlations. Together, these two analytical approaches contribute to a clearer understanding of the acquisition route by these learners, shedding light on the developmental stages in L2 acquisition.

5.4.1 Comprehension of verbal morphehms acquisition

The visualisation of the data in Figure 5.7 shows that higher mean values are observed with free morphemes (the copula 'be' and the auxiliary 'be'), with the highest values observed for the copula 'be'. This indicates that the comprehension of the copula 'be' is acquired before that of the auxiliary 'be', and therefore learners may produce the copula 'be' first, as observed in the production data in this thesis. Additionally, the comprehension of both free morphemes is acquired before that of bound morphemes, as shown in the bar chart. Lower mean values are observed with bound morphemes (the regular past tense *-ed* and the third-person singular *-s*), with the slightly lower mean observed for the third-person singular *-s*. This slight difference between the mean values in the comprehension of the regular past tense *-ed* and the third-person singular *-s* makes it difficult to determine which morpheme is comprehended first.

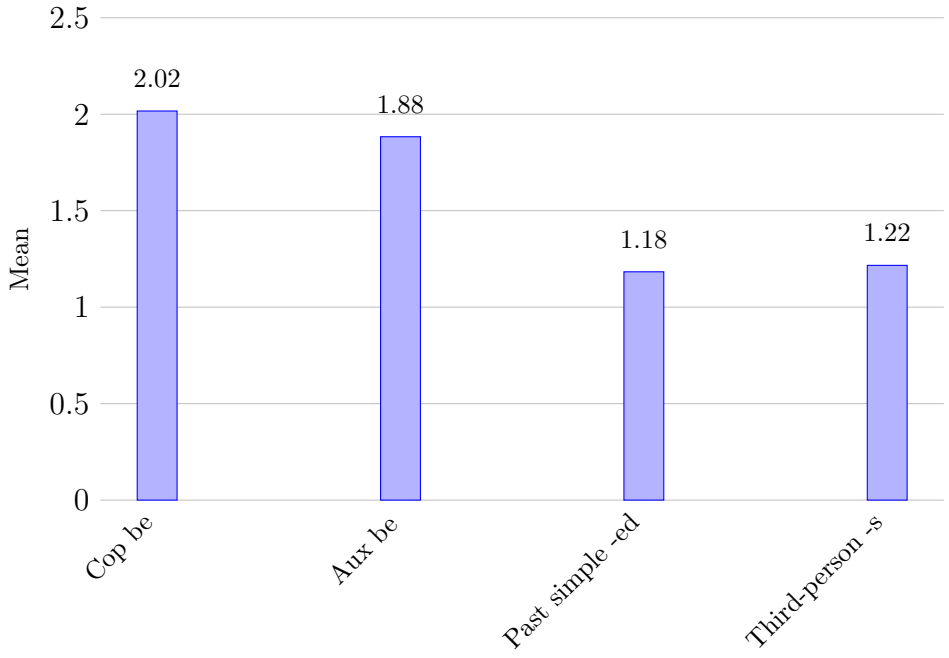


Figure 5.7: Mean for verbal morphemes

The relationship between the mean and SD values in Table 5.12 also indicates that the learners may comprehend the copula 'be' first. Although the variability observed between the mean and SD (M=2.02, SD=1.08) in relation to the comprehension of the copula 'be' suggests some inconsistency, the learners appear to be more consistent in the comprehension of the copula 'be' than that of the auxiliary 'be' (M=1.88, SD=1.04) and may produce the copula 'be' first. The relationship between the mean and SD values also shows that learners struggle more with the comprehension of the third-person singular *-s* than with the comprehension of the regular past tense *-ed*. Despite the slight difference in the mean shown above in the bar graph and in this table, the learners' scores in the third-person *-s* indicate higher variability (M=1.22, SD=1.09) than their scores in the regular past tense *-ed* (M=1.18, SD=0.93). This finding indicates that the learners may produce the regular past tense *-ed*, which functions as the head for TP, before the third-person *-s*, which functions as the head for AgrP, aligning with the stages of Organic Grammar where the TP is acquired before AgrP.

Cop 'be'	**Mean**	**SD**
Cop 'be'	2.02	1.08
Aux 'be'	1.88	1.04
Past simple *-ed*	1.18	0.93
Third-person *-s*	1.22	1.09

Table 5.12: Mean and SD for verbal morphemes

The results of the Friedman test and its subsequent Dunn's post hoc test partially support this finding. The Friedman test indicates a significant difference between the distribution of these morphemes, with a chi-squared value of 28.8303 and a p-value of 0.05. Dunn's post hoc test shows no difference between learners' scores in the copula 'be' and in the auxiliary 'be', but it shows that learners' scores in each free morpheme significantly differ from learners' scores in each bound morpheme. Specifically, the comparisons of the copula 'be' vs the regular tense (Z = 4.211, p = 0.0001), the copula 'be' vs the third-person singular (Z = 4.002, p = 0.0001), the regular past vs the auxiliary 'be' (Z = 3.523, p = 0.0004), and the third-person singular *-s* vs the auxiliary 'be' (Z = 3.314, p = 0.0007) are all significant. This finding indicates that learners may first comprehend the agreement features of the copula 'be' and the auxiliary 'be' before the tense feature (the *-ed* on main verbs) or the agreement feature on main verbs (the *-s* on main verbs).

Now let us discuss the relationship between OG stages and verbal morphemes, including negation.

5.4.2 Comprehension data and OG stages

This section presents the results of the relationship between OG stages and English negation, as well as the relationship between OG stages and free and bound morphemes. First, it discusses the relationship between OG stages and English negation, then the relationship between OG stages and free morphemes, and finally the relationship between OG stages and bound morphemes.

5.4.2.1 English negation comprehension across OG stages

According to Table 5.13, there appears to be a relationship between learners' OG stages and their scores in the comprehension of English negation constructions. As we can see, the mean value increases across OG stages. The data also show that the learners in the VP stage exhibit more variability in their scores, with the SD being close to the mean, compared to those in the other stages. On the other hand, the learners in the AgrP stage exhibit greater consistency in their scores, with the SD being far below the mean.

OG stages	N	Mean	SD
VP	14	1.86	1.29
NegP	26	2.38	0.98
TP	13	2.62	0.87
AgrP	7	2.86	0.38

Table 5.13: Mean and SD for English negation construction across OG stages

The scatterplot in Figure 5.8 indicates this relationship. As we can see, learners in the AgrP stage tend to consistently have higher scores compared to those in the other stages. The range is from 2 to 3, while the scores for the learners in the lower stages range from 0 to 3.

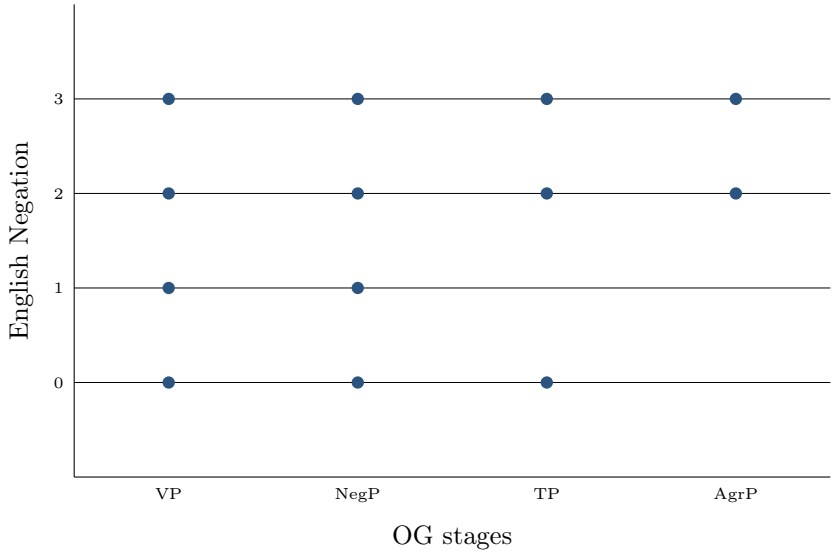


Figure 5.8: Relationship between English negation and OG stages

Analysis of Spearman's correlation corroborates this finding by indicating that there is a postive significant relationship at the significance level of 0.05 ($\rho = 0.281$, $p = 0.03$). This correlation reflects the data from the production tasks, where learners in the higher OG stages consistently use English constructions, typically employing 'not' or *do*-forms rather than 'no', as the learners in the lower stages do. Now, let us discuss the results of the relationship between learners' OG stage and their comprehension of verbal morphemes. First, I will present the relationship between free morphemes, followed by the relationship between bound morphemes.

5.4.2.2 Free morphemes comprehension across OG stages

The data from the learners across OG stages indicate that they behave similarly. However, AgrP-stage learners perform slightly better. As shown in Table 5.14, the mean of these morphemes increases noticeably, particularly among learners from higher OG stages (TP or AgrP). The SD also shows very little variability, especially among AgrP-stage learners. As we can see, the SD is considerably lower than the mean, which indicates some consistency in the data. This indicates that the learners in the AgrP stage show less variability than the other learners, reflecting their performance in the production task where they consistently used the copula 'be' and auxiliary 'be'.

Copula 'be' in relation to OG Stages			
OG Stages	N	Mean	SD
VP	14	1.93	0.92
NegP	26	1.92	1.20
TP	13	2.00	1.15
AgrP	7	2.57	0.79
Auxiliary 'be' in relation to OG Stages			
OG Stages	N	Mean	SD
VP	14	1.71	0.994
NegP	26	1.77	1.107
TP	13	1.77	1.013
AgrP	7	2.86	0.378

Table 5.14: Mean and SD for free morphemes in relation to OG stages

The data presented in Figure 5.9 reflect the similarity of these learners in response to OG stages. However, examining the data closely, we can see that the baseline scores are higher for learners with the AgrP stage.

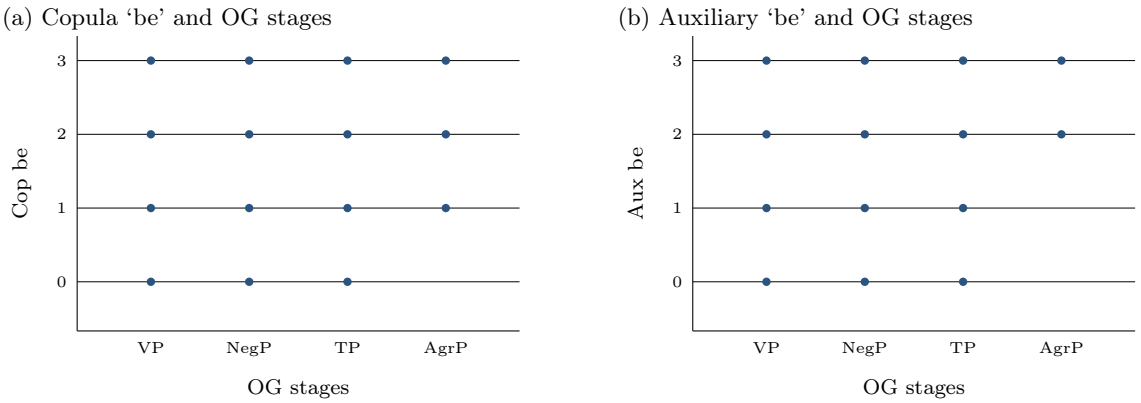


Figure 5.9: Relationship between English negation and OG stages

Spearman's correlation confirms this lack of association, revealing no correlation between OG stages and these morphemes, as shown in Table 5.15. This finding is expected, given the consistent performance of these morphemes by learners, as presented in Section 5.4.1, regardless of their OG stages.

Variables	Correlation	P-value
Cop 'is'	.124	.345
Cop 'are'	.172	.190
Aux 'is'	.207	.112
Aux 'are'	.292	.024

Table 5.15: Correlation between OG stages and learners' individual scores in the comprehension of free morphemes

5.4.2.3 Bound morphemes comprehension across OG stages

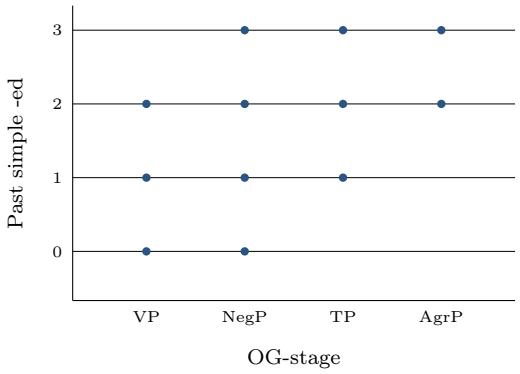
Unlike the mean values in learners' scores for free morphemes, the data in Table 5.16 show a noticeable increase in the mean values of learners' scores for bound morphemes based on their OG stages. This exactly reflects their performance in the production tasks, where TP- or AgrP-stage learners produce more bound morphemes compared to those in the VP or NegP stages. Similarly, less variability is observed with AgrP-stage learners' scores compared to the lower stages (particularly the VP or NegP stages). As we can see, the SD is very close to the mean in the VP and NegP stages, suggesting considerable variability in the data from these two stages.

Regular past simple *-ed* in relation to OG stages			
OG Stages	N	Mean	SD
VP	14	.71	.726
NegP	26	.96	.871
TP	13	1.62	.650
AgrP	7	2.14	1.069
Third-person singular *-s* in relation to OG stages			
OG Stages	N	Mean	SD
VP	14	.79	1.122
NegP	26	1.15	.925
TP	13	1.31	1.109
AgrP	7	2.14	1.215

Table 5.16: Mean and SD for bound morphemes in relation to OG Stages

The data in Figure 5.10 show no distinct relationship due to the uniform distribution of the data. However, the data in Figure 5.10a show that the lowest scores for the regular past tense *-ed* are observed among learners in the lowest stage (VP stage). This finding suggests that as learners' OG stage increases, their scores in the comprehension of the regular past tense *-ed* improve. This reflects their performance in the production tasks, where they used more tenses than those in the lower OG stages (VP or NegP).

(a) Past simple *-ed*



(b) Third-person *-s*

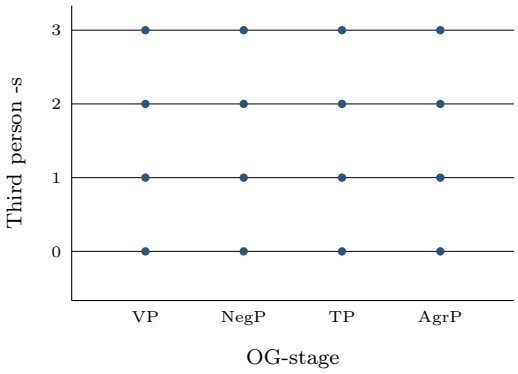


Figure 5.10: Percentage of verbal morphemes by NegP-stage learners

However, Spearman correlation analysis, as presented in Table 5.17, demonstrates a strong correlation between learners' OG stages and their comprehension scores of these bound morphemes. Specifically, the correlation is robust for the regular past tense *-ed*, with a p-value less than 0.001, and less pronounced for the third-person singular *-s*,

where the p-value is less than 0.05. This correlation likely reflects learners' performance in production tasks, aligning with our earlier discussion in this chapter, where learners in the TP or AgrP stages consistently produced more bound morphemes compared to those in the VP or NegP stages.

Variables	Correlation	P-value
Regular past tense *-ed*	.475	<.001
Third-person singular *-s*	.312	.015

Table 5.17: Correlation between OG stages and learners' individual scores in the comprehension of bound morphemes)

5.5 Copula 'be' as a potential trigger

The positive correlation between literacy and OG stages through Spearman's correlation and the effect of literacy on OG stages through ordinal regression led to an investigation of the triggers of morphosyntax and literacy. Following Hawkins' (2001) idea, the focus was on exploring the copula 'be' as a potential trigger in the development of negation constructions and functional projections. This examination of the copula 'be' as a trigger aligns with the results of the production and comprehension data in this thesis, which indicate that the copula 'be' appears to be acquired first, aligning with the literature. Additionally, the learners in this study are more consistent with the copula 'be' than with any other verbal morphemes. As mentioned in the methodology chapter, investigating these triggers is not the primary focus of the thesis, but it is still crucial for understanding more about the role of literacy in the acquisition of morphosyntax.

The investigations in this section will first include an examination of the relationship between the copula 'be' and verbal morphemes, including negation constructions, in both production data and comprehension data. If there is a positive relationship as expected, I will investigate whether the copula 'be' and literacy correlate in both production and comprehension data.

The analysis includes the mean and SD to explore whether the mean values of verbal morphemes change in relation to the copula 'be', with the SD used to assess variability. Visual presentations through scatter plots were used to illustrate any potential relationships. The analysis also includes Spearman's correlation to evaluate these relationships. A similar approach was employed for examining the relationship between the copula 'be'

and literacy levels, assessing whether the mean values for the copula 'be' vary according to literacy, with variability measured by SD and relationships visualised through scatter plots and correlations evaluated by Spearman's.

5.5.1 Copula be as a potential trigger: Production data

This section first presents the investigation of the relationship between learners' performance in the production tasks for the acquisition of the copula 'be' in relation to the acquisition of negation, and then the acquisition of the copula 'be' in relation to the acquisition of verbal morphemes.

5.5.1.1 Copula 'be' and negation construction

An examination of Table 5.18 shows that the mean scores for using 'no(t)' increase with some fluctuations as the learners achieve lower scores in the copula 'be', suggesting a negative relationship. However, the SD shows high variability, indicating that not all learners with lower scores in the copula 'be' use 'no(t)'. On the other hand, the mean values for using 'is no(t)', albeit with some fluctuations, or *do*-forms increase as learners' scores in the copula 'be' increase, especially for *do*-forms and with very little variability. This finding suggests that there is a positive relationship between *do*-forms, not necessarily 'is no(t)', and learners' scores in the copula 'be'.

Cop 'be'	**N**	**no(t)**		**is no(t)**		**do-forms**	
		Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**
0	11	4.91	3.94	0	0	1.6	3.58
1	2	3.50	4.95	0	0	4.5	4.95
3	1	0	-	0	0	8.0	-
4	9	4.56	4.10	2.9	1.07	1.93	3.29
5	3	0	0	1.0	1.4	7.0	1.41
6	1	0	-	1.0	1.4	7.0	1.41
8	33	0.36	1.08	0.68	2.04	6.85	2.50

Note: '-' refers to a missing value which occurs when there is only one observation, making SD calculations not applicable.

Table 5.18: Mean and SD for English negation construction in relation to the copula 'be' in production

Figure 5.11 also reflects this relationship. As shown in Figure 5.11a, learners with the lowest scores in the use of the copula 'be' (e.g., 0/8) tend to produce 'no(t)' more frequently, indicating a negative relationship. However, there are some instances where learners with the highest scores (8/8) also use 'no(t)'. Figure 5.11b, on the other hand, shows a positive relationship between learners' highest frequency use of the copula 'be'

and their use of ‘is no(t)’. However, there appears to be a consistently low frequency of using ‘no(t)’ even when the frequency of the copula ‘be’ increases (e.g., from 0 to 8).

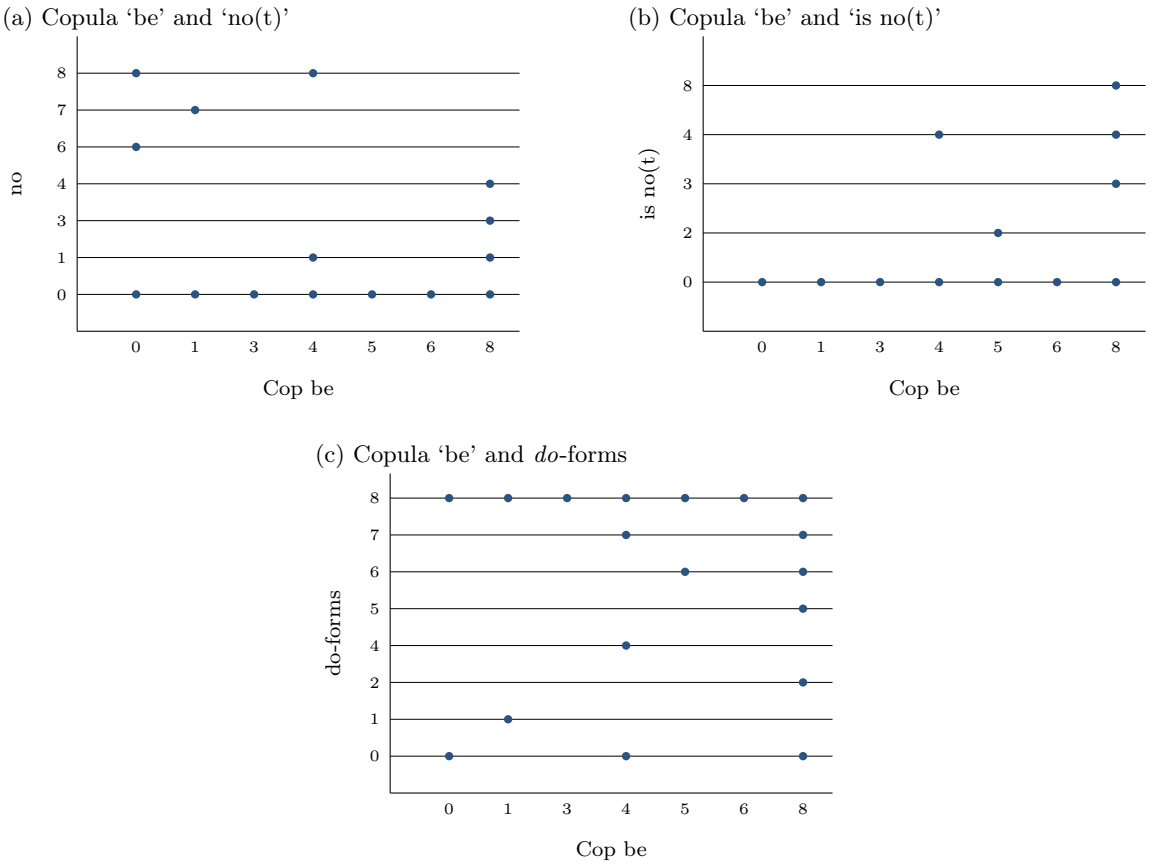


Figure 5.11: Relationship between copula ‘be’ and negation construction in production

The relationship between the use of the copula ‘be’ and the use of *do*-forms also appears to be positive and stronger than the relationship with ‘is no(t)’. As shown in Figure 5.11c, an increase in the use of the copula ‘be’ is accompanied by an increase in the use of *do*-forms. However, there is a higher use of *do*-forms even when learners produce the copula ‘be’ less frequently.

The statistical analysis in Table 5.19 reveals a significant negative correlation between the acquisition of the copula ‘be’ and the use of ‘no(t)’ as a negator. This means that when participants use the copula ‘be’ more often, they tend to use ‘no(t)’ less frequently to negate sentences. Therefore, there is an inverse relationship between the copula ‘be’ and ‘no(t)’ as negators. However, the analysis reveals a strong positive correlation between the use of the copula ‘be’ and ‘don’t’. As the frequency of the copula ‘be’ increases, the usage of ‘don’t’ tends to increase as well. So, there is a direct relationship between the copula ‘be’ and ‘don’t’. These correlations indicate that when learners tend to use the copula

'be' more frequently, they tend to use 'no(t)' less frequently, while simultaneously tending to use 'don't' more frequently. On the other hand, the analysis shows no relationship between the frequency of the copula 'be' and 'is no(t)' as a negator. This can be attributed to the low frequency rate in the use of 'is no(t)'.

Morphosyntactic elements	Correlation	P-value
no(t)	0.528	<0.001
is no(t)	0.113	0.391
do-forms	0.636	<0.001

Table 5.19: Correlation between the acquisition of copula 'be' and English negation

5.5.1.2 Copula 'be' and verbal morphemes

Similarly, the highest mean values for verbal morphemes (the auxiliary 'be', the irregular/regular past tense, and the third-person singular *-s*) are observed with the highest frequency of using the copula 'be'. In contrast, the lowest mean values are observed with the lowest frequency of using the copula 'be'. These results imply that there is a positive relationship between the acquisition of the copula 'be' and that of these verbal morphemes. However, high variability is observed with the lowest and highest mean values, making it difficult to infer a potential positive relationship. For example, learners with a higher frequency of using the copula 'be' vary considerably in using the third-person singular *-s* (M=1.82, SD=2.73).

Cop 'be'	**N**	**Aux be**		**Past *-ed***		**Third -s**	
		Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**
0	11	0.00	0.00	0.00	0.00	0.00	0.00
1	2	0.00	0.00	2.00	2.83	0.00	0.00
3	1	0.00	-	0.00	-	0.00	-
4	9	0.44	1.33	0.22	0.67	0.11	0.33
5	3	2.00	2.00	1.00	1.73	0.00	0.00
6	1	0.00	-	0.00	-	0.00	-
8	33	3.94	3.32	2.55	2.88	1.82	2.73

Note: '-' refers to a missing value which occurs when there is only one observation, making SD calculations not applicable.

Table 5.20: Mean and SD for verbal morphemes in relation to the copula 'be' (production)

The scatter plots in Figure 5.12 provide further support for the positive relationship between learners' performance in the copula 'be' and that of these verbal morphemes. As seen in the graphs, consistency in using the copula 'be' is accompanied by a notable increase in the use of the auxiliary 'be', the regular past tense *-ed*, and the third-person

singular -*s*. Lower production in using the copula 'be' is accompanied by a nearly complete absence of producing these morphemes.

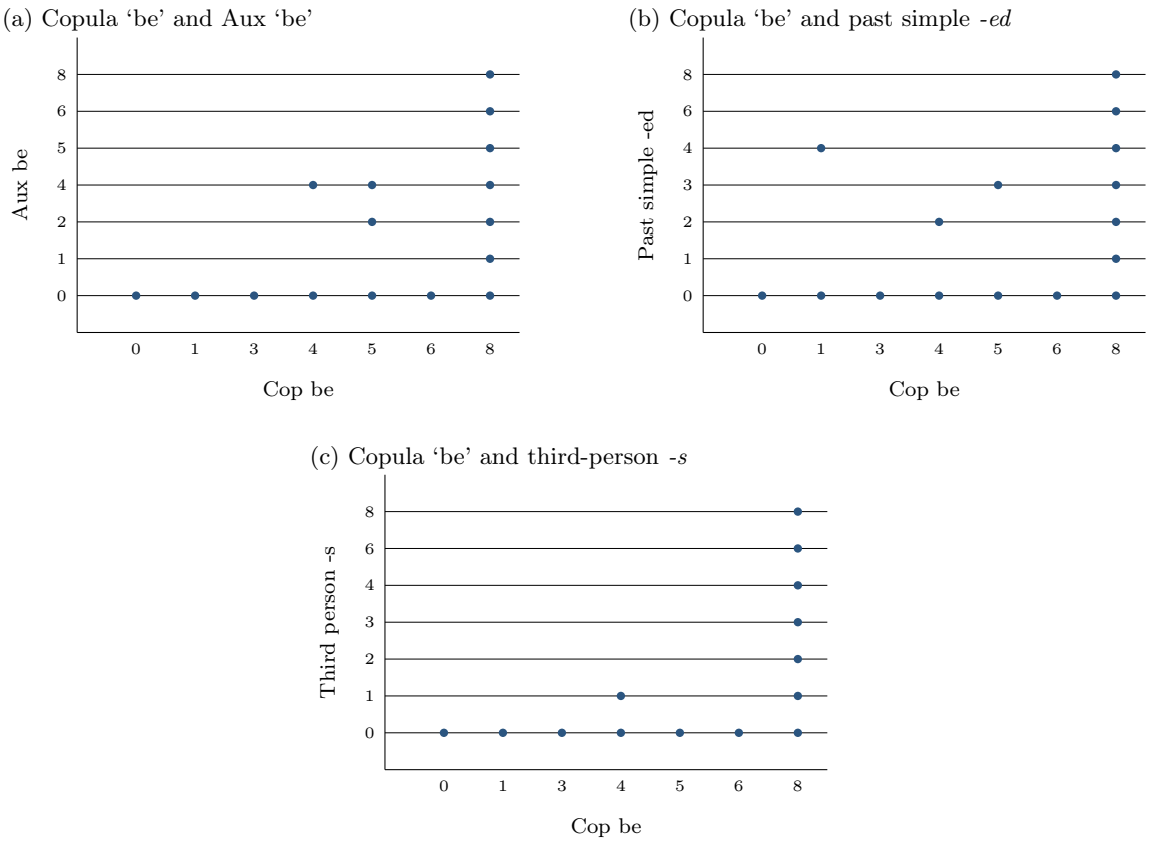


Figure 5.12: Relationship between the acquisition of copula 'be' and verbal morphemes in production

Spearman's correlation confirms this corresponding positive relationship. As shown in Table 5.21, the analysis indicates a significant relationship at the 0.01 significance level between the copula 'be' and these morphemes. This finding suggests that when learners produce more instances of the copula 'be', they are also expected to produce these morphemes.

Morphosyntactic elements	Correlation	P-value
Aux be	0.515	<0.001
-ed	0.516	<0.001
Third-person singular -*s*	0.546	<0.001

Table 5.21: Correlation between the acquisition of copula 'be' and verbal morphemes in production

5.5.2 Copula be as a potential trigger: Comprehension data

This section first presents the investigation of the relationship between learners' performance in the comprehension tasks for the acquisition of the copula 'be' in relation to the acquisition of negation, and then the acquisition of the copula 'be' in relation to the acquisition of verbal morphemes.

5.5.2.1 Copula 'be' and negation construction

The analysis of comprehension does not fully reflect this relationship. Although there are slight increases in mean values, there is notable fluctuation, as shown in Table 5.22. However, the highest mean value and the least variability (M=2.63, SD=0.79) are observed with scores of 3 for the copula 'be'. This means that when learners perform well on the copula 'be', they are expected to show little variability in their comprehension of negation.

Cop 'be'	**N**	**Mean**	**SD**
0	8	2.38	1.19
1	10	1.8	0.92
2	15	2.27	1.28
3	27	2.63	0.79

Table 5.22: Mean and SD for English negation construction in relation to the copula 'be' in comprehension

The scatter plots in Figure 5.13 do not reinforce any particular relationship between learners' scores in the copula 'be' and their scores in negation. The data is uniformly distributed, indicating that there is no relationship. Spearman's correlation also confirms this lack of relationship ($\rho = 0.233$, $p = 0.073$).

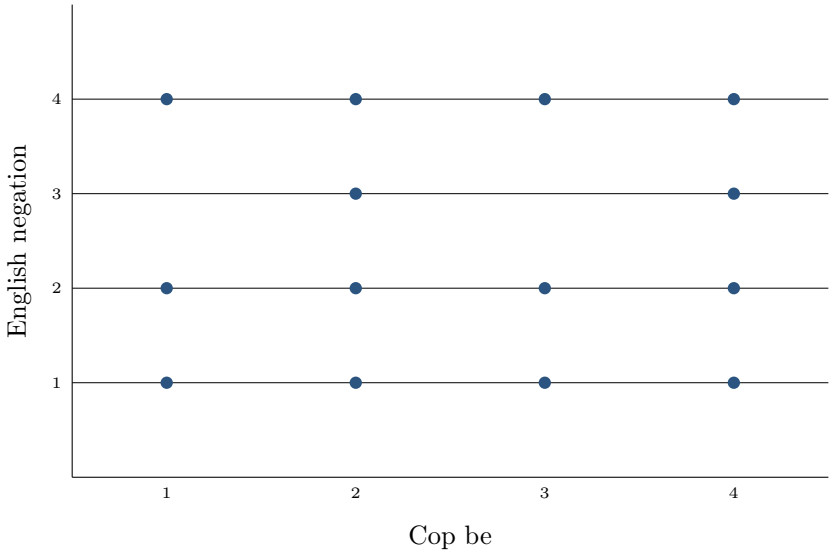


Figure 5.13: Relationship between copula 'be' and English negation in comprehension

5.5.2.2 Copula 'be' and verbal morphemes

The mean values, as shown in Table 5.23, for learners' comprehension of verbal morphemes relative to their scores in the copula 'be' do not exhibit consistent increasing patterns. Instead, they show notable fluctuations. However, the mean value of learners' scores tends to be higher when the copula 'be' score is 3, the highest possible score. This suggests a positive relationship between the acquisition of these morphemes and the acquisition of the copula 'be'. The SD data also support this observation, indicating less variability when the copula 'be' score is 3, particularly in the scores for the auxiliary 'be' and the regular past tense *-ed*, with SD values significantly below the mean. This finding suggests that learners with higher comprehension scores for the copula 'be' demonstrate more consistency compared to those with lower scores.

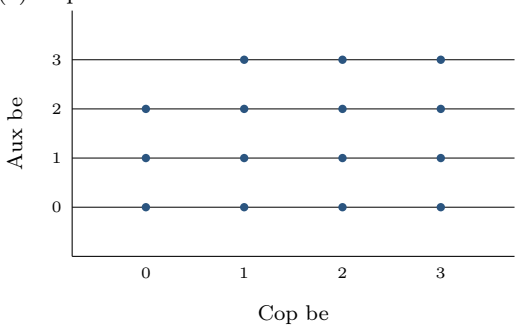
	Aux be				**Past tense -ed**				**Third-person -s**			
Cop 'be'	**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**	**0**	**1**	**2**	**3**
N	8	10	15	27	8	10	15	27	8	10	15	27
Mean	0.63	1.5	2.07	2.3	1.13	0.9	0.8	1.52	0.75	0.9	1.6	1.26
SD	0.74	1.18	0.88	1.13	1.15	0.99	0.94	0.75	0.89	0.74	0.99	1.26

Table 5.23: Mean and SD for auxiliary 'be', regular past tense *-ed*, and third-person singular *-s* in relation to copula 'be' in comprehension

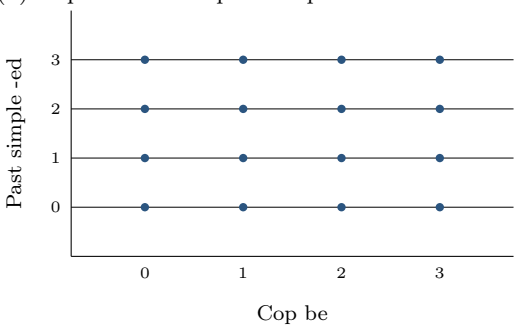
Similarly, the scatter plots in Figure 5.14 do not show increasing systematic trends, particularly in the learners' scores in the regular past tense *-ed*, as shown in Figure 5.14b,

which shows uniform distributions of the data. However, a closer look at Figures 5.14a and 5.14c reveals that learners' scores increase according to their scores in the copula 'be'. For example, learners with scores of 1 to 3 in the copula 'be' have higher scores in the auxiliary 'be' than those with scores of 0. Similarly, learners with scores of 2 and 3 in the copula 'be' perform better in the third-person singular *-s* than those with scores of 1 and 0.

(a) Copula 'be' and Aux 'be'



(b) Copula 'be' and past simple *-ed*



(c) Copula 'be' and third-person *-s*

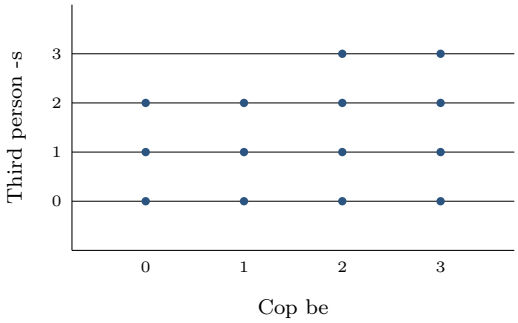


Figure 5.14: Relationship between the acquisition of copula 'be' and verbal morphemes in comprehension

Although the results from the mean and SD and the scatter plots do not show a distinct relationship, Spearman's correlation, as shown in Table 5.24, indicates a significant correlation between the copula 'be' and the auxiliary 'be', as well as between the copula 'be' and the regular past tense *-ed*, with p-values less than 0.05. This correlation could indicate that the copula 'be' triggers the acquisition of the auxiliary 'be' and the regular past simple *-ed*. On the other hand, the analysis shows no meaningful relationship between the copula 'be' and the third-person singular *-s*. This finding is still in line with the idea of triggers if we consider that the acquisition of the copula 'be' allows L2 learners to break into the inflectional systems. Both the auxiliary 'be' and the past tense *-ed* are inflectional morphemes.

Morphosyntactic elements	Correlation	P-value
Aux be	0.473	<0.001
Past simple -ed	0.287	0.02
Third-person singular -s	0.142	0.278

Table 5.24: Correlation between the acquisition of copula 'be' and verbal morphemes in comprehension

If this analysis is accurate and literacy affects the development of morphosyntax, as shown in Section 5.3.6, we expect that learners' scores for the copula 'be' in production and comprehension will increase in tandem with reading levels. So, let us now investigate this relationship before moving on to presenting the results of overgeneralisations.

5.5.3 Literacy and copula 'be'

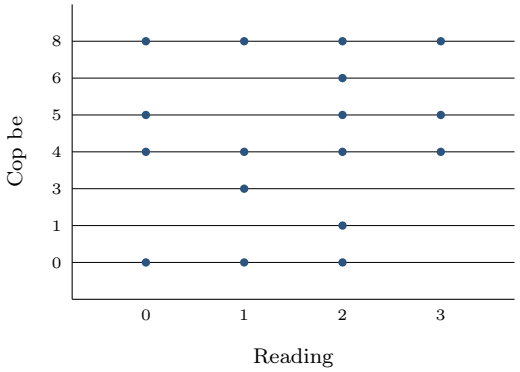
The descriptive statistics in Table 5.25 show that learners' mean scores change according to their literacy levels in an increasing pattern. As we can see, higher means are associated with higher levels of literacy. Additionally, learners with higher levels of literacy show far more consistency in the copula 'be' in both production and comprehension tasks, as indicated by the SD, which is well below the mean, especially for those with Literacy Level 4.

	Cop be (production)				**Cop be (comprehension)**			
Literacy	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
N	14	13	21	12	14	13	21	12
Mean	3.79	3.92	6.33	7.42	1.43	1.92	2.14	2.58
SD	3.68	3.28	2.73	1.38	1.16	1.12	1.06	0.67

Table 5.25: Mean and SD for copula 'be' in relation to literacy levels

On the other hand, Figure 5.15 shows no clear patterns of the relationship between literacy levels and the copula 'be'. However, a closer examination of these two graphs reveals that the baseline for learners at literacy levels 3 and 4 starts at scores of 2 and 3, respectively, in contrast to learners at lower literacy levels, whose baselines start at the lowest score of 0. This finding indicates that higher literacy levels are associated with higher baselines for the scores of these two morphemes.

(a) Literacy levels and the acquisition of the copula 'be' in production



(b) Literacy levels and the acquisition of the copula 'be' in comprehension

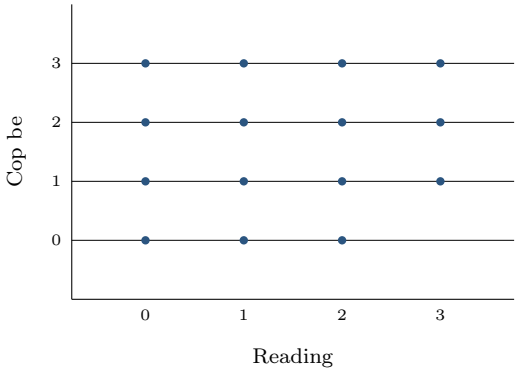


Figure 5.15: Relationship between literacy levels and the acquisition of the copula 'be' in production and comprehension

Spearman's correlation supports this positive relationship. As shown in Table 5.26, there is a significant relationship between literacy levels and learners' scores in the copula 'be' at the significance level of 0.01 in both production and comprehension. These correlations indicate that learners with higher levels of literacy tend to have higher scores in the production and comprehension of the copula 'be'. These correlations suggest that learners must be literate first to acquire the copula 'be', which in turn triggers the acquisition of inflectional morphemes. I will return to this topic in the following chapter.

	Correlation	P-value
Cop be (production)	0.448	<.001
Cop be (comprehension)	0.356	0.005

Table 5.26: Correlation between literacy levels and the acquisition of the copula 'be' in production and comprehension

5.6 Overgeneralisation by adult L2 immigrants

As discussed in Chapter 3, overgeneralisation in the second language acquisition of adult immigrants has been studied within three frameworks: Organic Grammar, the Minimalist Program, and the Basic Variety. For L2 English and L2 German, Vainikka and Young-Scholten (2019) and Vainikka *et al.* (2017) investigated overgeneralisation using the structure-building approach of Organic Grammar (Vainikka and Young-Scholten 1994, 1996a,b, 2011). Researchers of L2 Dutch, such as Julien *et al.* (2016) and Van de Craats and Van Hout (2010), explored overgeneralisation within the Minimalist Program (Chomsky 1995). Mocciaro (2019a,b) studied this phenomenon in L2 Italian based on the de-

velopmental path adopted in the Basic Variety by W. Klein and Perdue (1997). These researchers argue that overgeneralised forms serve as placeholders (Vainikka and Young-Scholten 2019; Vainikka *et al.* 2017) or dummy auxiliaries (Julien *et al.* 2016; Mocciaro 2019a,b; Van de Craats and Van Hout 2010; Verhagen 2013). In general, these approaches converge on the conclusion that overgeneralised forms fulfill specific syntactic functions.

My goal in this section is to confirm the specific stage at which learners start using overgeneralised forms, which has been previously examined. Then, I will investigate how these OG stages affect the choice of these forms. These forms differ in interesting ways and will be explored based on their OG stages presented in Section 5.3. This approach aims to provide new insights into the interplay between OG stages and the choice of these forms. To further understand overgeneralisation, I will investigate whether there is a correlation between the use of overgeneralised forms and the comprehension of morphosyntax. To my knowledge, this is the first study to evaluate this potential correlation.

5.6.1 Overgeneralisation at the early stages of Organic Grammar

As we recall from the OG stages placement in Section 5.3, analysis placed 13 learners in the VP stage because their VPs contain an optional specifier (the subject), bare verbs, and their complements (e.g., the direct object) with no functional morphology, apart from the copula 'is', which is also used as a default form (e.g., 'they is students'). Their negation construction contains only 'no' + verb or multi-word strings (such as 'I don't know', 'I don't like', or 'I don't') within utterances (see examples below). Four of these learners used overgeneralised forms as shown in Table 5.27. To capture the syntactic behavior of the copula 'be' as well as NegP development more explicitly, I categorised learners' production into three categories: 'is', 'are', and default 'is'. I also organised their NegP production into four categories based on negation constructions as well as NegP multi-word chunks (e.g., 'I don't know'). This categorisation aimed to help identify the exact OG stage where learners used single function words (e.g., 'for') or multi-word strings (e.g., 'you have') as overgeneralised forms, which were observed in their production. OG stages organise rows in all tables. The table also includes three learners from the NegP stage, but their performance will be discussed further below.

Name	OG Stage	copula 'be'			overgeneralised forms	NegP					NegP chunks
		is	*are*	default *is*		no	not	is+not/isn't	don't	doesn't	
Omar	VP	0	0	0	0	0	0	0	0	0	8
Mansoor	VP	4	0	4	4	0	0	0	0	0	8
Arshad	VP	4	0	4	11	0	0	0	0	0	8
Adeeb	VP	0	0	0	34	8	0	0	0	0	0
Zaid	NegP	4	0	4	7	0	0	0	8	0	0
Gharam	NegP	0	0	0	24	0	0	0	8	0	0
Sameer	NegP	4	2	3	3	0	0	0	6	0	2

Table 5.27: Relationship between the acquisition of copula 'be' and the use of overgeneralised forms

As can be seen in this table, and as the examples in the following sections show, none of these VP-stage learners had fully mastered all the patterns of negation structures. This suggests that the use of overgeneralisation begins to occur during a transitional phase between the VP and NegP stages. Table 5.27 shows a negative relationship between the overgeneralised use of single function words and multi-word strings and the acquisition of forms of the copula 'be'. For example, learners who produced overgeneralised forms have not yet acquired the copula 'be' or used it as a default form (e.g., 'they is students'). In the context of the copula 'be', two learners (Omar and Adeeb) showed no use of the copula verbs ('is' and 'are') and did not substitute them with any lexical items (e.g., the preposition 'for' or multi-word strings), as shown in (5.19). Instead, they omitted the copula entirely.

(5.19) Woman teacher. (Adeeb)

Target utterance: She/The lady is a teacher.

As previously mentioned, the table also includes three NegP-stage learners who used overgeneralised forms in their production data. The analysis shows that these three learners seemed to start to develop their NegP projection, which is indicated by the appearance of the sentential negator 'don't'. Now, let us move on to discussing the examples of these overgeneralised forms. First, I will discuss the occurrence of single function words, followed by multi-word strings in declarative sentences and multi-word strings in the development of negation (first in sentences with verbs and then in verb-less utterances).

5.6.1.1 Overgeneralisation of single function words

In this section, I present representative examples that contain single function words used as overgeneralised forms. I will begin with the first learner at the VP stage (Mansoor), who produced the pronoun 'I' five times in his utterances between the subject and the lexical verb, as in (5.20).

(5.20) This baby *I* drive care. (Mansoor)
Target utterance: The baby drives/is driving a car.

Another example of single function words used as overgeneralised forms is by Adeeb, whose production data includes non-target function words (particularly the preposition 'for'). As in (5.21), the sentences that Adeeb produced in all tasks, apart from the copula 'be', either in declarative or negative contexts, contain single function words.

(5.21) a. The girl *for* drink Coca Cola. (In the VP task) (Adeeb)
Target utterance: The boy drinks/is drinking cola.
b. Children no *for* clean. (In the NegP task)
Target utterance: The boy doesn't clean/is not cleaning the room.
c. Three boys *for* clean dishes. (In the present progressive tense *-ing* task)
Target utterance: Three boys/the boys are cleaning the dishes.
d. Every evening boy *for* clean table. (In the present third-person singular *-s* task)
Target utterance: Every evening, the boy sets the table.
e. Boat *for* water. (In the TP task)

The examples of the non-target function word 'for' in Adeeb's speech fragments do not always, as Lakshmanan (1993/94) claims, seem to be a case assigner, since he used it both when he knows the verb (the case assigner), as in (5.21a), and when he does not know it at all, as in (5.21b). The verb Adeeb uses in some of his sentences already functions as a case assigner.

Gharam, a NegP-stage learner, also used single function words as overgeneralised forms. She overused the pronoun 'his' in nearly all tasks, except for the copula 'be' and the TP tasks, as shown in (5.22). One might think that her use of this pronoun is simply a mispronunciation of the copula form 'is', but she did not use 'his' in the copula 'be'

tasks. She used 'his' three times in combination with the negator 'don't', as in (5.22d). She consistently used the negator 'don't', again as in (5.22d), regardless of the required agreement feature (that is, she overgeneralises 'don't' to 'doesn't').

(5.22) a. The girl *his* eat fish. (In the VP task) (Gharam)
Target utterance: The girl eats/is eating fish.
b. Son *his* play cat. (In the present progressive tense *-ing* task)
Target utterance: The girl is carrying a cat.
c. Man Monday *his* go to shopping. (In the third-person singular *-s* task)
Target utterance: On Mondays, the man goes to the supermarket.
d. Son *his* don't clean house. (In the NegP task task)
Target utterance: The boy does not/is not cleaning the room.

5.6.1.2 Overgeneralisation of multi-word strings

Similar to what has been observed in the literature discussed previously regarding the use of multi-word chunks, three learners at the early stages (VP and NegP) produced them. For example, the VP-stage learner (Arshad) used the phrase 'in the' after the main verb, as in (5.23a), and between the subject and the object, as in (5.23b). The example in (5.23d) clearly shows the exact syntactic position of the 'in the' as it appears before the main verb ('clean').

(5.23) a. The man drown (draw) *in the* brush. (In the VP task) (Arshad)
Target utterance: The boy paints/is painting a chair.
b. The woman *in the* brush in dog. (In the VP task)
Target utterance: The girl is washing/washes a dog.
c. Layla *in the* play box. (In the VP task)
Target utterance: Layla pushes/is pushing a toy box.
d. My children *in the* clean sink. (In the present progressive tense *-ing* task)
Target utterance: The children are cleaning the dishes.

Another instance of chunks comes from Zaid (a NegP-stage learner), who produced the phrase 'you have' in a couple of his utterances before the verb, as in (5.24a), and in his verbless utterances, as in (5.24b). The phrase 'you have' also appears in Sameer's data, as in (5.24c).

(5.24) a. Sarah *you have* eat fish. (In the VP task) (Zaid)

Target utterance: Sarah eats/is eating fish.

b. Layla *you have* horse.

Target utterance: Layla draws/is drawing a horse.

c. The baby *you have* drive car. (In the VP task) (Sameer)

Target utterance: The baby drives/is driving a car.

5.6.1.3 Overgeneralisation of multi-word expressions in the development of NegP

Cancino *et al.* (1978) observed in their study on English negation that their learners used the string 'I don't know' and identified it as monomorphic, therefore excluding it from the analysis. Based on their studies, which aimed to establish a path of negation development, this exclusion is practical because such a chunk would not provide accurate information on the use of 'don't'. However, my analysis shows that there is more to this. Three VP-stage learners (Arshad, Mansoor, and Omar) and one NegP-stage learner (Sameer) produced such chunks in the development of their negation constructions. As the data in Table 5.28 illustrates, these learners used three types of NegP chunks: 'I don't know', 'I don't like', and 'I don't'. The table also shows the frequency of these chunks.

Name	OG Stage	NegP chunks		
		'I don't know'	'I don't like'	'I don't'
Arshad	VP	7	0	1
Mansoor	VP	0	8	0
Omar	VP	6	2	0
Sameer	NegP	0	0	2

Table 5.28: NegP chunks

The rest of this section introduces some representative examples of NegP chunk use. I will begin with the use of the unanalysed multi-word string 'I don't know' produced by two VP-stage learners (Arshad and Omar). These two learners used it between the subject and the verb for negative phrases, as in (5.25).

(5.25) a. Boy *I don't know* clean house. (Arshad)

Target utterance: The boy does not clean/is not cleaning the house.

b. Daughter *I don't know* wash car. (Omar)

Target utterance: The girl does not wash/is not washing the car.

Another example of the production of chunks is the multi-word string 'I don't like', which two VP-stage learners (Mansoor and Omar) used. Again, 'I don't like' was used between the subject and the verb, as in (5.26).

(5.26) This girl *I don't like* clean car. (Mansoor)
Target utterance: The girl does not wash/is not washing the car.

The last example of chunks shows the use of the multi-word string 'I don't'. Sameer used this multi-word string twice before the verb, as in (5.27a). This multi-word string also appeared only once in Arshad's utterances, as in (5.27b).

(5.27) a. Boy *I don't* drive car. (Sameer)
Target utterance: The boy doesn't/is not driving a car.
b. Girl *I don't* clean and car. (Arshad)
Target utterance: The girl doesn't wash/isn't washing the car.

5.6.1.4 Overgeneralisation of NegP chunks in verb-less utterances

Just as in the cases of the production of function words in verb-less utterances, further analysis of the data shows that there are a couple of verb-less utterances produced by learners that contain the unanalysed multi-word string 'I don't know', as in (5.28). These strings often precede the object.

(5.28) a. Boy *I don't know* a car. (Arshad)
Target utterance: The boy doesn't wash/isn't washing the car.
b. This boy *I don't like* house. (Mansoor)
Target utterance: The boy does not clean/is not cleaning the room.

The words 'know' and 'like' in the strings 'I don't know' and 'I don't like' in verb-less utterances do not replace the verb; in other words, they are not the verb of the sentence because Mansoor used verbs after the phrase 'I don't like', as in (5.26), similar to how Adeeb used or omitted the main verb. So the 'I don't like' in his sentences is another indication that some learners used chunks as a strategy in the development of their morphosyntax. In his declarative sentences, he only produced one example that contains the multi-word string 'I like', as in (5.29).

(5.29) This boy *I like* games. (In the VP task) (Mansoor)
Target utterance: The boy plays/is playing games.

In sum, these learners have started to figure out the head of NegP. This is indicated by the appearance of the particle 'no' or the unanalysed expressions ('I don't know', 'I don't like' or 'I don't'). The majority of these learners produced multi-word strings as overgeneralised forms, whose syntactic slots often precede the main verbs. Now we will move on to overgeneralisation at higher stages of Organic Grammar.

5.6.2 Overgeneralisation at higher stages of Organic Grammar

According to the analysis in the preceding chapter, some learners from the NegP, TP, and AgrP stages used overgeneralised 'be' and/or personal subject pronouns either in the same context or in different contexts. Unlike the NegP-stage learners who used single function words or multi-word strings, these NegP-stage learners are more advanced, as indicated by their consistent use of the copula 'be' and their very limited use of auxiliary 'be'. The learners at the TP stage, in addition to their consistent use of the copula 'be', started to use auxiliary 'be' and, to a very limited extent, used the past tense or the third-person singular *-s*

5.6.2.1 Be-production

Table 5.29 details learners' performance in morphosyntax in the production tasks and the occurrence of 'be' overgeneralisation. As shown in this table, seven learners from the NegP stage, four from the TP stage, and two from the AgrP stage used the ungrammatical *be*-forms (usually 'is' and sometimes 'are') before the lexical verbs.

Name	OG Stage	TP(-ed)	AgrP			Be-production
			Aux be		3- sg -s	
			is	are		
Waleed	NegP	0/8	2/4	0/4	0/8	13/84
Nihad	NegP	0/8	0/4	4/4	0/8	14/84
Hind	NegP	0/8	4/4	4/4	0/8	9/84
Itab	NegP	0/8	0/4	0/4	0/8	13/84
Reham	NegP	0/8	2/4	4/4	0/8	17/84
Talal	NegP	0/8	0/4	0/4	0/8	3/84
Afnan	NegP	3/8	0/4	0/4	0/8	9/84
Maher	TP	0/8	0/4	1/4	0/8	11/84
Jasim	TP	4/8	1/4	4/4	0/8	4/84
Rabab	TP	1/8	3/4	3/4	1/8	18/84
Rihaab	TP	2/8	4/4	0/4	0/8	4/84
Nawal	AgrP	6/8	4/4	4/4	4/8	7/84
Zain	AgrP	8/8	4/4	4/4	8/8	2/84

Table 5.29: Be-production

In the rest of this section, I present some representative examples of the overgeneralisation of *be*-forms used before a lexical verb. For example, the NegP-stage learner Nihad used 'is' before the main verb, as in (5.30a). Another example comes from Maher, again at the NegP stage, who used 'are' before the main verb, as in (5.30b).

(5.30) a. The baby *is* drive a car. (In the VP task) (Nihad)
Target utterance: The baby drives/is driving a car.
b. The kids *are* play game. (In the VP task) (Maher)
Target utterance: The kids are playing video games.

In some cases, the production of 'is' before the verb seems to be an omission of *-ing*. For example, Rihaab displayed sufficient complete forms of the present progressive tense; however, she used this tense only once, as in (5.31a) regardless of its functions. The context in the example below required simple present tense, so third-person singular *-s*.

(5.31) a. Every morning, this boy usually *is* drinking coffee. (In the third-person *-s*) (Rihaab)
Target utterance: Every morning, this boy drinks coffee.

Rabab, who is at the TP stage, appeared to use the *be*-forms ('is' or 'are') as an overgeneralised form. She used it in the context of the present progressive tense *-ing*, as

in (5.32), before the light verb 'to have' (has or have). This tendency can be considered formulaic language for her, or she may have simply used these verbs to indicate possession in the present tense (e.g., 'She has a cat'). If so, she used *be*-forms ('is' or 'are') as an overgeneralised form because no missing *-ing* occurred.

(5.32) a. Girl *is has* cat playing. (In the present progressive tense *-ing* tasks) (Rabab)
Target utterance: The girl plays/is playing with the dog.

b. The children *are have* dog washing/cleaning. (In the present progressive tense *-ing* tasks)
Target utterance: The children wash/are washing the dog.

In the context of the NegP tasks, Rabab used 'is' twice before the negator 'don't', as in (5.33a). In these examples, the use of 'is' seemed to be an overgeneralised form that Rabab might have used to mark agreement. She did not use 'doesn't' in her utterances. She used 'no' as a sentential negator twice, as in (5.33b).

(5.33) a. boy *is don't* drive car. (Rabab)
Target utterance: The boy doesn't drive/is not driving the car.

b. Girl *is no* door painter.
Target utterance: The girl does not/is not painting the door.

5.6.2.2 Overgeneralisation of personal pronouns

Table 5.30 presents learners' performance according to the production tasks and lists the frequency of using personal pronouns or a combination of personal pronouns and *be*-forms as overgeneralised forms. The analysis also showed that five learners from the NegP stage, one learner from the TP stage, and two from the AgrP stage used subject personal pronouns before lexical verbs, or they used a combination of subject personal pronouns and *be*-forms before lexical verbs. However, the occurrence of a combination of personal subject pronouns and *be*-forms is very low compared to the occurrence of *be*-forms or the occurrence of personal pronouns as overgeneralised forms.

Some of these learners used both types of overgeneralisation. For example, they sometimes used *be*-forms and sometimes used pronouns, either in the same context or in different contexts.

Data analysis shows that eight learners produced non-target personal pronouns ('he', 'she', and 'they'), which are used before lexical verbs. As Table 5.30 shows, these learners, apart from Andleeba, are at the TP or AgrP stages of OG. The table also shows the combined use of personal pronouns (e.g., 'The boy *he is* clean the room') per learner.

Name	OG Stage	Frequency of personal pronouns			TP (-ed)	AgrP			Personal pronouns + copula
						Aux be		3- sg -s	
		He	She	They		is	are		
Maher	NegP	0/84	0/84	0/84	0/8	0/4	1/4	0/8	8/84
Jaber	NegP	11/84	12/84	1/84	0/8	0/4	1/4	0/8	0/84
Waleed	NegP	0/84	1/84	0/84	0/8	2/4	0/4	0/8	0/84
Andleeba	NegP	12/84	13/84	4/8	1/8	0/4	0/4	0/8	0/84
Afnan	NegP	6/84	4/84	0/84	4/8	0/8	0/8	0/8	4/84
Jasim	TP	0/84	0/84	0/84	0/8	1/4	4/4	0/8	1/84
Zain	AgrP	0/84	0/84	0/84	8/8	4/4	4/4	8/8	1/84
Nawal	AgrP	7/84	3/84	0/84	6/8	4/4	4/4	4/8	4/84

Table 5.30: Overgeneralisation of personal pronouns

Similar to the learners at the early stages of OG, Andleeba used the copula 'is' and the pronoun 'she' as default forms (e.g., 'She is two students') but did not use 'no' as a negator for NegP or multi-word strings such as 'I don't know'. Instead, she used 'don't', even though she generalised 'don't' to 'doesn't', which might be why she used the pronoun 'he' as an overgeneralised form, as in (5.34), instead of the preposition 'for', the conjunction 'and', or multi-word strings (e.g., 'you have'). This usage also indicates that the stage of Organic Grammar affects the choice of overgeneralised forms.

(5.34) a. The boy *he* don't drinking the milk. (Andleeba)

Regarding her morphosyntax, she appeared to struggle and did not produce any functional morphology other than *-ing*, albeit with the missing auxiliary 'be' form. In response to the present progressive tense *-ing*, she supplied 5 out of 10, as in (5.35). Her use of the *-ing* form did not seem to be an overgeneralisation because she used it only with the present progressive tense. In example (5.35b), she seemed to use the pronoun 'they' to agree with the plural subject.

(5.35) a. The girl *she* reading the book. (Andleeba)
Target utterance: The girl is reading a book.

b. The children *they* clean the dishes.
Target utterance: The girls are cleaning the dishes.

The example in (5.36) shows that she did not produce the plural-noun marker *-s* in obligatory contexts; however, she used the pronoun 'they'.

(5.36) The girl *they* cooking the food. (In the present progressive tense *-ing* task)
Target utterance: The girls are cooking the food.

There are five learners who used the personal pronouns and *be*-forms at the same time. For example, Afnan, who is at the NegP stage, used the pronoun 'he' and the verb 'is' in her sentences, as in (5.37). The use of 'is', as in (5.37b), can be interpreted as an incomplete form of the progressive tense (e.g., 'Emily is having her breakfast'); however, the context in the testing was about the third-person singular *-s*. A second possible interpretation is that 'is' and 'he' have different syntactic functions. The 'is' marks agreement, whereas the pronoun 'he' marks gender. In Arabic, lexical verbs are inflected for person, number, and gender. If this is correct, it means that L1 transfer occurs in the intermediate stages (Hawkins 2001). I will return to this in the following chapter.

(5.37) a. The boy *he is* clean his bedroom. (In the VP tasks) (Afnan)
Target utterance: The boys/They are cleaning the dishes.
b. Emily *she is* have her breakfast. (In the third-person singular *-s* tasks)
Target utterance: Emily eats her breakfast.

5.7 Comprehension data and overgeneralised forms

As presented in Chapter 4, the methodology included comprehension data obtained through computerised tasks (audio-picture matching questions). The purpose of these tasks is to test Vainikka and Young-Scholten's (2019) claim about overgeneralisation in the acquisition of morphosyntax in adult L2 learners. Under Organic Grammar, adult L2 learners acquire their L2 morphosyntax in a systematic manner. As a result of the interaction between Universal Grammar and the input, L2 learners first go through subsequent stages, beginning with the VP, followed by the NegP, then the intermediate stages (TP and AgrP), and finally the CP. The L2 learners' task is to search for the relevant syntactic head provided by Universal Grammar. Once the learners identify the relevant

head in the input, the syntactic projection is built for that head. According to Vainikka and Young-Scholten, failure to identify the relevant head may result in using non-target elements (e.g., overgeneralised forms) other than the target head. For example, Vainikka and Young-Scholten presume that overgeneralised forms are placeholders for the head of TP (i.e., the past simple marker *-ed*) or the head of AgrP (i.e., the third-person singular *-s*).

As we saw in Section 5.4, the learners struggle more with the comprehension of bound morphemes than with free morphemes. This finding alone does not confirm whether there is a relationship between the difficulty in comprehending bound morphemes and overgeneralisation. Therefore, this section aims to explore the potential relationship between the difficulty in comprehending bound morphemes and overgeneralisation.

As discussed above, overgeneralised forms in this thesis can be classified into three categories: the use of single function words (e.g., 'for' or multi-word sequences), *be*-forms, and subject pronouns. An investigation was conducted to explore the relationship between each bound morpheme and each occurrence of these overgeneralised forms. This investigation tests the hypothesis that learners who engage in overgeneralisation struggle with bound morphemes.

Additionally, the relationship between these overgeneralisations and free morphemes was examined to assess whether this relationship is reflected in learners' production data on overgeneralisation. As noted above, learners who use *be*-forms and personal pronouns as overgeneralised forms tend to use IP-related elements, especially the copula and auxiliary 'be'.

The analysis in this section included an investigation of the mean and SD values for the occurrence of each type of overgeneralisation in relation to bound morphemes. Specifically, I examine whether there is a negative relationship between the mean values and the increase in learners' scores for bound morphemes, as expected. Additionally, I will examine the mean values for the occurrence of each type of overgeneralisation in relation to learners' scores for free morphemes, with variability being addressed through the SD.

Spearman's rank correlation was conducted to detect any significant relationship between these morphemes and the occurrence of overgeneralisations. As mentioned in Chapter 4, Spearman's rank correlation is suitable when dealing with ordinal variables or when

the assumptions of Pearson correlation are not met. In our case, while the variables were continuous, the assumption of linearity was violated, necessitating the use of Spearman correlation. To better align with the ordinal nature of the data, the variables were ranked using SPSS.

This section is divided into three subsections, each presenting the analysis of the relationship between verbal morphemes (bound morphemes and free morphemes) and one type of overgeneralisation. Each subsection first examines the relationship between bound morphemes and overgeneralisation, followed by the relationship between free morphemes and overgeneralisation.

5.7.1 Overgeneralisation of single function words or multi-word strings and verbal morphemes

This section presents the results of using overgeneralisation of single function words or multi-word strings in relation to the acquisition of bound morphemes in the comprehension tasks. Then, it will present the results of using overgeneralisations in relation to the acquisition of free morphemes. As seen in the results of the production data, learners who used overgeneralisations exhibited a very low level of morphosyntax. Here, I expect that there is a mismatch between learners' scores in the acquisition of free or bound morphemes and the use of overgeneralisations, specifically with bound morphemes, as hypothesised in the methodology chapter.

5.7.1.1 Overgeneralisation of single function words or multi-word strings and bound morphemes

As we can see in Table 5.31, the mean value of the occurrence of this type is higher when learners' scores decrease in both the regular past tense *-ed* and the third-person singular *-s*. However, the SD shows high variability compared to the mean, indicating that learners with lower scores (e.g., 0 or 1) vary in using or not using this type of overgeneralisation.

Overgeneralisation of single function words or multi-word strings in relation to the regular past simple *-ed*			
Past simple *-ed*	N	Mean	SD
0	16	1.81	1.76
1	22	1.41	1.22
3	17	1.00	0.00
4	5	1.00	0.00
Overgeneralisation of single function words or multi-word strings in relation to the third-person singular *-s*			
Third-person *-s*	N	Mean	SD
0	20	1.9	1.92
1	17	1.24	0.56
2	13	1.00	0
3	10	1.00	0

Table 5.31: Mean and SD for overgeneralisation of single function words or multi-word strings in relation to bound morphemes

This variation is expected because the learners exhibit noticeable variation in the use of this type in the production tasks in this thesis. For example, as we saw in the production data in this thesis, not all learners at the same stages (e.g., VP-stage learners) used overgeneralised forms. This observation has also been reported in the literature on L2 lower-educated learners (e.g., Mocciaro 2019b) and on L1 children with respect to the overgeneralisation of the simple past tense *-ed* (e.g., Berko 1958).

On the other hand, the SD demonstrates that learners with higher scores (e.g., 2 or 3) show consistency in not using this type of overgeneralisation. In other words, when learners achieve higher scores in the comprehension of these two bound morphemes, the likelihood of avoiding this type of overgeneralisation decreases.

Figure 5.16 clearly reflects this finding. As shown in the scatter plots, there is a negative relationship: Lower scores are associated with a higher frequency of this type of overgeneralisation.

(a) Overgeneralisation of single function words or multi-word strings and the regular past tense *-ed*

(b) Overgeneralisation of single function words or multi-word strings and the third-person singular *-s*

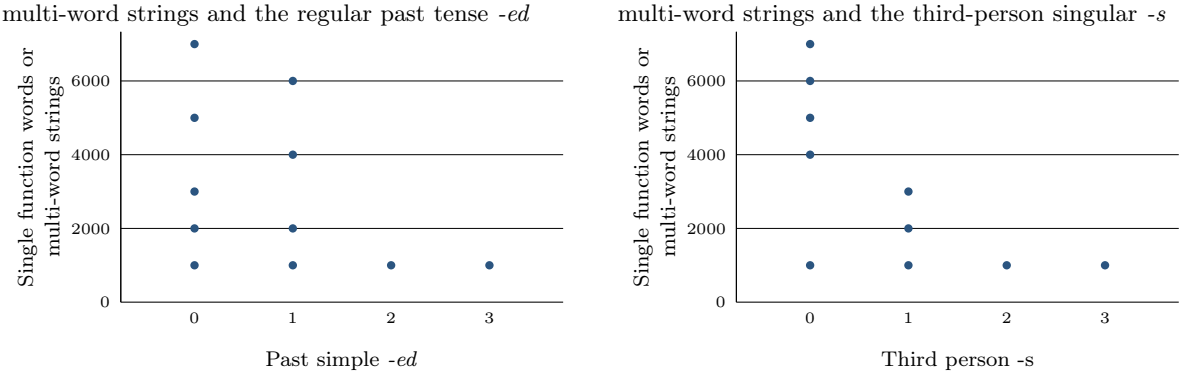


Figure 5.16: Relationship between overgeneralisation of single function words or multi-word strings and bound morphemes

The statistical analysis in Table 5.32 confirms these negative relationships by indicating a strong negative correlation between the past simple *-ed* and the occurrence of this type of overgeneralisations, as well as between the third-person singular *-s* and the occurrence of this type of overgeneralisations. The p-value is less than 0.05, confirming the hypothesis in Chapter 4 that learners who struggle with the comprehension of bound morphemes are more likely to use overgeneralisations.

Bound morphems	Correlation	P-value
Past simple *-ed*	-0.276	0.019
Third-person *-s*	-0.252	0.031

Table 5.32: Correlations between overgeneralisation of single function words or multi-word strings and bound morphemes)

Now let us move on to discuss the results of the relationship between this type of overgeneralisation and learners' comprehension of free verbal morphemes. More particularly, we will see if there is any pattern reflecting learners' use of overgeneralisation in production tasks.

5.7.1.2 Overgeneralisation of single function words or multi-word strings and free morphemes

According to Table 5.33, the mean value of using single function words or multi-word strings as overgeneralised forms shows varying patterns. For example, the mean value of these forms based on the copula 'be' exhibits fluctuation, making it difficult to discern positive or negative patterns. This fluctuation could be attributed to learners' Organic Grammar stages. Specifically, learners in the VP stage or at the entry of the subsequent OG stage (NegP) used these forms in their production and simultaneously exhibited fluctuation in the use of the copula 'be' in their production, as noted in this thesis.

Overgeneralisation of single function words or multi-word strings in relation to the copula 'be'			
Cop 'be'	N	Mean	SD
0	8	2.00	2.14
1	10	1.00	0.00
2	15	1.40	1.06
3	27	1.30	1.10
Overgeneralisation of single function words or multi-word strings in relation to the auxilary 'be'			
Aux 'be'	N	Mean	SD
0	8	1.00	0.00
1	12	1.67	1.78
2	19	1.47	1.12
3	21	1.24	1.09

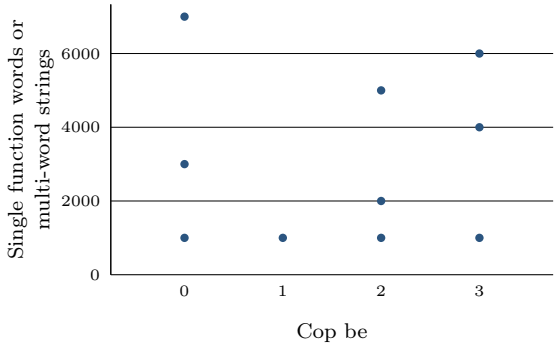
Table 5.33: Mean and SD for overgeneralisation of single function words or multi-word strings in relation to free morphemes

On the other hand, the mean value of using these forms in relation to the auxiliary 'be' demonstrates a decreasing trend, particularly from scores of 1 to 3. This finding could suggest that as learners' scores in the auxiliary 'be' increase, the occurrence of these forms decreases. This reflects learners' overgeneralisation of this type in production data. The learners who used this type did not produce the auxiliary 'be' at all. This analysis could be valid if we consider that the acquisition of the agreement feature of the auxiliary 'be' in this thesis occurs later, beyond the lower stages (VP or NegP) of Organic Grammar.

Additionally, this finding indicates that when these learners start to comprehend the auxiliary 'be', they may eliminate using single function words or multi-word strings as overgeneralised forms or possibly switch to using *be*-forms as overgeneralised forms, similar to the learners who used *be*-forms as overgeneralised forms and used some auxiliary 'be' in a target-like way. Let us see if this holds true when discussing the mean value of using *be*-forms as an overgeneralised form in relation to learners' comprehension of the auxiliary 'be'.

The scatter plots in Figure 5.17 do not reflect the finding of the mean value. On the other hand, these plots show a subtle positive relationship between learners' scores in the comprehension of these two morphemes and the use of this type of overgeneralisation. However, examining the graphs more closely, we can observe a positive relationship. At the same time, outliers are evident; the highest occurrence of these forms is observed with the lowest scores (0 or 1), which could affect the potential positive relationship.

(a) Overgeneralisation of single function words or multi-word strings and the copula 'be'



(b) Overgeneralisation of single function words or multi-word strings and the auxiliary 'be'

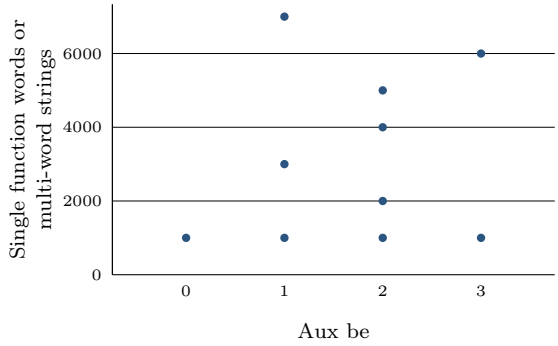


Figure 5.17: Relationship between overgeneralisation of single function words or multi-word strings and free morphemes

If there is a positive relationship, this could indicate that learners who use these forms appear to be in the early process of developing comprehension of inflectional morphemes, such as the copula 'be' or the auxiliary 'be', but are not yet able to produce these morphemes. Simultaneously, they struggle with the comprehension of bound morphemes, thereby having difficulty identifying syntactic heads (*-ed* for TP or *-s* for AgrP) in the input and use non-target elements to fill these syntactic heads. If this analysis is accurate, this finding supports the idea that overgeneralisation occurs at a stage after the VP has been established, since the agreement features of the copula 'be' or auxiliary 'be' are acquired beyond the bare VP stage of Organic Grammar.

However, the analysis in Table 5.34 indicates that there is no correlation between the overgeneralisation of single word or multi-word strings and each of these free morphemes, as evidenced by the p-value being greater than 0.05.

Morpho-syntactic elements	Correlation	P-value
Cop 'be'	-0.103	0.432
Aux 'be'	-0.044	0.736

Table 5.34: Correlations between overgeneralisation of single function words or multi-word strings and free morphemes

5.7.2 Overgeneralisation of 'be' and verbal morphemes

This section presents the results of using overgeneralisation of 'be' in relation to the acquisition of bound morphemes in the comprehension tasks. Then, it will present the results of using overgeneralisations in relation to the acquisition of free morphemes. As seen in the results of the production data, learners who used overgeneralisations exhib-

ited a moderate level of morphosyntax. Therefore, I expect that there is a negative relationship between learners' use of this type of overgeneralisation and their scores for bound morphemes, as hypothesised. In contrast, there is no relationship or a positive, but not negative, relationship between learners' use of this type of overgeneralisation and their scores in comprehension for free morphemes, as the learners who used *be*-forms as overgeneralised forms also use target-like *be*-forms.

5.7.2.1 Overgeneralisation of 'be' and bound morphemes

The descriptive statistics in Table 5.35 show no clear relationship between the overgeneralisation of 'be' and learners' scores on the comprehension of these morphemes, as indicated by the mean, which shows observed fluctuations. However, a closer examination reveals that the highest mean value of learners' overgeneralisation of 'be' is observed with their highest scores for the regular past simple *-ed*. This indicates no relationship between the difficulty of the regular past simple *-ed* in comprehension and the occurrence of overgeneralisation of 'be'. The high SD also suggests that while learners may score high in comprehension, they can vary significantly in their use of overgeneralisations. In other words, even learners with high comprehension scores may still overgeneralise the use of *be*-forms.

Overgeneralisation of 'be' in relation to the regular past simple *-ed*			
Past simple *-ed*	N	Mean	SD
0	16	2.25	2.74
1	22	1.77	1.85
2	17	2.82	3.30
3	5	3.60	3.21
Overgeneralisation of 'be' in relation to the third-person singular *-s*			
Third-person *-s*	N	Mean	SD
0	20	3.95	3.58
1	17	1.41	1.00
2	13	2.15	2.51
3	10	1.00	0.00

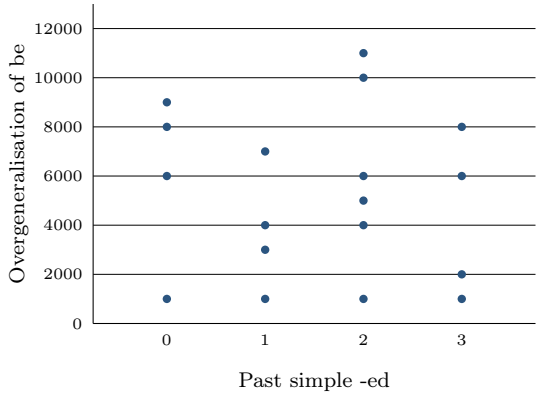
Table 5.35: Mean and SD for overgeneralisation of 'be' in relation to bound morphemes

On the other hand, the highest mean values for the overgeneralisation of 'be' are observed with the lowest scores for the third-person singular *-s*. This suggests a potential

relationship between the difficulty of the third-person singular *-s* and the occurrence of this type of overgeneralisation. Additionally, the SD indicates that learners vary in their use of 'be' as an overgeneralised form when their scores are very low. However, when their scores are higher, they are more likely to avoid using these forms.

The visual presentations in Figure 5.18 also assume these assumptions. While Figure 5.18a indicates no relationship, Figure 5.18b illustrates a clear negative relationship between the occurrence of 'be' forms as overgeneralised forms and learners' scores in comprehension.

(a) Overgeneralisation of 'be' and the regular past simple *-ed*



(b) Overgeneralisation of 'be' and the third-person singular *-s*

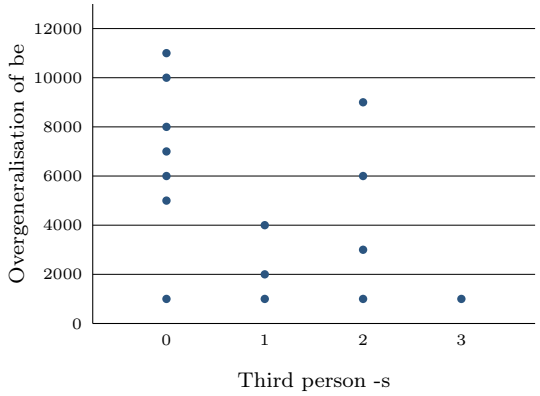


Figure 5.18: Relationship between overgeneralisation of 'be' and bound morphemes

Spearman analysis confirms this relationship. As Table 5.36 shows, there is a strong negative correlation between the third-person singular *-s* and the occurrence of overgeneralisation of 'be', with a p-value of 0.01. On the other hand, there is no correlation between the past simple marker *-ed* and the overgeneralisation of 'be', indicating no connection between these two variables.

Bound morphems	Correlation	P-value
Past simple *-ed*	.185	.156
Third-person *-s*	-.356	.005

Table 5.36: Correlations between overgeneralisation of 'be' and bound morphemes

5.7.2.2 Overgeneralisation of 'be' and free morphemes

Table 5.37 shows that a higher frequency of overgeneralisation of 'be' is observed with the lowest scores in the copula or auxiliary 'be'. In contrast, a lower frequency is observed with the highest scores. This finding indicates that when learners' comprehension of these

morphemes increases, the likelihood of using 'be' as an overgeneralised form decreases. However, the high variability of the SD in relation to the mean does not support this inverse relationship.

Overgeneralisation of 'be' in relation to the copula 'be'			
Cop 'be'	N	Mean	SD
0	8	3.63	3.78
1	10	1.30	0.95
2	15	1.53	2.07
3	27	2.81	2.86
Overgeneralisation of 'be' in relation to the auxiliary 'be'			
Aux 'be'	N	Mean	SD
0	8	4.00	4.17
1	12	2.08	2.94
2	19	2.16	2.19
3	21	2.05	2.13

Table 5.37: Mean and SD for overgeneralisation of 'be' in relation to free morphemes

The data representations in Figure 5.19 also do not show any potential relationship, indicating that the occurrence of overgeneralisation of 'be' does not change noticeably with learners' scores in the comprehension of these two morphemes.

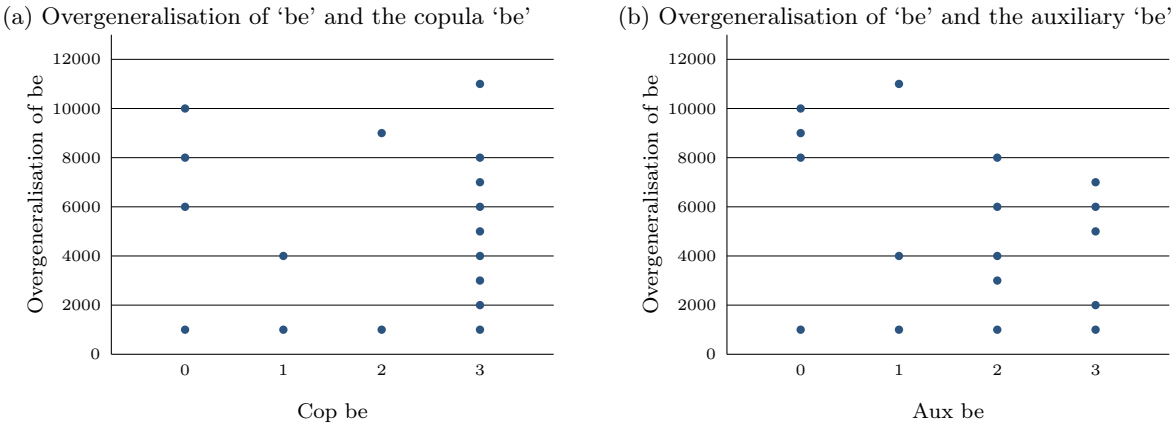


Figure 5.19: Relationship overgeneralisation of 'be' and free morphemes

Analysis of Spearman's correlation in Table 5.38 detected no meaningful relationship between learners' use of overgeneralisations of 'be' and learners' comprehension of these two free morphemes.

Free morphemes	Correlation	P-value
Cop 'be'	0.128	0.330
Aux 'be'	-0.078	0.555

Table 5.38: overgeneralisation of 'be' and free morphemes

5.7.3 Overgeneralisation of subject personal pronouns and verbal morphemes

This section presents the results of using overgeneralisation of subject personal pronouns in relation to the acquisition of bound morphemes in the comprehension tasks. Then, it will present the results of using overgeneralisations of this type in relation to the acquisition of free morphemes. As seen in the results of the production data, learners who used overgeneralisation of subject personal pronouns exhibited a moderate level of morphosyntax. Therefore, I expect that there is a negative relationship between learners' use of this type of overgeneralisation and their scores for bound morphemes, as hypothesised. In contrast, there is no relationship or a positive, but not negative, relationship between learners' use of this type of overgeneralisation and their scores in comprehension for free morphemes, as the learners who used subject personal pronouns as overgeneralised forms also use target-like *be*-forms.

5.7.3.1 Overgeneralisation of subject personal pronouns and bound morphemes

The mean values of learners' overgeneralisation of subject personal pronouns in relation to their scores in the regular past tense *-ed* as well as their scores in the third-person singular *-s* show an inverted pattern compared to the relationship between the other two types of overgeneralisation and learners' scores in these two bound morphemes. A higher frequency of overgeneralisation is observed with the highest scores for these morphemes. However, there are fluctuations in the mean values of learners' overgeneralisation of 'be' in relation to their scores in the regular past tense *-ed* but not with the relationship between overgeneralisation of subject personal pronouns and learners' scores in the third-person singular *-s*.

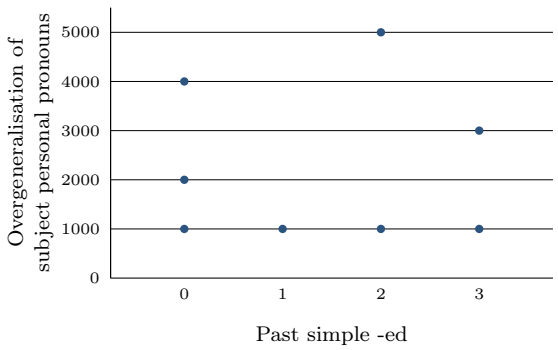
Overgeneralisation of subject personal pronouns in relation to the regular past simple *-ed*			
Past simple *-ed*	N	Mean	SD
0	16	1.25	0.77
1	22	1.00	0.00
2	17	1.24	0.97
3	5	1.40	0.89
Overgeneralisation of subject personal pronouns in relation to the third-person singular *-s*			
Third-person *-s*	N	Mean	SD
0	20	1.05	0.22
1	17	1.12	0.49
2	13	1.23	0.83
3	10	1.40	1.26

Table 5.39: Mean and SD for overgeneralisation of subject personal pronouns in relation to verbal bound morphemes

This finding suggests that the difficulty of bound morphemes in comprehension is not associated with learners' use of overgeneralisations. Instead, high levels of comprehension of these morphemes are positively associated with a high frequency of overgeneralisation of pronouns. However, the standard deviation shows significant variability, especially for higher frequencies of overgeneralisation of pronouns in relation to higher scores. Therefore, individual differences among learners with higher scores could influence this potential positive relationship between these two morphemes and learners' use of overgeneralisation.

This positive relationship between the occurrence of overgeneralisation of 'be' and learners' scores in the third-person singular *-s* is clearly shown in Figure 5.20b. On the other hand, Figure 5.20a shows no relationship between the regular past tense *-ed* and overgeneralisation of personal subject pronouns.ovegernisation of pesrional subject pronouns.

(a) Overgeneralisation of subject personal pronouns and the regular past tense *-ed*



(b) Overgeneralisation of subject personal pronouns and the third-person singular *-s*

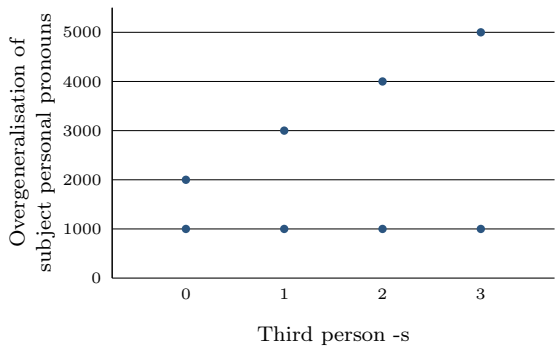


Figure 5.20: Relationship between overgeneralisation of subject personal pronouns and bound morphemes

Despite the variability in the SD compared to the mean, which is shown in the graph, this positive relationship corroborates learners' performance in the production data in this thesis. Learners who use personal pronouns as overgeneralised forms show some evidence that they have acquired IP-related elements (the copula 'be' or auxiliary 'be') and, to a lesser extent, the regular past tense or the third-person singular *-s*. However, Spearman correlation indicates that there is no association between these bound morphemes and this type of overgeneralisation. As shown in Table 5.40, no significant correlation is detected between these bound morphemes and the use of personal pronouns as overgeneralised forms. The p-value exceeds the significance level of 0.05.

Bound morphemes	Correlation	P-value
Past simple *-ed*	.004	.156
Third-person *-s*	.078	.553

Table 5.40: Relationship between overgeneralisation of subject personal pronouns and bound morphemes

5.7.3.2 Overgeneralisation of subject personal pronouns and free morphemes

Similarly, the highest mean value of using this type of overgeneralisation in relation to the copula 'be' and auxiliary 'be', as shown in Table 5.41, is observed with the highest scores in these two morphemes. Higher variability is also observed with higher means, indicating individual differences in using this type of overgeneralisation.

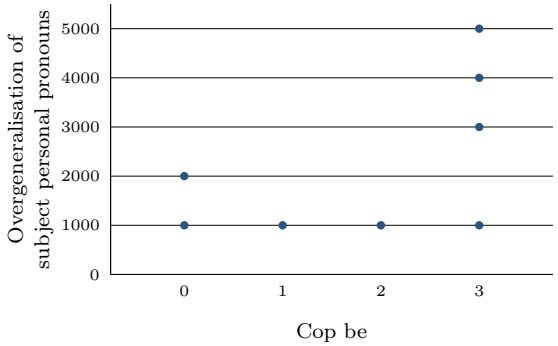
Overgeneralisation of subject personal pronouns in relation to the copula 'be'			
Cop 'be'	N	Mean	SD
0	8	1.13	0.35
1	10	1	0
2	15	1	0
3	27	1.33	1
Overgeneralisation of subject personal pronouns in relation to the auxiliary 'be'			
Aux 'be'	N	Mean	SD
0	8	1.13	0.35
1	12	1	0
2	19	1	0
3	21	1.43	1.12

Table 5.41: Mean and SD for overgeneralisation of subject personal pronouns and free morphemes

This finding indicates a positive relationship between overgeneralisation and learners' comprehension in the use of free morphemes.

The scatter plots in Figure 5.21 also support these relationships. As we can see, there are positive relationships between learners' scores in free morphemes in comprehension and their use of overgeneralisation of personal pronouns.

(a) Overgeneralisation of subject personal pronouns and the copula 'be'



(b) Overgeneralisation of subject personal pronouns and the auxilary 'be'

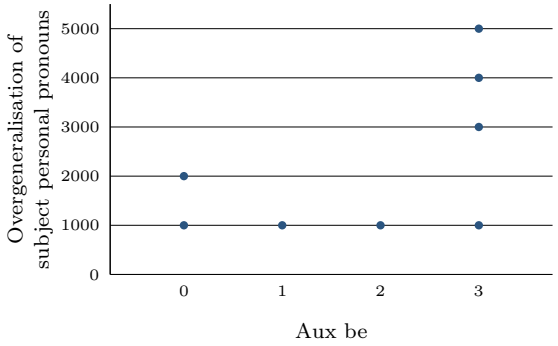


Figure 5.21: Relationship between overgeneralisation of subject personal pronouns and free morphehms

However, Spearman correlation analysis shows no significant correlation between learners' scores of these free morphemes in comprehension and the overgeneralisation of personal pronouns. As shown in Table 5.42, the p-values are higher than the significance level of 0.05.

free morphemes	Correlation	P-value
Cop 'be'	.105	.425
Aux 'be'	.140	.286

Table 5.42: Relationship between overgeneralisation of subject personal pronouns and free morphemes

5.8 Summary

In this chapter, I have investigated whether the participants in this study follow the predicted route of acquisition under Organic Grammar through production and comprehension data. This was accomplished using specific criteria for implicational analysis, aimed at detecting developmental stages based on the predicted stages of Organic Grammar. I also included comprehension tasks to gain further insight into the path of acquisition. Specifically, I examined whether the learners performed similarly in the accuracy of comprehension tasks and production tasks. This involved investigating the acquisition of comprehension of free verbal morphemes and bound verbal morphemes, both of which are important factors in establishing OG stages. The results of the production data

show that learners follow the predicted OG stages, and the comprehension data reflect this route in that (1) learners' scores for free morphemes are higher in frequency than those for bound morphemes. This variation has been confirmed by a Friedman test, with Dunn's test used as a post hoc test. (2) While learners' scores in free morphemes do not show a strong correlation with OG stages, bound morphemes do, as indicated by Spearman.

Exploring the potential relationship between literacy and other linguistic factors, the results indicate that literacy levels and native language schooling are strongly and positively correlated with the development of morphosyntax. However, the other non-linguistic factors showed no significant relationship. Ordinal regression analysis reveals that it is literacy, specifically Literacy 1 and 2 levels, rather than other factors, that delays progression between OG stages.

To gain a deeper understanding of the role of literacy in the acquisition of morphosyntax, this study, though not primarily focused on this aspect, investigates the relationship between the copula 'be' as a potential trigger, other verbal morphology (the auxiliary 'be', the regular past tense *-ed*), and negation, as well as the relationship between the copula 'is' and literacy in both production and comprehension data. In the production data, the results show that the copula 'is' correlates with all these variables, while in the comprehension data, the copula 'is' only correlates with the auxiliary 'is' and the regular past tense *-ed*. The results also show a positive significant relationship between the copula 'be' and literacy in both production and comprehension data.

In the rest of this chapter, I have examined the use of overgeneralised forms as a secondary aim, using Organic Grammar's structure-building approach. The goals were, first, to identify the specific OG stage at which learners begin using these forms and to determine whether OG stages influence the choice of these forms, as they may reflect the learners' developmental stage, and second, to investigate whether there is a correlation between the use of overgeneralised forms and the comprehension of morphosyntax. The aim of the correlation analysis was to determine whether the use of overgeneralised forms is related to learners' comprehension of morphosyntax. The results of the production data show that overgeneralisation begins to occur during the transitional phase between the VP and NegP stages, with learners using different forms of overgeneralisation based on their OG stages. The results of the comprehension data indicate that while learners'

use of single function words or multi-word strings as overgeneralised forms correlates with the regular past tense *-ed* or the third-person singular *-s*, learners' use of 'be' as overgeneralised forms correlates only with their scores in the third-person singular *-s*. On the other hand, learners' use of overgeneralised forms, whether single or multi-word strings, 'be', or personal pronouns, shows no correlation with the use of the regular past tense *-ed* or the third-person singular *-s* in comprehension.

Moving on to the discussion of these results, which will be examined in relation to the route of acquisition, including the role of literacy and overgeneralisation.

Chapter 6. Discussion

6.1 Introduction

This chapter is divided into three sections. Section 6.2 discusses the results of morphosyntax in terms of (1) the route of morphosyntax acquisition and (2) the role of literacy in the development of morphosyntax. As in previous studies, the route of acquisition appears to be consistent among all learners in this current study. In line with the existing literature on low-educated learners, literacy appears to affect the rate of language development. Therefore, this section will explore a new possible explanation for how literacy could affect the development of morphosyntax, drawing on insights from the distinction between 'learned linguistic knowledge' (Schwartz 1993) and linguistic competence, as well as the distinction between language learning and language acquisition (Krashen 1985). In Section 6.3, I will discuss the choice of overgeneralised forms in relation to the OG stages, which appear to influence the choice of these forms. As mentioned in the preceding chapter, this study aims to offer a new perspective on the relationship between stages of Organic Grammar and the selection of overgeneralised forms. In this section, I will also discuss the results in relation to the concepts of systematicity and creativity, which were previously hinted at by Mocciaro (2019b) and Vainikka and Young-Scholten (2019). However, this section goes one step further by providing additional evidence. This includes examining the syntactic slot for these forms and noting that no lexical material is ever inserted between these overgeneralised forms and the verbs, along with new insights from the analysis of comprehension data. Furthermore, this section contributes to our understanding of systematicity and creativity by providing more evidence through a comparison of my data with overgeneralisation in L1 acquisition. Before closing this section, I will present the use of multi-word strings as placeholders versus Exemplar as chunks. This contrast provides us with new perspectives on the role of overgeneralisation of chunks in the development of morphosyntax. I also suggest possible reasons why VP-

stage learners use multi-word strings as placeholders. Section 6.4 will discuss whether the use of subject personal pronouns as overgeneralised forms is a result of L1 transfer. This section raises some caution about the existing tentative interpretations because my data reveal different patterns in terms of how subject pronouns are used as placeholders.

6.2 The route of acquisition and the role of literacy in the development of morphosyntax

The first aim of this study is to investigate the route of acquisition by low-educated/low-literate learners within the framework of Organic Grammar, according to which the acquisition of morphosyntax proceeds through stages. The acquisition process begins with the bare VP stage, which consists solely of lexical elements, followed by the NegP stage, intermediate stages (TP and AgrP), and eventually the CP stage. As previously discussed in this thesis, the first question I seek to answer is: Do adult immigrants with varying levels of native language schooling/literacy follow the same acquisition path? This section will provide an answer to that question.

6.2.1 The route of acquisition

The implication analysis in the previous chapter indicates that the learners follow the route of acquisition under Organic Grammar. For example, the utterances produced by the majority of learners in the VP stage consist solely of lexical items that make up the bare VP. Their data show a noticeable absence of IP-related elements, whether they are free or bound morphemes. As mentioned in the previous chapter, IP-related elements are not predicted in the VP stage. However, three learners used the copula 'is', which was also used as a default form. Regardless of the precise stage of these three learners, the first verbal morpheme to be acquired is the copula 'is'. This early use of the copula 'is', compared to the other verbal morphemes, is supported by studies on the acquisition of English auxiliaries as well as studies on children with L2.

The learners who were placed in the NegP stage not only produced the copula 'is', but also produced the copula 'are'. For example, most of these learners at this stage consistently used the copula 'is', and some of them used the copula 'are'. This indicates that some of these learners appear to have acquired the agreement paradigm for the copula 'be' because most of these learners who produced the copula 'are' also produced the copula

'is'. The auxiliary 'be' is also attested by these learners but is still low compared to the copula 'be'. This indicates that the accuracy in the copula 'be' is ahead of the accuracy in the auxiliary 'be'. This finding aligns with Hawkins's (2001) theoretical perspective that learners acquire the copula 'be' before the auxiliary 'be' because the copula 'be' links the noun phrase with the subject, in contrast to the auxiliary 'be', which links the subject with verb+ing, making it more difficult to produce.

The use of L2 word order (the next sub-stage of the VP stage), the different forms of 'be' (such as 'is' and 'are'), although not always consistent, and the do-forms in negation all indicate that they have acquired the NegP. However, because they do not show any evidence of TP (such as using regular or irregular past tense) or fully fledged agreement (like the third-person singular *-s* on the main verbs), they were placed in the NegP stage. This finding refers to the first implicational ordering that coincides with Organic Grammar, where NegP is immediately acquired after the bare VP.

The learners in the TP stage have presented evidence that demonstrates their acquisition of TP by producing the past tense. Their data also suggest that they have acquired VP and NegP. This is consistent with the sequence of acquisition proposed by Organic Grammar, where TP is acquired after NegP, which is assumed to be acquired after the bare VP. In contrast to learners in the previous stage, these advanced learners demonstrate greater consistency in using both the copula 'be' and the auxiliary 'be' but not the third-person singular *-s*. This indicates that the agreement feature for free morphemes has begun to be established. This finding is in line with the finding of Zobl and Liceras's (1994) presented in Chapter 2 that free morphemes are acquired before bound morphemes. However, it should be noted that there are still some learners in this stage who may occasionally omit the auxiliary 'be'. This is to be expected, as their acquisition of full-fledged agreement is still in progress.

The learners in the AgrP stage appear to have acquired the full agreement paradigm. For example, they have shown evidence of acquiring the agreement feature for free morphemes (the copula 'be' and the auxiliary 'be') as well as the agreement feature on the main verb. Unlike the learners in the previous stage, these learners consistently produce the copula 'be' and the auxiliary 'be' correctly. They have also demonstrated evidence of acquiring VP, NegP, and TP. Again, this finding corroborates the order of stages of Organic Grammar, which posits that AgrP is acquired later than VP, NegP, and TP. The

production frequency for bound morphemes is still lower than that for free morphemes, indicating that bound morphemes are more difficult to acquire than free morphemes. This observation aligns with the findings of some researchers who have explored this pattern and offered explanations. For example, White (1992: 280) suggested that 'the more economical derivation (Chomsky, 1989) raising auxiliaries out of the VP to up agreement in AGR/T was acquired before the less economical affix-lowering of agreement which is required in the case of main verbs in English'. Zobl and Liceras (1994: 173) made a similar suggestion, saying that for L2 learners 'thematic head-movement (AUX BE) is unmarked vis-à-vis affix-movement'. That is, free verbal morphemes are easier to use than bound morphemes. As we recall from Chapter 2, Vainikka and Young-Scholten (1998) made a related proposal. They argue that the acquisition of free inflectional morphemes rather than bound morphemes triggers the acquisition of morphosyntax. This is attributed to the difficulty of phonology. Free morphemes consist of at least a phonological foot, whereas bound morphemes consist of units smaller than a foot. Again, this indicates that bound morphemes are more difficult to acquire than free morphemes.

Analysis of comprehension data also reflects the difficulty of bound morphemes in relation to free morphemes. As seen in the results chapter, each free morpheme is consistently higher in frequency than the bound morphemes, and the statistical analysis also showed that each free morpheme is significantly different from each bound morpheme. This indicates that learners may first produce free morphemes before moving on to use bound morphemes. The analysis also shows that the learners were more accurate in the copula 'be' in relation to their parallel free morpheme (the auxiliary 'be'). Although it is possible that learners start using the copula 'be' before using other free morphemes, the analysis does not show significant differences in accuracy between the copula 'be' and other free morphemes. Similarly, learners in the first three stages (VP, NegP, and TP) were more accurate in the regular simple past tense compared to the third-person singular *-s*. Again, one would assume that learners may use the regular simple past tense *-ed* before the third-person singular *-s*, but the statistical analysis does not show significant differences. Therefore, the comprehension data in this study reflect the production data only in terms of the difficulty of bound morphemes compared to free morphemes.

The investigation of the relationship between learners' scores in comprehension and their OG stages reflects the observed developmental path of acquisition seen in the pro-

duction data. For example, in production data, the acquisition of the agreement features of the copula 'be' occurs before that of the auxiliary 'be'. The analysis does not show that the acquisition of the copula 'be' correlates with OG stages because learners are consistent in their comprehension of the copula 'be' regardless of their OG stages. However, the acquisition of the auxiliary 'be' correlates with OG stages. This indicates that learners' comprehension of the auxiliary 'be' increases as their OG stages increase. This relationship corroborates the results of the production data, where learners in higher stages of OG (TP or AgrP) tend to produce the auxiliary 'be' more frequently than those in lower stages of OG (VP or NegP). Furthermore, an increase in learners' comprehension scores of bound morphemes is accompanied by progression in their OG stages. As explained in the theoretical background presented in Chapter 2, bound morphemes like the simple past *-ed* and the third-person singular *-s* function as heads for TP and AgrP, respectively, and are acquired later in language development compared to the bare VP and NegP stages. This suggests that learners who demonstrate higher OG stages (TP or AgrP) in production data also exhibit a higher level of comprehension for these bound morphemes. In contrast, learners in the lower OG stages (VP or NegP) are more likely to show lower levels of comprehension for these morphemes.

The analysis of comprehension data also reflects the implicational ordering of morphemes observed in their production data. For example, learners who demonstrate a good level of comprehension of bound morphemes also exhibit a higher level of comprehension of free morphemes, which are typically acquired before bound morphemes.

As discussed in Chapter 4, implicational analysis was introduced in part to help minimise the potential limitations associated with cross-sectional studies. As we saw, this analysis could help us detect an OG stage by knowing the preceding and following stages based on the omission or the use of particular verbal morphemes. However, the omission of particular verbal morphemes is not guaranteed to mean that these learners have not acquired the relevant syntax, a factor in determining a stage under Organic Grammar. As discussed earlier, L2 learners may show fluctuations while developing their interlanguage. Furthermore, implicational analysis does not guarantee that L2 learners go through specific stages. For example, the implicational analysis for AgrP stage learners does not solely provide evidence that these learners started their acquisition with the bare VP, followed by NegP, then TP, and finally AgrP.

However, if this implicational analysis actually helped us determine the OG stages, considering the existing proposal about the difficulty of bound morphemes versus free morphemes for L2 learners as well as the analysis of comprehension data, we can say that literacy may affect the rate, not the route. Now we will move on to the following subsection, in which I will discuss the role of literacy in the development of morphosyntax and provide new possible explanations for how literacy could affect the development of morphosyntax based on the distinction between 'learned linguistic knowledge' (Schwartz 1993) and linguistic competence, as well as the distinction between language learning and language acquisition (Krashen 1985). This subsection will also present additional evidence for the role of literacy through the idea of morphosyntactic triggers (Vainikka and Young-Scholten 1998).

6.2.2 The role of literacy

As shown in the previous chapter, there is a correlation between OG stages and reading levels, as well as between OG stages and other non-linguistic variables such as native language schooling, ESOL instruction, residence in the UK, age, and gender. Statistical analysis indicates that literacy and the amount of native language schooling are strongly correlated with OG stages, while the other non-linguistic factors show no correlation.

Ordinal regression analysis presented in the results chapter reveals that only Reading Levels 1 and 2 negatively affect OG stages, reducing the likelihood that learners will progress to higher stages. This finding suggests that these two levels impede the transition between OG stages. Therefore, we expect that learners with lower literacy levels will exhibit lower levels of morphosyntax (lower OG stages). In contrast, Reading Levels 3 and 4 have no negative impact, suggesting that these levels might either facilitate the transition between OG stages or at least not obstruct it. Consequently, we anticipate that learners with higher literacy levels will display higher levels of morphosyntax (higher OG stages). This observation could explain why more than 60% of learners are categorised within the first two stages (bare VP and NegP), despite the amount of ESOL instruction and the duration of their stay in the UK.

However, as discussed in Chapter 2, instructed classroom learners exhibited faster language development than naturalistic learners. Researchers attribute this variation to the metalinguistic information that learners receive in the classroom. In traditional

classrooms, teachers provide learners with metalinguistic information about grammatical properties, followed by a set of exercises. For example, learners are explicitly taught that the morpheme *-s* is attached to verbs to mark agreement with singular subjects in the context of habitual actions. A closer examination of the production data from AgrP-stage learners shows that ESOL instruction, together with literacy levels, plays a significant role in the development of morphosyntax. For example, AgrP-stage learners, except for Eman, who had taken online English courses before arriving in the UK, receive instruction and have reading levels between 3 and 4. This indicates that literacy is essential for learners to fully benefit from classroom instruction.

If we consider Schwartz's (1993) 'learned linguistic knowledge' mentioned in Chapter 2, we can conclude that the correlation indicates that literacy and the amount of native language schooling are more likely to facilitate the learning (not necessarily linguistic competence or acquisition) of morphosyntax. Recall that Schwartz distinguishes between learned linguistic knowledge and linguistic competence. Learned linguistic knowledge is the result of explicit instruction and negative data, whereas linguistic competence is a result of primarily linguistic data. Explicit instruction and negative data cannot feed into linguistic competence. In other words, 'learned linguistic knowledge' and linguistic competence are two distinct processes. Of note, Schwartz's (ibid.) distinction is in line with Krashen's (1985) 'learning vs acquisition' in which acquisition is the result of natural interaction with the language through meaningful communication, which stimulates developmental processes similar to those of L1 acquisition, while learning is the result of classroom experience in which the learner is required to focus on form and study the rules of the target language in a systematic way. As with Schwartz, Krashen argue that learning and acquisition cannot be integrated.

As we saw in the previous chapter, the analysis of production data shows that free morphemes are acquired before bound morphemes, and the copula 'be' is the first verbal morpheme to be acquired among all learners. Additionally, the analysis of comprehension data shows that free morphemes are acquired before bound morphemes, and the copula 'be' is the first free morpheme to be acquired. The analysis in this study shows that literacy positively correlates with the acquisition of the copula 'be', which is strongly correlated with other verbal morphemes in production and, to some extent, in comprehension.

Moreover, the acquisition of the copula 'be' shows a negative correlation with the frequency of the particle 'no' before the main verb (e.g., 'the boy no drink milk') and a positive correlation with the frequency of *do*-forms (e.g., 'the boy don't/doesn't drink milk'). This is similar to what Hawkins observed in terms of the correlation between the acquisition of the copula 'be' and the development of English negation. As stated in Chapter 2, the proportion of copula 'be' coincides with the acquisition of 'not' and 'don't' as heads of NegP. The only difference between my analysis and Hawkins's (2001) analysis is that the copula 'be' does not correlate with the particle 'not'. This might be because my learners use the particle 'don't' a lot, as my analysis shows.

Again, considering the idea of triggers for the development of morphosyntax, we can conclude that literacy indirectly affects the acquisition of morphosyntax (Vainikka and Young-Scholten 1998). Recall that Vainikka and Young-Scholten propose that free morphemes trigger the acquisition of morphosyntax, and these morphemes should be salient in the input. This aligns with Hawkins's (2001) claim that the acquisition of the copula 'be' triggers the acquisition of IP, as detailed in Chapter 2. However, this positional indirect effect of literacy on the acquisition of morphosyntax needs to be addressed not only through the morphosyntax but also with linguistic segmentations, where meta-linguistics again plays a role, helping learners identify the triggers in the input through literacy.

Researchers should also take into account other non-linguistic factors such as trauma, ongoing stress after resettlement, and motivation. As discussed in the literature review, these factors are important for language acquisition in general and morphosyntax acquisition in particular. However, research in this area remains limited, especially regarding the impact of trauma on language development. Therefore, confirming the relationship between literacy and the acquisition of morphosyntax requires substantial effort.

6.3 OG stages and overgenralisation

As discussed in the previous chapter, this study aims to offer a fresh perspective on the interplay between OG stages and the selection of overgeneralised forms. This perspective will be explored by addressing the following research question, as previously mentioned in this thesis.

1. Do the stages of Organic Grammar influence the choice of overgeneralised forms? If so, how?

Therefore, this section will provide answers to these two research questions.

6.3.1 Do the stages of Organic Grammar influence the choice of overgeneralised forms? If so, how?

The type of overgeneralised forms appears to reflect the learners' stage of development (i.e., OG stage). The learners in this study used sub-patterns of overgeneralised forms. As we have seen in the previous chapter, the learners in the VP stage used single function words ('for') or multi-word strings as overgeneralised forms. These learners who used these forms have not yet fully developed their acquisition of NegP and copula 'be'; therefore, they were assigned to the bare VP stage. They also used the expressions ('I don't know', 'I don't like', or 'I don't') in the task aimed to elicit negation. For copula 'be', they often used 'is' as a default form (i.e., they overgeneralise 'is' to 'are'). However, learners with higher OG stages used the copula 'be' and personal pronouns as overgeneralised forms. I believe that this choice of overgeneralised forms is ascribed to the nature of their stage. To wit, the bare VP stage lacks any IP-related elements (e.g., the copula 'be' or auxiliary 'be' or the bound functional morphemes such as the past tense marker *-ed* or the third-person singular *-s*); therefore, learners at this stage are more likely to use other forms that they were previously exposed to in the input. These forms change according to the elements available in the subsequent stage. As noted in the previous chapter, the learners who were placed at the stage of NegP or TP used *be*-forms or the subject personal pronouns as overgeneralised forms, which are available at these stages.

The analysis of comprehension data in this thesis supports this. For example, the scatter plots show a clear pattern: as the use of *be*-forms or subject personal pronouns as overgeneralised forms increases, learners' scores in the comprehension of the copula 'be' or the auxiliary 'be' tend to increase. This indicates that when learners comprehend or produce these IP-related elements, they are more likely to use *be*-forms or personal pronouns as overgeneralised forms and do not use single function words (e.g., 'for') or multi-word strings. However, this relationship has not been confirmed by Spearman analysis, as the p-value exceeds the significance level of 0.05, indicating the need for further research to validate these findings.

The choice of overgeneralised forms has also been observed by Mocciaro (2019b), who reported that overgeneralisation of *fare* ('to do') is connected to the transition from basic

to post-basic varieties, while overgeneralisation of *essere* ('to be') may appear only at a slightly more advanced stage when learners are in the process of acquiring the copula 'be'. This is also expected under the prediction of placeholders in L2 English; the copula 'be' is a candidate as a placeholder for the head of TP or AgrP. My data and this prediction contradict Julien *et al.*'s (2016) findings in terms of the choice of dummy *zijn* ('be') and *gaan* ('go'), as they claimed that learners with lower levels of language acquisition used both dummy auxiliaries while more advanced learners continued to use *gaan*. Thus, they proposed two phases with respect to dummy use: an initial phase in which both dummy *zijn* and *gaan* are used, and a following phase in which dummy *zijn* occurs less often but dummy *gaan* continues to be used. I believe this assumption should be reconsidered as Julien *et al.* (ibid.) only investigated whether *zijn* and *gaan* represent syntactic functions. Their assumption is based on language proficiency rather than on a theory of second language development (e.g., the Basic Variety or Organic Grammar) in which we can determine a specific stage or a point of language development when and where overgeneralised forms begin to appear and disappear.

The third question I seek to answer in this thesis is: What is the nature of overgeneralisation by low-educated L2 learners? The answer to this question yielded the following section, in which I will discuss the results in relation to systematicity and creativity, which were previously hinted at by Mocciaro (2019b) and Vainikka and Young-Scholten (2019). However, this section goes a bit further by providing additional evidence. This includes discussing the syntactic slot for these forms and noting that no lexical material is ever inserted between these overgeneralised forms and the verbs. Furthermore, this section adds to our understanding of systematicity and creativity by gathering more evidence through a comparison of my data with overgeneralisation of the past tense *-ed* in L1 acquisition.

6.3.2 Systematicity and creativity (UG operative in SLA)

The use of overgeneralised forms in this study could point to systematicity and creativity in language acquisition. For example, the syntactic slot for these forms often lies between the subject and the verb, regardless of the choice of these forms (single function words or multi-word strings), and no lexical material is ever inserted between these overgeneralised forms and the verbs.

The onset of using overgeneralized forms reflects systematicity, as these forms emerge after the bare VP. According to the results of this study, seven of the VP-stage learners who used these overgeneralized forms appear to have already established their English VP paradigm. This usage indicates that these forms occur after the bare VP has been established, as they never appeared among learners who transferred their L1 bare VP. As discussed in the previous chapter, the bare VP-stage learners who used the preposition 'for' or multi-word strings (e.g., 'you have' or 'in the') showed no evidence of IP-related elements. This supports the conclusion that the use of overgeneralized forms begins immediately following the bare VP, rather than in the NegP stage, thus confirming my hypothesis.

As mentioned previously, one of the main objectives of the OG placement is to accurately determine the stage at which learners begin to use overgeneralized forms, which is established through implicational analysis. In Section 6.2.1, I acknowledged that relying solely on the omission of specific morphemes from a single data collection (i.e., cross-sectional data) may not be sufficient to definitively identify an OG stage, as learners' interlanguage may fluctuate. Additionally, the analysis of cross-sectional data alone does not guarantee that overgeneralizations begin after the VP stage, highlighting the need for longitudinal data to confirm this conclusion.

This belief, in general, is consistent with Vainikka *et al.*'s (2017) observation that the use of overgeneralised forms occurs after the bare VP stage. A similar claim was also made by Mocciaro (2019b) in that non-target constructions emerge as soon as L2 Italian learners enter the post-basic variety, a stage after the basic variety that is similar to the bare VP stage of Organic Grammar.

Echoing Julien *et al.* (2016), Mocciaro (2019a), Vainikka and Young-Scholten (2019), and Vainikka *et al.* (2017), these overgeneralised forms are placeholders. In addition to the specific syntactic position and choice of overgeneralised forms, the lack of inflection also supports the use of these forms as placeholders. All verbs that follow these forms always appear in their base forms. Plus, inflected verbs never occurred with these forms (e.g., 'the boy you have played...' or 'the boy he played...').

Furthermore, similar to what has been observed by Vainikka and Young-Scholten (2019) and Vainikka *et al.* (2017), the participants in this study did not use content words as placeholders. Instead, they used function ones, which indicates that mental

linguistic mechanisms operate subconsciously in L2 acquisition. I attribute the choice of function words to their grammatical functions, unlike content words, indicating that the learners in this study use them for grammatical purposes (that is, as placeholders for morphosyntactic elements). This evidence alone refers to the operation of UG as adopted under Organic Grammar, which suggests that all learners, including L1 children and younger and older L2 learners, use the potential syntactic heads provided by Universal Grammar. The language learner subconsciously searches for syntactic heads in the input. Once the learner identifies the head, Universal Grammar provides the tools to subconsciously build the relevant syntactic projection for that head. Failure to identify the relevant heads results in the use of overgeneralised forms. As just mentioned, the analysis of production data in this study shows that overgeneralised forms are placed before bare verbs. This indicates that overgeneralisation occurs when learners fail to identify the relevant head. That is, overgeneralised forms are used as placeholders for the syntactic head (e.g., the third-person singular *-s*). Overgeneralised forms as placeholders have also been noted by Julien *et al.* (2016: 54) in that the use of dummy auxiliary (i.e., placeholders) is 'a structural step in the acquisition of finiteness'. Julien *et al.* report that learners used overgeneralised forms as place markers before producing lexical verbs with agreement. This is also attested by Bernini (1989) and recently by Mocciaro (2019a,b). Bernini claims that overgeneralised forms create a provisional analytical stage before the acquisition of finiteness. Mocciaro argues that the learners who produce overgeneralised forms have reached at least an initial level of post-basic variety, which is indicated by using past participle forms and the copula 'be'.

Overgeneralised forms among L2 adult immigrant learners are comparable to those among L1 children in terms of creativity and systematicity in child L1 acquisition. Recalling from Chapter 3, children acquired morphosyntactic elements in successive stages. Children at the first stage supply the correct bound morphemes, followed by a stage where they start to extend these morphemes to irregular verbs, resulting in a non-target form (e.g., 'goed'). At the final stage, overgeneralisation begins to disappear. This process of acquisition is taken as strong evidence for the operation of UG in L1 acquisition and, thereby, creativity. For example, the construction ('goed') is not a result of imitation, as this construction is not found in the input. Native speakers do not produce this construction. This can be applied to overgeneralisation in L2 acquisition; learners' constructions

(e.g., 'Subject + you have + Verb + Object') are not found in the input.

The striking difference between L1 and L2 acquisition in terms of overgeneralisation is that L2 learners overgeneralise free morphemes, unlike L1 child learners, who overgeneralise bound morphemes. The learners in this study used only free morphemes as overgeneralised forms. This has important implications for ideas of Vainikka and Young-Scholten (1998) regarding morphosyntactic triggers under Organic Grammar. Recall from Chapter 3 and this chapter, to account for the difference between the full success of language acquisition between child L1 and younger and adult L2 learners, Vainikka and Young-Scholten propose that bound morphemes trigger the acquisition of morphosyntax in L1, while free morphemes do so in L2. Adult L2 learners struggle more than child L2 learners to identify projection heads due to difficulties in acquiring L2 phonology, which causes difficulty distinguishing suffixes such as *-ed* from the verb stem (Vainikka and Young-Scholten 1998; Zobl and Liceras 1994). The difficulty in distinguishing heads may be exacerbated for less-literate learners because they have less visual reinforcement of linguistic forms. Now we will discuss the likelihood that these forms serve as placeholders, based on insights from the analysis of comprehension data. As noted in this thesis, comprehension data were used as a contribution to production data and to Vainikka *et al.*'s (2017) study.

6.3.3 Evidence for the idea of placeholder: insight from comprehension data

In addition to the evidence from production data suggesting that these forms may function as placeholders, the statistical analysis in this thesis offers relative support for the notion that overgeneralised forms (specifically those used by lower OG stages) might serve as placeholders for bound morphemes (the syntactic heads for TP or AgrP). As we observed in the results chapter, there is a strong negative correlation between the overgeneralisation of single function words or multi-word strings and learners' scores in the comprehension of both the regular past simple *-ed* and the third-person singular *-s*. Therefore, we expect that learners with lower comprehension levels of these morphemes might use overgeneralised forms as placeholders. Similarly, the analysis provides some support for the idea of placeholders, indicating a strong negative correlation between the overgeneralisation of *be*-forms and learners' scores in the comprehension of the third-person singular *-s*, but not with that of the regular past simple *-ed*. Based on this finding,

we expect that when learners have lower comprehension levels of the third-person singular *-s*, they specifically use *be*-forms as placeholders for this morpheme, not for the regular past simple *-ed*. This distinct pattern suggests that although the use of single function words or multi-word strings as overgeneralised forms negatively correlates with both *-ed* and *-s*, they particularly serve as indicators of placeholders for the third-person singular *-s*. That is, these overgeneralised forms are placeholders marking agreement, not tense. However, research on overgeneralisation by low-educated learners or on the overgeneralisation of 'be' by child L2 learners is not conclusive about what these forms are placeholders for (agreement or tense). To my knowledge, no studies specifically address this distinction, except for the observations provided in Chapter 3.

However, the analysis of the use of personal pronouns as overgeneralised forms shows no support for the idea that these forms may act as placeholders for the syntactic heads. As we saw in the previous chapter, neither the scatter plot nor the correlation test shows a relationship. This contradicts the hypothesis in this thesis that overgeneralised forms negatively correlate with learners' scores on comprehension of the bound morphemes, as well as the tentative analysis from García Mayo *et al.* (2005) and Vainikka and Young-Scholten (2019) that personal pronouns, when overgeneralised, are used as placeholders for agreement. I will come back to this later when I discuss the use of these two forms and L1 transfer in Section 6.4.

Now, we will move on to the next section, which presents a comparison between the use of formulaic chunks through the Exemplar within the Usage-based framework and the use of chunks as placeholders under Organic Grammar. This comparison provides more insight into how the overgeneralisation of chunks comes into play in the development of morphosyntax from an Organic Grammar perspective.

6.3.4 Exemplar as chunks vs. chunks as placeholders

Like what has been observed by Vainikka *et al.* (2017), my data demonstrates that learners at the bare VP stage used multi-word strings, which were placed between the subject and bare verbs in the same way as single function words. Therefore, I believe these multi-word strings are also placeholders. Additional but strong evidence is the use of the strings ('I don't know', 'I don't like', or 'I don't') in the negation tasks, which apparently are used as sentential negators (e.g., 'the boy I don't know drink milk'). Recall from Chapter 3,

some researchers within the Usage-based framework, such as N. C. Ellis (2002), Eskildsen (2012), and Myles (2004), argue that formulaic language serves as a database for the acquisition of morphosyntax. These researchers believe that much of our language use is formulaic, in the sense that we recycle phrasal constructions that we have memorised from previous usage (see, e.g., Wulff 2008). In this respect, our language processing is not confined to these constructions. Certain constructions, such as the slot-and-frame greeting pattern 'Good + (time-of-day)', generate instances such as 'Good morning' and 'Good afternoon'. Others still are abstract, broad-ranging, and generative, such as the schemata that represent more complex morphological (e.g., '[NounStem-PL]'), syntactic (e.g., '[Adj Noun]'), and rhetorical (e.g., the iterative listing structure, '[the ( ), the ( ), the ( ), . . ., together they . . .]') patterns. According to exemplar theory, constructions such as 'Good + (time of day)', '[Adj Noun]', or '[NounStem-PL]' progressively arise over time when the learner's language system recognises the regularities shared by exemplars and abstracts them. One example, as reviewed in Chapter 3, is the study by Eskildsen (2012), who claims that the multi-word expression 'I don't know' serves as an exemplar, followed by a gradual increase that expanded to other verbs and pronouns ('Subject + don't + Verb'). This construction is different from what I have found in my data in that my participants placed the strings 'I don't know' in the bare VP, resulting in the following construction: ('Subject + I don't know + Verb + Object'). This usage was also observed in Vainikka *et al.* (2017), in which their learners use the expression 'I don't' between the subject and the verb ('Subject + I don't + Verb + Object'). Therefore, I argue that overgeneralised forms can also involve multi-word strings as used by low-literate learners in the current study.

The multi-word strings ('you have' or 'in the') in my study also differ from the sort of multi-word chunks to which Myles (2004) refers. As noted in Chapter 3, Myles traced verbs that originated in chunks and observed that these verbs behave differently from other verbs, as they always appear in a finite form (i.e., an inflected form). In contrast to non-chunk verbs, which start life uninflected, all instances of these verbs in the three chunks or outside the chunks appear in a finite form. On the other hand, the multi-word strings in my data are added to the bare VP as extra forms and seem to be used to mark syntactic functions. Another important difference is that the chunks I refer to do not always contain verbs. For example, the string 'in the', which is frequently used before

bare verbs. A similar instance was reported in Vainikka *et al.* (2017).

Therefore, the use of multi-word strings in my data is not to feed into grammar. Instead, it is used as placeholders, and the use of multi-word strings by VP-stage learners supports Myles's (2004: 139) proposal that 'early grammars consist of lexical projections and formulaic sequences'. This, in turn, supports Organic Grammar in terms of the lack of functional elements at the bare VP stage.

6.3.5 Possible explanations for using multi-word strings as placeholders

This section provides new explanations for why some VP-stage learners use multi-word strings as placeholders for morphosyntactic elements. These explanations are based on the concepts of salience and noticeability within the Usage-based framework.

6.3.5.1 Salience and noticeability

As previously discussed, VP-stage learners in this study use multi-word strings as over-generalised forms, unlike higher-stage learners, who use only single function words. In Section 6.3, I have suggested that learners select forms based on the availability of functional elements in their OG stages. If that stage lacks functional elements (e.g., the bare VP stage), they use other elements available in the input. In this section, I argue that these forms are not only available in the input but also salient and noticeable. Under the Usage-based approach, formulaic language is more salient orthographically, auditorily, and perceptually than morphosyntactic elements. It is more likely that salient forms would be perceived, attended to, and processed successfully (Wulff 2019). Therefore, perceptual salience may explain 'success and failure' in learning forms through the approach of Focus on Form (M. Long 2015: 60). According to Loewen and Sato (2018), the saliency of linguistic elements is indeed a mediating factor in L2 development during communicative interactions. Furthermore, the communicative value of formulaic language might be another possible reason for using multi-word strings as placeholders. Formulaic language, as opposed to morphosyntactic elements (e.g., the third-person singular *-s*) that are redundant in message comprehension, is vital for communicating meaning (Wulff 2019). The communicative value of linguistic elements has been tied to their developmental patterns in L2 learning (Loewen and Sato 2018). According to VanPatten's (2007) 'lexical preference principle', L2 learners analyse lexical items for meaning prior

to morphosyntactic forms.

Another plausible explanation for using multi-word strings as placeholders might be because formulaic language is more noticeable. The noticeability of forms is influenced by the type of linguistic purpose (Loewen and Sato 2018). Stimulated recall data revealed that corrective feedback for lexical errors was more consistently noticed than that for morphosyntactic errors (Mackey *et al.* 2000). Similarly, using a dual-mode system that differentiates between a memory-driven exemplar-based system and a rule-based analytic system, Y. Yang and Lyster (2010) observed that rote-learned items were more noticeable than rule-based ones. Their findings demonstrated that corrective feedback focused on rote-learned items is more effective, regardless of the type of corrective feedback. Previous research has demonstrated that vocabulary items are more prominent than morphosyntax (Loewen and Sato 2018). This research contributes to our understanding of how formulaic language, as larger units of lexical items, influences learners' degree of noticeability and is used as placeholders for morphosyntactic elements. The salience and noticeability of forms as overgeneralised forms have also been suggested to explain why learners with a low level of morphosyntax use them rather than functional elements such as the copula 'be'. As noted in Chapter 3, Mocciaro (2019a) suggest that the reason why learners with a low level of morphosyntax use light verbs ('to go') rather than functional elements such as the copula 'be' is because they are more salient in the input. Moving on to the last section, which raises some caution about the current assumptions regarding the use of subject pronouns as placeholders, because the data in this study reveals different patterns.

6.4 Overgeneralisation and L1 influence

Is there evidence of L1 influence with respect to overgeneralisation? The participants in this study used subject personal pronouns ('he', 'she', 'they') before lexical verbs. Some of them used the pronoun together with *be*-forms again before the bare verb (e.g., 'the boy he is drink milk' or 'the girl she is drink milk'). If we look closely at the analysis of Arabic, one would assume that *be*-forms and the personal pronouns may have different functions. *Be*-forms mark the third-person agreement in English, and the personal pronouns mark gender as a result of transferring the characteristics of the agreement morphemes from Arabic, whose verbs are inflected for number, person, and gender.

The use of the pronouns as overgeneralised forms has been reported and interpreted from two different perspectives: the transfer view (García Mayo *et al.* (2005)) and the non-transfer view (Vainikka and Young-Scholten (2019)). Of note, the use of pronouns in García Mayo *et al.* (2005) as well as Vainikka and Young-Scholten (2019) never occurred in combination with *be*-forms, unlike my data. In the transfer view, García Mayo *et al.* interpreted both ('he' and 'is') as placeholders for agreement as a consequence of transferring the characteristics of the agreement morphemes from their first language. In the non-transfer view, Vainikka and Young-Scholten tentatively interpreted this pronoun as a placeholder for the third-person singular *-s*. However, the results do not support this assumption, and show no relationship.

One might assume that learners using such pronouns are more advanced than those who use single function words or multiple word strings or *be*-forms. They may be in the process of comprehending bound morphemes, but they are not yet producing them consistently. However, my data show that these pronouns sometimes use *be*-forms or a combination of personal pronouns and *be*-forms. However, this does not accurately confirm whether these pronouns are placeholders for the functional bound morphemes. Using them as placeholders can be considered a reasonable perspective, since these learners change these pronouns according to the subject (plural or singular).

Overgeneralisation of the personal pronouns has important implications for Hawkins's (2001) Modulated Structure Building, which is similar to Organic Grammar but allows for the transfer of functional projections at the relevant stage of L2 development. Hawkins argues for L1 influence in terms of subject-verb agreement, although he provides little evidence for such a transfer from L1. The influence of L1 on subject-verb agreement is the only argument in his book that deals with L1 transfer at the sentential level (i.e., not DP; see Vainikka and Young-Scholten's [2003b] review of Hawkins' book). The evidence provided by Hawkins is based on Stauble's (1984) comparison, which shows that Spanish L1 speakers acquire the copula and subject-verb agreement in L2 English before Japanese L1 speakers, indicating L1 influence. However, Vainikka and Young-Scholten (2003b) point out two main issues in this argument: (1) the sample size is relatively small, with only six speakers in each group and two hours of data collection for each speaker; (2) the performance levels were based on negation data from each speaker. It would be intriguing to see whether learners whose L1 has no explicit morphological marker (e.g., Chinese) use

these pronouns as placeholders or whether such use is only found in learners whose L1 has explicit morphological markers. If such a use of pronouns is found among learners whose L1 has no explicit morphological markers, this would only be the interaction between UG and the input under the structure building of Organic Grammar, and L1 would have no role to play.

6.5 Summary

In this chapter, I have discussed the results of morphosyntax in terms of the route and rate of morphosyntax acquisition. Based on implicational analysis, I have shown that the route of acquisition appears to be consistent among the learners recruited for this study. In line with the existing literature on low-educated learners, literacy seems to delay the process of acquiring morphosyntax. However, confirming this direction of influence requires further research, including the consideration of non-linguistic factors such as trauma, stress, and lack of motivation, as well as the investigation of metalinguistic awareness and literacy on the development of morphosyntax.

This chapter has also examined the results of the use of overgeneralised forms in relation to the stages of Organic Grammar. I propose that OG stages appear to influence the choice of overgeneralised forms, as these forms reflect the learners' OG stage. As discussed, learners use different overgeneralised forms based on their OG stages. These forms result from a failure to identify the relative syntactic head in the input, with evidence showing that these forms precede bare verbs, that no lexical items are placed between these forms and bare verbs, and that their use negatively correlates with lower learners' scores in bound morphemes (the regular past tense *-ed* or the third-person singular *-s*). Therefore, I suggest that these forms are placeholders for the heads of TP or AgrP. I have demonstrated that overgeneralised forms could indicate systematicity and creativity in language acquisition. In this chapter, I have also discussed the possible role of literacy in using multi-word strings as overgeneralised forms, arguing that low-literate or VP-stage learners tend to use more identifiable elements such as multi-word strings as overgeneralised forms, in contrast to those with higher literacy levels or at higher OG stages, who tend to use IP-related elements.

Chapter 7. Conclusion

The aim of this study is twofold: the first is to confirm whether low-literate learners follow the route of acquisition under Organic Grammar as identified by Vainikka *et al.* (2017). The second aim is to investigate overgeneralisation by these learners using the developmental path of acquisition under Organic Grammar. Data come from a larger sample size (60) of learners through production tasks and comprehension tasks (newly designed computerised match questions). The production tasks included picture descriptions designed to test the acquisition of morphosyntax based on the stages of Organic Grammar. For example, in this task, I tested the word order in L2 English, which constitutes the VP stage, and the English negation construction in declaratives, which constitutes the NegP stage. For the acquisition of functional morphology, I looked at the acquisition of the irregular past tense, which indicates the acquisition of TP, and the agreement features of the copula 'be' and the auxiliary 'be', as well as the acquisition of the third-person singular *-s* on the main verbs, which all indicate the acquisition of AgrP.

The comprehension tasks included an online link for the audio-picture match questions to test the acquisition of English negation in declaratives, agreement of the copula 'be' and auxiliary 'be', as well as agreement on the main verbs (the third-person singular *-s*) and the acquisition of tense (regular past tense *-ed*). The aim is to determine whether learners' performance in comprehension reflects their performance in production and to investigate if there is a relationship between low levels of comprehension and the use of overgeneralisation.

A more in-depth analysis was conducted to assign stages. For example, unlike Vainikka *et al.*, this study employed implicational analysis to determine stages by setting specific criteria for morphosyntactic features, which helped us detect the implicational ordering of stages by identifying the preceding and following stages. In addition to the general criteria of OG stages in L2 English, this study used Déprez and Pierce's (1993) predictions in terms of word order in negative constructions in L1 English to determine whether the

learner's negation is in or outside the VP, particularly in the early stages of L2 acquisition. The analysis shows an implicational ordering of OG stages, starting with the VP stage, followed by the NegP stage, and the intermediate stages (TP and AgrP), supporting my hypothesis that the learners will follow the predictable OG stages.

Ordinal regression analysis was conducted to investigate the relationships between OG stages as a dependent categorical variable and literacy, including other non-linguistic factors as independent variables, to examine how these factors affect the correlations observed through Spearman's correlation. The results show that only literacy among these variables has a negative effect on the progression of learners through the stages of Organic Grammar, confirming my hypothesis that lower literacy levels negatively affect progression through OG stages.

As a secondary investigation into the developmental path of OG stages, the study compared learners' comprehension of the verbal morphemes with their OG stages in production. This investigation was preceded by an initial analysis, including the mean and SD, to explore the distribution of the verbal morphemes among those learners. This analysis also included the Friedman test to assess the differences between these morphemes. The results showed that the mean values for free morphemes were higher and more consistent than those for bound morphemes, and the test indicated a significant difference between each free morpheme and each bound morpheme, whereas there was no significant difference among the free morphemes and among the bound morphemes. This aligns with the results of the production data, where learners are more consistent in using free morphemes than bound morphemes.

Then, the study compared learners' performance in verbal morphemes according to their OG stages. This included examining the relationship between learners' scores in comprehension in relation to their OG stages through the mean and SD for variability and using scatterplots to visualise any potential relationship, followed by Spearman's correlation to assess correlations. The results show that while the mean values for learners' scores in free morphemes do not change significantly according to their OG stages, the mean values for learners' scores in bound morphemes increase in tandem with the mean values for comprehension scores. Spearman's correlation detected a significant correlation between bound morphemes and OG stages.

As previously discussed, these bound morphemes are functional morphology and func-

tion as heads for TP and AgrP. Therefore, these results reflect learners' performance, where learners at higher stages use bound morphemes more than those in the lower stages. These results provide further support for my hypothesis that learners will follow predictable stages.

Based on the idea that free morphemes, rather than bound morphemes, serve as triggers for the development of morphosyntax, I examined the relationship between the copula 'be' as a potential trigger using both production and comprehension data, alongside other verbal morphemes investigated in this study (the auxiliary 'be', the regular past tense *-ed*, and the third-person singular *-s*), including the acquisition of negation. The study explored this relationship through the analysis of means and SDs and scatterplots to visualize the relationships, followed by Spearman's correlation to test the associations. The same analytical approach was used to investigate the relationship between literacy development and the acquisition of the copula 'be' in both production and comprehension data.

The analysis of the production data shows that the copula 'be' strongly correlates with the acquisition of negation and other verbal morphemes, suggesting that the acquisition of the copula 'be' could function as a trigger for the acquisition of inflectional morphemes and thus the activation of functional projections (e.g., TP or AgrP). While the comprehension data do not fully mirror this broad relationship, they still show significant correlations between the copula 'be' and both the auxiliary 'be' and the regular past tense *-ed*, but no significant correlations were found with negation or the third-person singular *-s*. Despite the confined scope of these associations, this finding supports the idea of triggers if we consider that acquiring the copula 'be' allows L2 learners to break into the inflectional system, as both the auxiliary 'be' and the past tense *-ed* are inflectional morphemes. The analysis of the relationship between literacy development and the acquisition of the copula 'be' shows a strong correlation in both production and comprehension data, indicating that literacy may reinforce the visibility of the copula 'be' in the input, thereby triggering the acquisition of functional projections.

To investigate whether there is a relationship between learners' use of each type of overgeneralisation and each verbal morpheme, I conducted a Spearman's correlation. This was preceded by an investigation through the mean and SD, and then through scatter plots to visualise the relationships. The results show that there is a strong correlation be-

tween learners' overgeneralisations of single function words or multi-word strings and each bound morpheme, confirming my hypothesis that learners who engage in overgeneralisation struggle with the acquisition of bound morphemes in comprehension. In contrast, the results show that there is no correlation between the use of overgeneralisation of this type and learners' comprehension of free morphemes, indicating no relationship. This is expected when considering that free morphemes are presumed to be acquired before the acquisition of bound morphemes, aligning with the results in this thesis.

The results also show that there is a significant relationship between the third-person singular *-s* and overgeneralisation of 'be', but not with the regular past tense *-ed* and overgeneralisation of 'be'. This partially confirms my hypothesis, but indicates that learners in general tend to use these forms (single function words or multi-word strings or overgeneralisation of 'be') as placeholders for the third-person, not for the regular past tense. This is even though the results show a negative correlation between the regular past tense *-ed* and single function words or multi-word strings as overgeneralised forms. This is also a plausible interpretation when considering that the third-person singular *-s* is acquired later than the regular past tense *-ed*, as suggested in the literature and observed in the production tasks in this current study.

The results do not show any relationship between learners' use of overgeneralisation of 'be' and learners' scores of free morphemes in comprehension. As expected in the results chapter, we may find no relationship or a positive but not negative correlation, and this is because (1) free morphemes, as just mentioned, are acquired early, and (2) the learners who used *be*-forms as overgeneralised forms also used target-like *be*-forms.

Finally, the results show that there is no connection between learners' use of personal subject pronouns as overgeneralised forms and bound morphemes, nor with free morphemes. As discussed in the preceding chapter, the non-relationship observed in this thesis contradicts the claim that these pronouns are used as placeholders for subject-verb agreement. However, this is a reasonable interpretation since the learners switch using these pronouns according to the subject (e.g., 'the boy he play' or 'the girl she play'). My concern here is not about agreement; rather, it is about whether their use is an effect of L1 transfer, as explained in the discussion chapter.

With respect to the non-relationship between overgeneralisation of these forms and learners' scores in comprehension, this is expected. As noted in this thesis, learners tend

to overgeneralise 'be' while simultaneously using free morphemes in a target-like way.

7.1 Main research findings

Taking into account the limitations of cross-sectional studies in terms of the developmental stages discussed in this thesis and the results of this study based on implicational ordering, as well as the comprehension data, the main findings indicate the following:

1. The route of acquisition appears to be fixed regardless of the development of literacy or other non-linguistic factors.
2. Lower literacy levels could negatively affect the progression between OG stages. That is, it could hinder the progression between OG stages.
3. OG stages could affect the choice of overgeneralised forms. As stated in the previous chapter, the use of single function words (e.g., 'for') or multi-word strings (e.g., 'you have') seems to be peculiar to the VP-stage learners, while learners with higher stages use different types of placeholders (e.g., the copula 'be' or personal pronouns), thus supporting the recent predictions of placeholders through the stages of Organic Grammar (Vainikka and Young-Scholten 2019).
4. It appears that overgeneralisation does not alter the general developmental path of acquisition (i.e., the stages of Organic Grammar).
5. Similarly to what has been observed by Vainikka *et al.* (2017) and Myles (2004), overgeneralisation can involve morphemes, words, and multi-word sequences.
6. Overgeneralised forms are more likely to act as temporary placeholders that learners temporarily use as they work on identifying the relevant syntactic heads in the input. That is, failure to identify the relevant heads results in using placeholders: The analysis shows a negative relationship between the use of placeholders and comprehension of the relevant morphosyntax, and the use of placeholders increases as comprehension scores decrease. The temporary use of placeholders has also been attested by Julien *et al.* (2016: 54) in that the use of dummy auxiliary (i.e., placeholders) is 'a structural step in the acquisition of finiteness'.

7.2 Limitations

The limitations of the present study are summarised as follows:

1. The impact of Covid-19 has affected data collection in terms of the target learner. The intended aim was to involve illiterate learners versus low-literate learners to investigate whether the route of acquisition by illiterate learners versus low-literate learners differs, as claimed by Vainikka *et al.* (2017). Therefore, the study cannot make the claim that literate and illiterate learners follow the same developmental path of acquisition as suggested under Organic Grammar.
2. As detailed in Chapter 2, the role of input is largely ignored within the generative approach. This causes methodological issues that affect any conclusions about the role of literacy in the development of morphosyntax. This is because low-educated L2 learners are assumed not to interact with native speakers.

7.3 Directions for future research

I believe the findings of this current study raise interesting questions.

1. Will a longitudinal study of naturalistic L2 literate versus illiterate learners show a fixed OG developmental path for L2 English acquisition? As noted in this thesis, the data for this study were cross-sectional and were collected from only low-educated/low-literate learners. It is worth repeating that numerous scholars (e.g., Hawkins 2001) have asserted that the paths of L2 acquisition are the same regardless of the context in which L2 acquisition takes place. As previously discussed in this thesis, implicational analysis was introduced to minimise the potential limitations of cross-sectional analysis with respect to fluctuation in L2 acquisition, but conducting detailed longitudinal investigations is still crucial to confirm the stages and sequences proposed by Organic Grammar theory. Longitudinal investigations will also confirm the appearance and disappearance of overgeneralised forms through the stages of Organic Grammar.
2. Can the use of the placeholder 'he' be found among learners whose L1s have the same structure as English functional projections or among learners whose L1s do not have any explicit morphological markers (e.g., Chinese)? If so, this use of 'he' as a placeholder is not traceable to L1.

Appendix A. Consent form and questionnaire

Consent form (adopted from Newcastle University) I, the undersigned, confirm that (please tick box as appropriate):

1.	I have read and understood the information about the project, as provided in the Information Sheet dated _____.	☐
2.	I have been given the opportunity to ask questions about the project and my participation.	☐
3.	I voluntarily agree to participate in the project.	☐
4.	I understand I can withdraw at any time without giving reasons and that I will not be penalised for withdrawing nor will I be questioned on why I have withdrawn.	☐
5.	The procedures regarding confidentiality have been clearly explained (e.g. use of names, pseudonyms, anonymisation of data, etc.) to me.	☐
6.	If applicable, separate terms of consent for interviews, audio, video or other forms of data collection have been explained and provided to me.	☐
7.	The use of the data in research, publications, sharing and archiving has been explained to me.	☐
8.	I understand that other researchers will have access to this data only if they agree to preserve the confidentiality of the data and if they agree to the terms I have specified in this form.	☐
9.	I, along with the Researcher, agree to sign and date this informed consent form.	☐

Participant:

Name of Participant	Signature	Date

Researcher:

Name of Participant	Signature	Date

Name: Age: Gender:

Questionnaire

1. Where were you born? (Country and region).

   .....
   .....
   .....

2. Where have you lived? (Give countries and regions and length of time at each location).

   .....
   .....
   .....

3. When did you start learning English?

   .....
   .....
   .....

4. What programme level of English are you in (e.g., Pre-entry 1, 2, or 3)?

   .....
   .....
   .....

5. How many weeks/months/years did you receive formal English instruction?

   .....
   .....
   .....

6. How many hours did you receive formal English instruction per week?

   .....
   .....
   .....

7. Have you ever been to an English-speaking country before arriving in the UK (e.g., USA)? YES ☐ NO ☐
If yes, which one(s) and how long did you spend there?

.....
.....
.....

8. Which other countries have you visited?

.....
.....
.....

9. What other languages do you know? To what level? (1 = beginner, 10 = native) (Please provide ANY formal language teaching you have had including school lessons, private study, etc.).

.....
.....
.....

10. Does any member of your family speak English as a native or second language? YES ☐ NO ☐
If yes, give more detail:

.....
.....
.....

11. Do you know anyone whose native language is English (e.g., friend, relative)? YES ☐ NO ☐
If yes, give more detail:

.....
.....
.....

12. Do you watch TV/films or listen to music in English? YES ☐ NO ☐
If yes, give more detail (e.g., which shows/films/music, how often):

.....
.....
.....

13. Did you receive formal schooling in your country? YES ☐ NO ☐
If yes, how many years did you receive and were they continuous or interrupted?

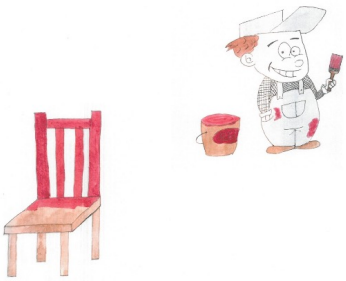
.....
.....
.....

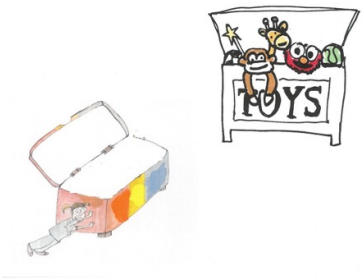
Appendix B. Stimulus

B.1 Testing materials for production tasks

B.1.1 VP task

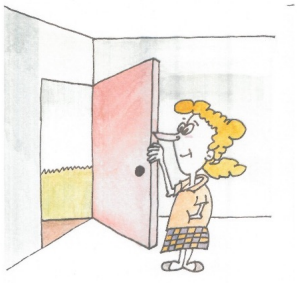
As stated in the methodology chapter, the learner's task is to describe the actions in these picture sheets. The researcher provides only the beginning of each sentence. For example, 'The girl...'

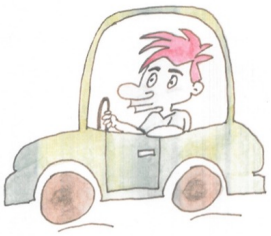


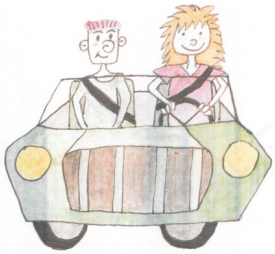

TOYS

B.1.2 NegP task

In this task, the researcher says the beginning of each sentence, and the learner's task is to describe these groups of picture descriptions. For example, 'the boy...' and the learner completes the sentence. Similarly, for the picture without action, the learners are expected to produce a negative sentence. For example, 'The boy doesn't clean the room.'

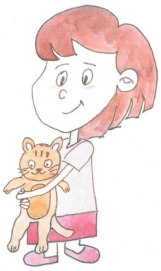


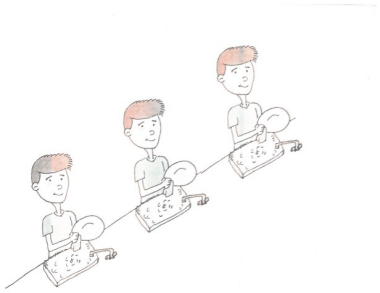


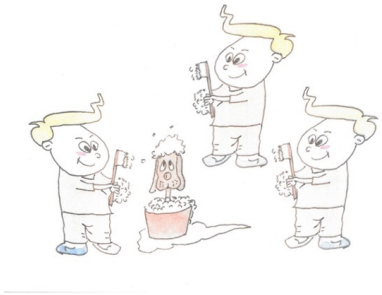


B.1.3 Task for progressive tense

The learner's task is to describe the action in response to the question. For example, 'What is this girl doing?' The learner is expected to produce a sentence with the progressive tense. For example, 'The girl is reading a book.'



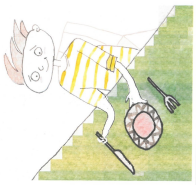


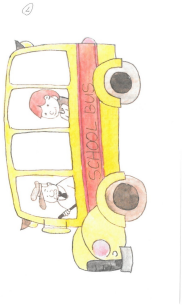


B.1.4 Tasks for third person singular -s on main verbs

The learner is to describe the action based on what they see and the beginning of each sentence. Each two groups have the same character, and the researcher says for the first picture (e.g., 'Every morning, Sam...'), and the learner is expected to produce a sentence

(e.g., ‘Every morning, Sam drinks coffee’). Then the researcher says, ‘Then Sam or He...’ for the second picture, and the learner is expected to say, ‘Sam smokes a cigarette’.





B.1.5 Tasks for past simple tense: Repeated from the methodology chapter

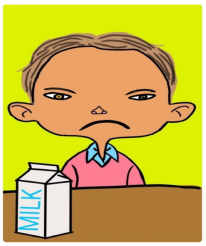
The learner's task is to retell a story after seeing it for the first time. For example, after the learner views a series of subsequent actions in the picture sheets, the researcher says

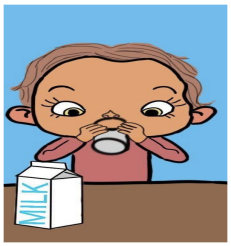
(e.g., 'Last week'... as an indication to use the past tense). Then the learner's task is to describe the incident in the past.

B.2 Testing materials for comprehension tasks

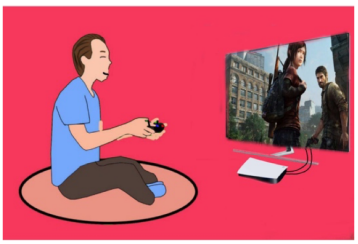
B.2.1 Negation

The learner's task is to select the photo based on the stimulus they hear. For example, the learner listens to a negative sentence (e.g., 'He doesn't drink milk').







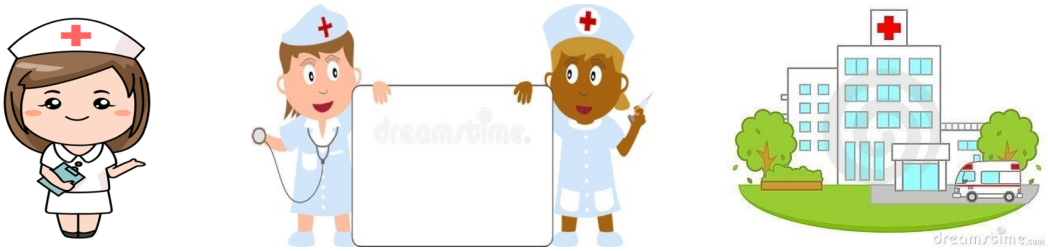
B.2.2 Copula 'be'

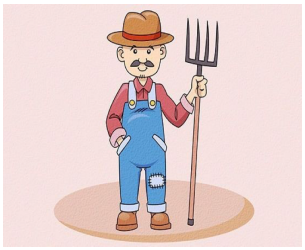
B.2.2.1 Copula 'is'

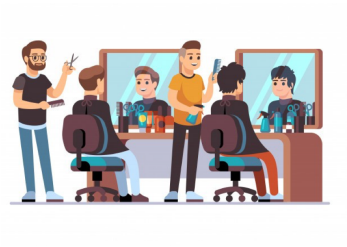
The learner's task is to select the photo based on the stimulus they hear. For example, the learner listens to '... is a policeman'. '...' is a pause.

B.2.2.2 Copula 'are'

The learner's task is to select the photo based on the stimulus they hear. For example, the learner listens to '... are nurses'. '...' is a pause.



B.2.3 Auxiliary 'be'

B.2.3.1 Auxiliary 'is'

The learner's task is to select the photo based on the stimulus they hear. For example, the learner listens to '... is eating breakfast'. '...' is a pause.

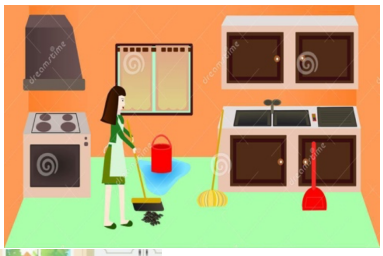







B.2.3.2 Auxiliary 'are'

The learner's task is to select the photo based on the stimulus they hear. For example, the learner listens to '... are cleaning the kitchen'. '...' is a pause.







B.2.4 Third person singular -s

The learner's task is to select the photo based on the stimulus they hear. For example, the learner listens to '... plays PlayStation'. '...' is a pause.

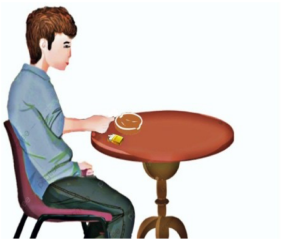


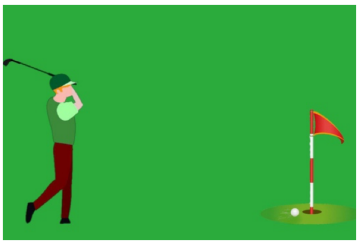




B.2.5 Regular past tense -ed

The learner sees a series of subsequent actions in each group of these photos while simultaneously listening to a stimulus. For example, they hear 'He smoked a cigarette'. Then, the learner's task is to select the target photo based on the stimulus they hear.



References

Adrian, J. A., J. Alegría, and J. Morais. 1995. 'Metaphonological abilities of Spanish illiterate adults'. *International Journal of Psychology* 30: 329–351.

Ahn, S. 2003. 'A note on the topic-comment stage in korean EFL syntactic development'. *Studies in Generative Grammar* 13: 369–382.

Ahn, S. 2006. 'The grammar of verb be in early Korean EFL interlanguages'. *Studies in Generative Grammar* 16: 769–782.

Altenberg, B. 1998. 'On the phraseology of spoken English: The evidence of recurrent word combinations'. In A. P. Cowie (ed.). *Phraseology: Theory, Analysis, and Application*. Oxford: Oxford University Press. 161–186.

Andersen, R. W. 1978. 'An implicational model for second language research'. *Language Learning* 28: 221–282.

Andorno, C., G. Bernini, and A. G. Ramat. 2003. 'Sintassi'. In G. R. Anna (ed.). *Vesro l'italiano. Percorsi e Strategie di Acquisizione*. Roma: Carocci. 70–115.

Bailey, N., C. Madden, and S. D. Krashen. 1974. 'Is there a "natural sequence" in adult second language learning?' *Language Learning* 24: 235–243.

Banfi, E. and G. Bernini. 2003. 'Il verbo'. In G. R. Anna (ed.). *Vesro l'italiano. Percorsi e Strategie di Acquisizione*. Roma: Carocci. 70–115.

Bardovi-Harlig, K. 2009. 'Conventional expressions as a pragmalinguistic resource: Recognition and production of conventional expressions in L2 pragmatics'. *Language Learning* 59: 755–795.

Bardovi-Harlig, K. 2014. 'Awareness of meaning of conventional expressions in second-language pragmatics'. *Language Awareness* 23: 41–56.

Bardovi-Harlig, K. and D. Stringer. 2017. 'Unconventional expressions: Productive syntax in the L2 acquisition of formulaic language'. *Second Language Research* 33: 61–90.

Bardovi-Harlig, K. and H. E. Vellenga. 2012. 'The effect of instruction on conventional expressions in L2 pragmatics'. *System* 40: 77–89.

Bartke, S. and J. Siegmüller. 2004. *Williams Syndrome across languages*. Amsterdam: John Benjamins.

Baynham, M. 2006. 'Agency and contingency in the language learning of refugees and asylum seekers'. *Linguistics and Education* 17: 24–39.

Becker, J. 1975. 'The phrasal lexicon'. In R. Shank and B. Nash-Webbe (eds). *Theoretical Issues in Natural Language Processing*. Cambridge, MA: Bolt, BaranekNewman. 60–63.

Bellugi, U., K. Van Hoek, D. Lillo-Martin, and L. O'Grady. 1993. 'The acquisition of syntax in young deaf singers'. In D. Bishop and K. Mogford (eds). *Language Development in Exceptional Circumstances*. Hove: Lawrence Erlbaum. 132–150.

Benazzo, S. and M. Starren. 2007. 'L'émergence de moyens grammaticaux pour exprimer les relations temporelles en L2'. *Acquisition et Interaction en Langue étrangère* 25: 129–157.

Benmamoun, E. 1993. 'Functional and inflectional morphology problems of projection, representation and derivation.' Unpublished PhD dissertation, University of Southern California.

Benmamoun, E. 2000. *The Feature Structure Of Functional Categories: A Comparative Study Of Arabic Dialects*. Oxford: Oxford University Press.

Berko, J. 1958. 'The child's learning of English morphology'. *Word* 14: 150–177.

Bernini, G. 1989. 'Strategie di costruzione dei paradigmi verbali in italiano lingua seconda'. *Quaderni del Dipartimento di Linguistica e Letterature Comparate Dell'Università di Bergamo* 5: 195–207.

Biber, D. 1991. *Variation Across Speech and Writing*. Cambridge: Cambridge University Press.

Bickerton, D. 1975. *Dynamics of A Creole System*. Cambridge: Cambridge University Press.

Bigelow, M., R. Delmas, K. Hansen, and E. Tarone. 2006. 'Literacy and the processing of oral recasts in SLA'. *TESOL Quarterly* 40: 665–689.

Bigelow, M. and E. Tarone. 2004. 'The role of literacy level in second language acquisition: Doesn't who we study determine what we know?' *TESOL Quarterly* 38: 689–700.

Bishop, D. 2001. *Language and Cognitive Processes in Developmental Disorders*. Hove: Psychology Press.

Bliss, H. 2006. 'L2 acquisition of inflectional morphology: Phonological and morphological transfer effects'. In M. Grantham O'Brien, C. Shea, and J. Archibald (eds). *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference*. Somerville, MA: Cascadilla Proceedings Project. 1–8.

Blom, E. and S. De Korte. 2011. 'Dummy auxiliaries in child and adult second language acquisition of Dutch'. *Lingua* 121: 906–919.

Bonilla, C. L. 2015. 'Instructing stages of Processability Theory in L2 Spanish'. In K. Baten, A. Buyl, K. Lochtman, and M. Van Herreweghe (eds). *Theoretical and methodological developments in Processability Theory*. John Benjamins. 205–238.

Brattico, P. 2009. 'Long-distance case assignment in Finnish and the theory of phases'. *Biolinguistica Fennica Working Papers* 1: 79–104.

Brattico, P. 2010. 'Multiple long-distance case assignment and its syntax'. *Manuscript, University of Jyväskylä*:

Brown, R. and C. Hanlon. 1970. 'Derivational complexity and order of acquisition in child speech'. In J. R. Hayes (ed.). *Cognition and the Development of Language*. New York: Joensuu University Press. 11–53.

Brown, R. 1973. *A First Language: The Early Stage*. Cambridge, MA: Harvard University Press.

Cancino, H., E. J. Rosansky, and J. H. Schumann. 1978. 'The acquisition of English negatives and interrogatives by native Spanish speakers'. In E. M. Hatch (ed.). *Second Language Acquisition: A Book of Readings*. Rowley, MA: Newbury House Publishers, Inc. 207–380.

Carroll, S. 1999. 'Input and SLA: Adults' sensitivity to different sorts of cues to French gender'. *Language Learning* 49: 37–92.

Carroll, S. 2001. *Input and Evidence: The Raw Material of Second Language Acquisition*. Amsterdam: John Benjamins.

Carter, R. 1998. *Mapping the Mind*. London: Weidenfeld and Nicolson.

Cazden, C. B. 1968. 'The acquisition of noun and verb inflections'. *Child Development* 39: 433–448.

Chafe, W. L. 1976. 'Givenness, contrastiveness, definiteness, subjects, topics, and point of view'. In C. N. Li (ed.). *Subject and Topic*. New York: Academic Press. 25–55.

Chen, D. 2017. 'Early stages of adult L2 acquisition of mandarin: An Organic Grammar approach'. Unpublished PhD dissertation, Newcastle University.

Chomsky, N. 1957. 'Logical structures in language'. *American Documentation* 8: 284–301.

Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, N. 1986. *Language and Problems of Knowledge*. Cambridge, Massachusetts: MIT Press.

Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chun, A., R. R. Day, A. Chenoweth, and S. Luppescu. 1982. 'Errors, interaction, and correction: A study of native-normative conversations'. *TESOL Quarterly* 16: 537–547.

Cintrón-Valentín, M. C. and N. C. Ellis. 2016. 'Salience in second language acquisition: Physical form, learner attention, and instructional focus'. *Frontiers in Psychology* 7: 12–84.

Clahsen, H. 1984. 'The acquisition of German word order: A test case for cognitive approaches to L2 development'. In R. Andersen (ed.). *Second languages*. Rowley, MA: Newbury House. 219–42.

Clahsen, H. 1983. 'Some more remarks on the acquisition of German negation'. *Journal of Child Language* 10: 465–469.

Clahsen, H. 1990. 'Constraints on parameter setting: A grammatical analysis of some acquisition in stages in German child language'. *Language Acquisition* 1: 361–391.

Clahsen, H., S. Eisenbeiss, and A. Vainikka. 1994. 'The seeds of structure; A syntactic analysis of the acquisition of case marking'. *Language Acquisition Studies in Generative Grammar* 8: 85–118.

Clahsen, H., J. M. Meisel, and M. Pienemann. 1983. *Deutsch Als Zweitsprache: Der Spracherwerb Ausländischer Arbeiter [translation needed]*. Tübingen: Gunter Narr.

Clahsen, H. and M. Penke. 1992. 'The acquisition of agreement morphology and its syntactic consequences: New evidence on German child language from the Simone-corpus'. In M. Jürgen (ed.). *The Acquisition of Verb Placement*. Dordrecht: Kluwer. 181–223.

Clark, R. and I. Roberts. 1993. 'A computational model of language learnability and language change'. *Linguistic Inquiry* 24: 299–345.

Clark, R. 1974. 'Performing without competence'. *Journal of Child Language* 1: 1–10.

Condelli, L., H. S. Wrigley, and K. Yoon. 2002. 'What Works Study for Adult ESL Literacy Students'. *Final Report. Washington, DC: American Institutes for Research-Aguirre International.*:

Cooke, M. 2006. 'When I wake up I dream of electricity: The lives, aspirations and "needs" of adult ESOL learners'. *Linguistics and Education* 17: 56–73.

Corder, S. P. 1967. 'The significance of learner's errors.' *International Review of Applied Linguistics* 5: 161–170.

Cox, M. 2005. 'L2 English morpheme acquisition order: The lack of consensus examined from a case study of four L1 Chinese pre-school boys'. *Working Papers in Educational Linguistics* 20: 59–78.

Craats, I. van de, J. Kurvers, and M. Young-Scholten. 2005. 'Research on low-educated second language and literacy acquisition'. In I. van de Craats, J. Kurvers, and M. Young-Scholten (eds). *Low-educated Second Language and Literacy Acquisition.* Utrech: Landelijki Onderzoekschool Taalwetenschap (LOT). 45–68.

d'Avis, F.-J. and P. Gretsch. 2011. 'Variations on "variation": On the acquisition of complementizers in German'. In R. Tracy and E. Lattey (eds). *How Tolerant is Universal Grammar? Essays on Language Learnability and Language Variation.* Tübingen: Niemeyer. 59–109.

Dąbrowska, E. 2012. 'Different speakers, different grammars: Individual differences in native language attainment'. *Linguistic Approaches to Bilingualism* 2: 219–253.

Dąbrowska, E. and J. Street. 2006. 'Individual differences in language attainment: Comprehension of passive sentences by native and non-native English speakers'. *Language Awareness* 28: 604–615.

De Villiers, P. G. and J. M. de Villiers. 1973. 'A cross-sectional study of the acquisition of grammatical morphemes in child speech'. *Journal of Psycholinguistic Research* 2: 267–278.

Déprez, V. and A. Pierce. 1993. 'Negation and functional projections in early grammar'. *Linguistic Inquiry* 24: 25–67.

Dimroth, C. 2002. 'Topics, assertions, and additive words: how L2 learners get from information structure to target-language syntax'. *Languages* 40: 891–923.

Dimroth, C. 2008. 'German at different ages: Negation and finiteness in adult, adolescent and child learners'. *Language Learning* 58: 117–150.

Dornyei, Z. 2009. *The Psychology of Second Language Acquisition.* Oxford: Oxford University Press.

Dulay, H. and M. Burt. 1973. 'Should we teach children syntax?' *Language Learning* 23: 245–258.

Dulay, H. and M. Burt. 1974. 'Errors and strategies in child second language acquisition'. *TESOL Quarterly* 8: 129–136.

Dulay, H. and M. Burt. 1974. 'Natural sequences in child second language acquisition'. *Language Learning* 24: 37–53.

Eamer, A., S. Fernando, and A. E. King. 2017. 'Still on the margins: Migration, English language learning, and mental health in immigrant psychiatric patients'. *Diaspora, Indigenous,Minority Education* 11: 190–202.

Ehri, L. 1994. 'Development of the ability to read words'. In R. Ruddell, M. Ruddell, and H. Singer (eds). *Theoretical Models and Processes of Reading.* Newark, DE: International Reading Assiociation. 323–358.

Ellis, N. C. 2002. 'Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition'. *Studies in Second Language Acquisition* 24: 143–188.

Ellis, N. C. 2003. 'Constructions, chunking, and connectionism: The emergence of second language structure.' In C. Doughty and M. Long (eds). *Handbook of Second Language Acquisition.* Oxford: Blackwell. 33–68.

Ellis, N. C. 2008. 'Optimizing the input: Frequency and sampling in usage-based and form-focussed learning'. In M. H. Long and C. Doughty (eds). *Handbook of Second and Foreign Language Teaching.* Oxford: Blackwell. 34–64.

Ellis, N. C. 2012. 'Formulaic language and second language acquisition: Zipf and the phrasal teddy bear'. *Annual Review of Applied Linguistics* 32: 17–44.

Ellis, N. C., R. Simpson-Vlach, and C. Maynard. 2008. 'Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics, and TESOL'. *TESOL Quarterly* 42: 375–396.

Ellis, R. 1990. *Instructed Second Language Acquisition.* Oxford: Blackwell.

Ellis, R. 1999. *Form-focused Instruction and Second Language Learning.* London: Lawrence Erlbaum.

Epstein, S. D., S. Flynn, and G. Martohardjono. 1996. 'Second language acquisition: Theoretical and experimental issues in contemporary research'. *Behavioral and Brain Sciences* 19: 677–714.

Erman, B. and B. Warren. 2000. 'The idiom principle and the open choice principle'. *Text & Talk* 20: 29–62.

Eskildsen, S. W. and T. Cadierno. 2007. 'Are recurring multi-word expressions really syntactic freezes? Second language acquisition from the perspective of usage-based linguistics'. In M. Nenonen and S. Niemi (eds). *Collocations and Idioms 1: Papers from the First Nordic Conference on Syntactic Freezes*. Joensuu, Finland: Joensuu University Press. 19–20.

Eskildsen, S. W. 2009. 'Constructing another language—Usage-based linguistics in second language acquisition'. *Applied Linguistics* 30: 335–357.

Eskildsen, S. W. 2012. 'L2 negation constructions at work'. *Language Learning* 62: 335–372.

Eubank, L. 1992. 'Verb movement, agreement and tense in L2 acquisition'. In J. M. Meisel (ed.). *The acquisition of Verb Placement: Functional Categories and V2 Phenomena in Language Development*. Dordrecht: Kluwer. 225–244.

Eubank, L. 1993. 'On the transfer of parametric values in L2 development'. *Language Acquisition* 3: 183–208.

Eubank, L. 1996. 'Negation in early German-English interlanguage: More valueless features in the L2 initial state'. *Second Language Research* 12: 73–106.

Fleta, M. T. 2003. 'Is-insertion in L2 grammars of English: A step forward between Developmental stages?' In J. M. Liceras, H. Zobl, and H. Goodluck (eds). *Proceedings of the 6th Generative Approaches to Second Language Acquisition Conference*. Somerville, MA: Cascadilla Proceedings Project. 85–96.

Fodor, J. A. 1983. *The Modularity of Mind*. Cambridge: MIT Press.

Foster, D. and P. Bolton. 2018. *Adult ESOL in England*. Last accessed 31 July 2020. URL: `https://commonslibrary.parliament.uk/research-briefings/cbp-7905/`.

Foster, P. 2001. 'Rules and routines: A consideration of their role in the task-based language production of native and non-native speakers.' In M. Bygate, P. Skehan, and M. Swain (eds). *Researching Pedagogic Tasks: Second Language Learning, Teaching, and Testing*. Harlow: Longman. 75–94.

Friederici, A. D. 2017. *Language in our Brain: The Origins of a Uniquely Human Capacity*. Cambridge: MIT Press.

García Mayo, M. P., A. L. Ibarrola, and J. M. Liceras. 2005. 'Placeholders in the English interlanguage of bilingual (Basque/Spanish) children'. *Language Learning* 55: 445–489.

Gibson, E. and K. Wexler. 1994. 'Triggers'. *Linguistic Inquiry* 25: 407–454.

Goldin-Meadow, S. and C. Mylander. 1993. 'Beyond the input given: The child's role in the acquisition of language'. In P. Bloom (ed.). *Language Acquisition: Core Readings*. London: Wheatsheaf. 507–542.

Goldschneider, J. M. and R. M. De Keyser. 2001. 'Explaining the "natural order of L2 morpheme acquisition" in English: A meta-analysis of multiple determinants'. *Language Learning* 51: 1–50.

Gombert, J.-E. 1992. *Meta-linguistic Development*. London: Harvester Wheatsheaf.

Gopnik, M. and M. B. Crago. 1991. 'Familial aggregation of a developmental language disorder'. *Cognition* 39: 1–50.

Gordon, D. M. 2011. 'Trauma and second language learning among Laotian refugees'. *Journal of Southeast Asian American Education & Advancement* 6: 1–17.

Goswami, U. and P. Bryant. 1990. *Phonological Skills and Learning to Read*. London: Lawrence Erlbaum.

Grondin, N. and L. White. 1996. 'Functional categories in child L2 acquisition of French'. *Language Acquisition* 5: 1–34.

Grondin, N. and L. White. 1996. 'Functional categories in child L2 acquisition of French'. *Language Acquisition* 9: 20–42.

Guasti, M. T. 2002. *Language Acquisition: The Growth of Grammar*. Cambridge: MIT Press.

Guilfoyle, E. and M. Noonan. 1992. 'Functional categories and language acquisition'. *Canadian Journal of Linguistics* 37: 241–272.

Haegeman, L. 1995. *The Syntax of Negation*. Cambridge: Cambridge University Press.

Haegeman, L. and J. Gueron. 1999. *English Grammar: A Generative Perspective*. Oxford: Black Well.

Hakuta, K. 1974. 'Prefabricated patterns and the emergence of structure in second language acquisition'. *Language Learning* 24: 287–297.

Hakuta, K. 1976. 'A case study of a Japanese child learning English as a second language'. *Language Learning* 26: 321–351.

Hall, T. 2010. 'L2 Learner-made formulaic expressions and constructions'. *Studies in Applied Linguistics and TESOL* 10: 1–18.

Halliday, M. A. 1994. *An Introduction to Functional Grammar*. London: Edward Arnold.

Hasselgren, A. 1994. 'Lexical teddy bears and advanced learners: A study into the ways Norwegian students cope with English vocabulary'. *International Journal of Applied Linguistics* 4: 237–258.

Hatch, E. 1972. 'Some studies in language learning'. *UCLA Workpapers in Teaching English as a Second Language* 9: 29–36.

Hawkins, R. 2001. *Second Language Syntax: A Generative Introduction*. Cambridge: Wiley-Blackwell.

Haznedar, B. 1997. 'L2 acquisition by a Turkish-speaking child: Evidence for L1 influence'. In E. Hughes, M. Hughes, and A. Greenhill (eds). *Proceedings of the 21st Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press. 245–256.

Haznedar, B. 2001. 'The acquisition of the IP system in child L2 English'. *Studies in Second Language Acquisition* 23: 1–39.

Haznedar, B. 2003. 'The status of functional categories in child second language acquisition: Evidence from the acquisition of CP'. *Second Language Research* 19: 1–41.

Haznedar, B. and B. D. Schwartz. 1997. 'Are there optional infinitives in child L2 acquisition'. *Proceedings of the 21st Annual Boston University Conference on Language Development* 21: 257–268.

Hill, J. H. 1970. 'Foreign accents, language acquisition, and cerebral dominance revisited'. *Language Learning* 20: 237–248.

Hinton, D. E., A. Rasmussen, L. Nou, M. H. Pollack, and M.-J. Good. 2009. 'Anger, PTSD, and the nuclear family: A study of Cambodian refugees'. *Social Science & Medicine* 69: 1387–1394.

Hornstein, N. and D. Lightfoot. 1981. 'Introduction'. In N. Hornstein and D. Lightfoot (eds). *Explanation in Linguistics: The Logical Problem of Language Acquisition*. New York: Longman. 9–31.

Horsman, J. 2000. *Too Scared to Learn: Women, Violence, and Education*. Mahwah, NJ: Lawrence Erlbaum.

Howarth, P. 1998. 'The phraseology of learners' academic writing'. In A. P. Cowie (ed.). *Phraseology: Theory, Analysis, and Application*. Oxford: Oxford University Press. 161–186.

Huang, J. S. 1971. 'A Chinese child's acquisition of English syntax'. Unpublished PhD dissertation, University of California, Los Angeles.

Ingram, D. 1989. *First Language Acquisition: Method, Description and Explanation*. Cambridge: Cambridge University Press.

Ionin, T. and K. Wexler. 2001. 'L1 Russian children learning English: Tense and overgeneralization of be'. In A. Notfound (ed.). *Proceedings of the 2000 Second Language Research Forum*. Somerville, MA: Casca- Dilla Press. 76–94.

Ionin, T. and S. Montrul. 2010. 'The role of L1 transfer in the interpretation of articles with definite plurals in L2 English'. *Language Learning* 60: 877–925.

Jackendoff, R. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.

Jackendoff, R. 1977. *X-Bar Theory: A Study of Phrase Structure*. Cambridge, MA: MIT Press.

Jenkins, L. 2000. *Biolinguistics*. Cambridge: Cambridge University Press.

Jordens, P. 1996. 'Input and instruction in second language acquisition'. In P. Jordens and J. Lalleman (eds). *Investigating Second Language Acquisition*. Berlin: De Gruyter. 407–449.

Jordens, P. and C. Dimroth. 2006. 'Finiteness in children and adults learning Dutch'. In N. Gagarina and I. Gülzow (eds). *The acquisition of Verbs and their Grammar: The Effect of Particular Languages*. Dordrecht: Springer. 173–200.

Julien, M., R. Van Hout, and I. Van de Craats. 2016. 'Meaning and function of dummy auxiliaries in adult acquisition of Dutch as an additional language'. *Second Language Research* 32: 49–73.

Kahoul, W. 2014. 'Arabic and Chinese learners' production, perception and processing of past tense and verbal agreement morphology in L2 English'. Unpublished PhD dissertation, Newcastle University.

Kahoul, W., A. Vainikka, and M. Young-Scholten. 2018. 'The mystery of the missing inflections'. In C. Wright, T. Piske, and M. Young-Scholten (eds). *Mind Matters in SLA*. Bristol: Multilingual Matters. 176–189.

Kail, M. 2002. 'Sentence processing studies and linguistic literacy'. *Journal of Child Language* 29: 449–488.

Karmiloff-Smith, A. 1986. 'Some fundamental aspects of language acquisition after five'. In P. Fletcher and M. Garman (eds). *Studies in Language Acquisition*. Cambridge: Cambridge University Press. 455–474.

Kato, M. A. 1999. 'Strong and weak pronominals in the null subject parameter'. *Probus* 11: 1–27.

Kegl, J. 1994. 'The Nicaraguan sign language project: An overview'. *Signpost* 7: 24–31.

Kempen, G. 1998. 'Comparing and explaining the trajectories of first and second language acquisition: In search of the right mix of psychological and linguistic factors'. *Bilingualism: LanguageCognition* 1: 29–30.

Kim, K. 2011. 'Overgenerated be from topic marker to verbal inflection'. *Foreign Language Education* 14: 34–55.

Klein, E., I. Stoyneshka, K. Adams, Y. Pugach, S. Solt, and T. Rose. 2004. 'Past tense affixation in L2 English'. In L. Brugos, L. Micciulla, and C. E. Smith (eds). *Proceedings Supplement of the 28th Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press. 553–564.

Klein, W., R. Dietrich, and C. Noyau. 1995. 'Conclusions'. In R. Dietrich, W. Klein, and C. Noyau (eds). *The Acquisition of Temporality in a Second Language*. Amsterdam: John Benjamins. 261–280.

Klein, W. and N. Dittmar. 1999. *Developing Grammars: The acquisition of German Syntax by Foreign Workers*. Oxford: Black Well.

Klein, W. and C. Perdue. 1992. *Utterance structure: Developing grammars again*. Co., Amsterdam: Benjamins.

Klein, W. and C. Perdue. 1997. 'The Basic Variety (or: Couldn't natural languages be much simpler?)' *Second Language Research* 13: 301–347.

Krashen, S. 1985. *The Input Hypothesis*. London: Longman.

Krashen, S. and R. Scarcella. 1978. 'On routine and Pattern in Language Acquisition and performance'. *Language Learning* 28: 283–300.

Kurvers, J., W. Stockmann, and I. van de Craats. 2010. 'Predictors of success in adult L2 literacy acquisition'. In T. Wall and M. Leong (eds). *Low-Educated Second Language and Literacy Acquisition*. Calgary: Bow Valley College. 47–62.

Kvam, P., B. Vidakovic, and S.-j. Kim. 2022. *Nonparametric Statistics with Applications to Science and Engineering with R*. Hoboken: John Wiley & Sons.

Laka Mugarza, M. I. 1990. 'Negation in syntax–on the nature of functional categories and projections'. Unpublished PhD dissertation, Massachusetts Institute of Technology.

Lakshmanan, U. 2014. 'Functional categories and related mechanisms in child second language acquisition'. In S. Flynn, G. Martohardjono, and W. O'Neil (eds). *The Generative Study of Second Language Acquisition*. New York: Psychology Press. 281–306.

Lakshmanan, U. 1993/94. 'The Boy for the cookie–Some evidence for the nonviolation of the Case Filter in child second language acquisition'. *Language Acquisition* 3: 55–91.

Lantolf, J. P. 2000. *Sociocultural Theory and Second Language Learning*. Oxford: Oxford University Press.

Lardiere, D. 2000. 'Mapping features to forms in second language acquisition.' In J. Archibald (ed.). *Second Language Acquisition and Linguistic Theory*. Oxford: Blackwell. 102–29.

Lardiere, D. 2003. 'Second language knowledge of [+/-past] vs. [+/-finite]'. In J. Liceras, H. Goodluck, and H. Zobl (eds). *Proceedings of the 6th Generative Approaches to Second Language Acquisition*. Somerville, MA: Cascadilla Press. 176–189.

Lardiere, D. 1998. 'Case and tense in the 'fossilized'steady state'. *Second Language Research* 14: 1–26.

Lardiere, D. 1998. 'Dissociating syntax from morphology in a divergent L2 end-state grammar'. *Second Language Research* 14: 359–375.

Lebeaux, D. 1989. 'Language acquisition and the form of the grammar'. Unpublished PhD dissertation, University of Massachusetts.

Li, C. N. and S. A. Thompson. 1976. 'Subject and topic: A new typology'. In C. N. Li (ed.). *Subject and Topic*. New York: Academic Press. 457–589.

Liu, G. 1991. 'Interaction and second language acquisition: A case study of a Chinese child's acquisition of English as a second language.' Unpublished PhD dissertation, LaTrobe University.

Loewen, S. and M. Sato. 2018. 'Interaction and instructed second language acquisition'. *Language Teaching* 51: 285–329.

Long, M. H. 1990. 'The least a second language acquisition theory needs to explain'. *TESOL Quarterly* 24: 649–666.

Long, M. H., S. Inagaki, and L. Ortega. 1998. 'The role of implicit negative feedback in SLA: Models and recasts in Japanese and Spanish'. *The Modern Language Journal* 82: 357–371.

Long, M. 2015. *Second Language Acquisition and Task-based Language Teaching*. Malden, MA: John Wiley & Sons.

Lorenzo, G. and V. M. Longa. 2003. 'Minimizing the genes for grammar. The minimalist program as a biological framework for the study of language'. *Lingua* 113: 643–657.

Loureiro, C., L. Braga, L. Souza, G. Filho, E. Queiroz, and G. & Dellatolas. 2004. 'Degree of illiteracy and phonological and metaphonological skills in unschooled adults'. *Brain and Language* 89: 499–502.

Lyster, R. 1998. 'Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classrooms'. *Language Learning* 48: 183–218.

Lyster, R. and L. Ranta. 1997. 'Corrective feedback and learner uptake: Negotiation of form in communicative classrooms'. *Studies in Second Language Acquisition* 19: 37–66.

Mackey, A., S. Gass, and K. McDonough. 2000. 'How do learners perceive interactional feedback?' *Studies in Second Language Acquisition* 22: 471–497.

Mackey, A. and J. Philp. 1998. 'Conversational interaction and second language development: Recasts, responses, and red herrings?' *The Modern Language Journal* 82: 338–356.

McDonald, S. 2000. 'Trauma and second language learning'. *Canadian Modern Language Review* 56: 690–696.

Meisel, J. M. 1997. 'The acquisition of the syntax of negation in French and German: Contrasting first and second language development'. *Second Language Research* 13: 227–263.

Meisel, J. M., H. Clahsen, and M. Pienemann. 1981. 'On determining developmental stages in natural second language acquisition'. *Studies in Second Language Acquisition* 3: 109–135.

Miller, J. 2002. 'Questions about constructions'. *Journal of Child Language* 29: 449–488.

Mishra, R. K., N. Singh, A. Pandey, and F. Huettig. 2012. 'Spoken language-mediated anticipatory eye movements are modulated by reading ability: Evidence from Indian low and high literates'. *Journal of Eye Movement Research* 5: 1–10.

Mobaraki, M. 2007. 'The acquisition of English by two Farsi-speaking children'. Unpublished PhD dissertation, Durham University.

Mocciaro, E. 2019. 'Emerging constructions in the L2 Italian spoken by low literate migrants'. *Languages* 4: 1–24.

Mocciaro, E. 2019b. 'The acquisition of L2 Italian morphosyntax by low- or non-literate adult learners'. Unpublished PhD dissertation, University of Palermo.

Mollica, R. F., G. Wyshak, and J. Lavelle. 1987. 'The psychosocial impact of war trauma and torture on Southeast Asian refugees.' *The American Journal of Psychiatry* 144: 1567–1572.

Molony, C. 1977. '"Ich bin sprechen Deutsch aber": the sequence of verb and word order acquisition of an American child learning German'. In B. VanPatten, G. D. Keating, and S. Wulff (eds). *German in Contact with other Languages*. Von Privat, Bewertungen: Scriptor Verlag Kronberg/Ts. 274–295.

Morais, J., P. Bertelson, L. Cary, and J. Alegría. 1986. 'Literacy training and speech segmentation'. *Cognition* 24: 45–64.

Mourssi, A. 2013. 'Crosslinguistic influence of L1 (Arabic) in acquiring linguistic items of L2 (English): An empirical study in the context of Arab learners of English as undergraduate learners.' *Theory & Practice in Language Studies* 3: 397–403.

Muftah, M. 2016. 'English past tense morphology in adult Arab EFL learners: Mental representation mechanism and types of errors'. *Journal of Advances in Linguistics* 6: 38–52.

Myles, F. 2004. 'From data to theory: The over-representation of linguistic knowledge in SLA'. *Transactions of the Philological Society* 102: 139–168.

Myles, F., J. Hooper, and R. Mitchell. 1998. 'Rote or rule? Exploring the role of formulaic language in classroom foreign language learning'. *Language Learning* 48: 323–364.

Myles, F., R. Mitchell, and J. Hooper. 1999. 'Interrogative chunks in French L2: A basis for creative construction?' *Studies in Second Language Acquisition*: 49–80.

Nattinger, J. R. and J. S. DeCarrico. 1992. *Lexical Phrases and Language Teaching.* Oxford: Oxford University Press.

Nicholl, C. and A. Thompson. 2004. 'The psychological treatment of Post Traumatic Stress Disorder (PTSD) in adult refugees: A review of the current state of psychological therapies'. *Journal of Mental Health* 13: 351–362.

Olson, D. 2002. 'What writing does to the mind'. In E. Amsel and J. Byrnes (eds). *Language, Literacy, and Cognitive Development: The Development and Consequences of Symbolic Communication.* Mahwah, NJ: Lawrence Erlbaum Association. 153–166.

Ong, W. J. 1988. *Orality and Literacy.* London: Routledge.

Ortega, L. and M. H. Long. 1997. 'The effects of models and recasts on the acquisition of object topicalization and adverb placement in L2 Spanish'. *Spanish Applied Linguistics* 1: 65–86.

Osborne, J. 2008. 'Phraseology effects as a trigger for errors in L2 English: The case of more advanced learners'. In F. Meunier and S. Granger (eds). *Phraseology in Foreign Language Learning and Teaching.* Amsterdam: John Benjamins. 67–83.

Ouhalla, J. 1993. 'Functional categories, agrammatism and language acquisition'. *Linguistische Berichte* 143: 3–36.

Paradis, M. 2004. *A Neurolinguistic Theory of Bilingualism.* Amsterdam: John Benjamins.

Pawley, A. and F. H. Syder. 1983. 'Two puzzles for linguistic theory: Nativelike selection and nativelike fluency'. In J. C. Richards and R. W. Schmidt (eds). *Language and Communication.* London: Longman. 191–225.

Peters, A. M. 1983. *The Units of Language Acquisition.* Cambridge: Cambridge University Press.

Philp, J. 2003. 'Constraints on "noticing the gap": Nonnative speakers' noticing of recasts in NS-NNS interaction'. *Studies in Second Language Acquisition* 25: 99–126.

Piaget, J., B. Inhelder, and P. Fraise. 1969. 'Intellectual operations and their development'. *Experimental Psychology* 33: 144–205.

Pienemann, M. 1998. *Language Processing and Second Language Development: Processability Theory.* Amsterdam and Philadelphia: John Benjamins.

Piske, T., I. R. MacKay, and J. E. Flege. 2001. 'Factors affecting degree of foreign accent in an L2: A review'. *Journal of Phonetics* 29: 191–215.

Platzack, C. 1990. 'A Grammer without Functional Categories: A Syntactic Study of Early Swedish Child Language'. *Nordic Journal of Linguistics* 13: 107–126.

Pollock, J.-Y. 1989. 'Verb movement, universal grammar, and the structure of IP'. *Linguistic Inquiry* 20: 365–424.

Preston, D. 2000. 'Three kinds of sociolinguistics and SLA: A psycholinguistic perspective'. In F. Swierzbin, M. Anderson, C. Klee, and E. Tarone (eds). *Social and Cognitive Factors in Second Language Acquisition: Selective Proceedings of the 1999 Second Language Research Forum*. Cascadilla: Somerville, MA. 140–158.

Prévost, P. and L. White. 2000. 'Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement'. *Second Language Research* 16: 103–133.

Pylyshyn, Z. W. 1980. 'Computation and cognition: Issues in the foundations of cognitive science'. *Behavioral and Brain Sciences* 3: 111–132.

Radford, A. 1988. 'Small children's small clauses'. *Transactions of the Philological Society* 86: 1–43.

Radford, A. 1990. *Syntactic Theory and the Acquisition of English Syntax*. Oxford: Blackwell.

Radford, A. 1995. 'Children: Architects or brickies?' In M. Dawn and M. Susan (eds). *Proceedings of the 19th Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press. 1–19.

Ravem, R. 1974. 'The development of wh-questions in first and second language learners'. In J. C. Richards (ed.). *Error Analysis: Perspectives on Second Language Acquisition*. London: Longman. 134–55.

Ravid, D. and L. Tolchinsky. 2002. 'Developing linguistic literacy: A comprehensive model'. *Journal of Child Language* 29: 417–447.

Read, C., Z. Yun-Fei, N. Hong-Yin, and D. Bao-Qing. 1986. 'The ability to manipulate speech sounds depends on knowing alphabetic writing'. *Cognition* 24: 31–44.

Reis, A. and A. Castro-Caldas. 1997. 'Illiteracy: A cause for biased cognitive development'. *Journal of the International Neuropsychological Society* 3: 444–450.

Rizzi, L. 1993/4. 'Some notes on linguistic theory and language development: The case of root infinitives'. *Language Acquisition* 3: 371–393.

Roberts, C., M. Cooke, M. Baynham, and J. Simpson. 2007. 'Adult ESOL in the United Kingdom: Policy and research'. *AMEP Research Centre* 22: 18–31.

Robinson, P. 1996. *Consciousness, Rules and Instructed Second Language Acquisition*. Bern: Peter Lang.

Romaine, S. 1984. *The Language of Children and Adolescents: The Acquisition of Communicative Competence*. Oxford: Blackwell.

Rosansky, E. J. 1976. 'Methods and morphemes in second language acquisition research 1'. *Language Learning* 26: 409–425.

Sakas, W. and J. D. Fodor. 2001. 'The structural triggers learner'. In S. Bertolo (ed.). *Language Acquisition and Learnability*. Cambridge: Cambridge University Press. 172–233.

Santoro, N. 1997. 'Why won't they talk: The difficulties of engaging victims of trauma in classroom interaction'. *TESOL in Context* 7: 14–18.

Scarcella, R. 1976. 'Watch up! Working Papers on Bilingualism'. *Working Papers on Bilingualism* 19: 79–88.

Schellekens, P. 2007. *The Oxford ESOL Handbook*. Oxford: Oxford University Press.

Schimke, S. 2009. 'The acquisition of finiteness by Turkish learners of German and Turkish learners of French: Investigating knowledge of forms and functions in production and comprehension'. Unpublished PhD dissertation, Max Planck Institute for Psycholinguistics, Nijmegen.

Schmidt, R. W. 1990. 'The role of consciousness in second language learning'. *Applied Linguistics* 11: 129–158.

Schmidt, R. W. 1994. 'Deconstructing consciousness in search of useful definitions for applied linguistics'. *Consciousness in Second Language Learning* 11: 237–326.

Schwartz, B. D. 1993. 'On explicit and negative data effecting and affecting competence and linguistic behavior'. *Studies in Second Language Acquisition* 15: 147–163.

Schwartz, B. D. 1997. 'On the basis of the Basic Variety'. *Second Language Research* 13: 386–402.

Schwartz, B. D. and R. Sprouse. 1994. 'Word order and nominative case in nonnative language acquisition: A longitudinal study of (L1 Turkish) German interlanguage'. In T. Hoekstra and B. D. Schwartz (eds). *Language Acquisition Studies in Generative Grammar*. Amsterdam: Casca- Dilla Press. 317–68.

Schwartz, B. D. and R. Sprouse. 1996. 'L2 cognitive states and the full transfer/full access model'. *Second Language Research* 12: 40–72.

Selinker, L. 1974. 'interlanguage'. *International Review of Applied Linguistics* 10: 209–231.

Sharwood Smith, M. 1994. *Second Language Learning: Theoretical Foundations*. London: Longman.

Shibata, M. 2006. 'Topic marking in English composition by Japanese EFL learners'. *Scripsimus* 14: 1–26.

Shin, J. 2001. 'L1 influence in foreign language learning: Topic-prominence in Korean EFL learners' interlanguage grammar'. *Foreign Language Education* 8: 1–21.

Slobin, D. I. 1997. *The Crosslinguistic Study of Language Acquisition: Expanding the Contexts*. Mahwah, NJ: Lawrence Erlbaum.

Smith, N. and I. Tsimpli. 1995. *The Mind of a Savant: Language Learning and Modularity*. Oxford: Blackwell.

Smith, N., I. Tsimpli, G. Morgan, and B. Woll. 2010. *The Signs of a Savant: Language against the Odds*. Cambridge: Cambridge University Press.

Söndergaard, H. P. and T. Theorell. 2004. 'Language acquisition in relation to cumulative posttraumatic stress disorder symptom load over time in a sample of resettled refugees'. *PsychotherapyPsychosomatics* 73: 320–323.

Sorhus, H. B. 1977. 'To hear ourselves—Implications for teaching English as a second language'. *ELT Journal* 31: 211–221.

Spada, N. and P. M. Lightbown. 1999. 'Instruction, first language influence, and developmental readiness in second language acquisition'. *The Modern Language Journal* 83: 1–22.

Starren, M. 2001. 'The second time: The acquisition of temporality in Dutch and French as a second language'. Unpublished PhD dissertation, Tilburg University.

Stauble, A.-M. 1984. 'A comparison of a Spanish-English and a Japanese-English second language continuum: Negation and verb morphology'. In R. Anderson (ed.). *Second Languages: A Cross-linguistic Perspective*. Rowley, MA: Newbury House. 323–353.

Stevens, C. A. 2001. 'Perspectives on the meanings of symptoms among Cambodian refugees'. *Journal of Sociology* 37: 81–98.

Tager-Flusberg, H. 1994. 'The relationship between language and social cognition: Lessons from autism'. In Y. Levy (ed.). *Other Children, Other Languages*. Hove: Erlbaum. 359–381.

Tarone, E. and M. Bigelow. 2005. 'Impact of literacy on oral language processing: Implications for SLA research'. *Annual Review of Applied Linguistics* 25: 77–97.

Tarone, E., M. Bigelow, and K. Hansen. 2009. *Literacy and Second Language Oracy*. Oxford: Oxford University Press.

Turner, S. 2018. 'Understanding the mental health needs of refugees: Trauma, stress, and coping'. *Journal of Refugee Studies* 31: 329–345.

Unsworth, S. 2005. 'Child L2, adult L2, child L1: Differences and similarities. A study on the acquisition of direct object scrambling in Dutch'. Unpublished PhD dissertation, Utrecht University.

Vainikka, A. 1993/4. 'Case in the development of English syntax'. *Language Acquisition* 3: 257–325.

Vainikka, A. and M. Young-Scholten. 1994. 'Direct access to X'-theory: Evidence from Korean and Turkish adults learning German'. In T. Hoekstra and B. D. Schwartz (eds). *Language Acquisition Studies in Generative Grammar*. Amsterdam: John Benjamins. 265–316.

Vainikka, A. and M. Young-Scholten. 1996. 'Gradual development of L2 phrase structure'. *Second Language Research* 12: 7–39.

Vainikka, A. and M. Young-Scholten. 1996. 'The early stages in adult L2 syntax: Additional evidence from Romance speakers'. *Second Language Research* 12: 140–176.

Vainikka, A. and M. Young-Scholten. 1998. 'Morphosyntactic triggers in adult SLA'. In M.-L. Beck (ed.). *Morphology and its Interfaces in Second Language Knowledge*. Amsterdam: John Benjamins. 89–114.

Vainikka, A. and M. Young-Scholten. 2004. 'Organic Grammar and SLA'. Talk Presented at the University of Pittsburgh.

Vainikka, A. and M. Young-Scholten. 2005. 'The roots of syntax and how they grow: Organic Grammar, the Basic Variety and Processability Theory'. In S. Unsworth, T. Parodi, A. Sorace, and M. Young-Scholten (eds). *Paths of Development*. Amsterdam: John Benjamins. 77–106.

Vainikka, A. and M. Young-Scholten. 2007. 'The role of literacy in the development of L2 morpho-syntax from an Organic Grammar perspective'. In N. Faux (ed.). *Research, Practice and Policy for Low-educated Second Language and Literacy Acquisition for Adults. Proceedings of the 2006 LESLLA Conference*. Richmond, VA: Literacy Institute at Virginia Commonwealth University. 123–184.

Vainikka, A. and M. Young-Scholten. 2010. 'All acquisition begins with the projection of a bare verb phrase'. *Applied Psycholinguistics* 31: 332–239.

Vainikka, A. and M. Young-Scholten. 2011. *The Acquisition of German: Introducing Organic Grammar*. Berlin: De Gruyter.

Vainikka, A. and M. Young-Scholten. 2019. 'Finding their heads: How immigrant adults posit L2 functional projections'. In D. Arteaga (ed.). *L2 Grammatical Representation and Processing: Theory and Practice*. Bristol: Multilingual Matters. 116–138.

Vainikka, A., M. Young-Scholten, C. Ijuin, and S. Jarad. 2017. 'Literacy in the development of L2 English morphosyntax'. In M. Sosinski (ed.). *Language and Literacy Teaching LESLLA Students*. Granada: University of Granada. 239–50.

Van de Craats, I. 2011. 'A LESLLA corpus: L1 obstacles in the learning of L2 morphosyntax'. In C. Schöneberger, I. van de Craats, and J. Kurvers (eds). *Low-Educated Adult Second Language and Literacy Acquisition. Proceedings of the 6th Symposium*. Nijmegen: Centre for Language Studies. 33–48.

Van de Craats, I. and R. Van Hout. 2010. 'Dummy auxiliaries in the second language acquisition of Moroccan learners of Dutch: Form and function'. *Second Language Research* 26: 473–500.

Van de Craats, I., J. Kurvers, and M. Young-Scholten. 2006. 'First-time L2 readers: Is there a critical period?' In I. Van de Craats, J. Kurvers, and M. Young-Scholten (eds). *Low-educated Second Language and Literacy Acquisition*. Utrech: Landelijki Onderzoekschool Taalwetenschap (LOT). 45–68.

Van Kampen, J. 1997. *First Steps in Wh-movement*. Delft: Eburon.

VanPatten, B. 2007. 'Input processing in adult second language acquisition'. In B. VanPatten, G. D. Keating, and S. Wulff (eds). *Theories in Second Language Acquisition: An Introduction*. New York: Routledge. 115–135.

VanPatten, B., S. R. R. Jessica Williams, and O. Mark. 1996. *Form-Meaning Connections in Second Language Acquisition*. London: Lawrence Erlbaum.

Verhagen, J. 2009. 'Finiteness in Dutch as a second language'. Unpublished PhD dissertation, Free University of Amsterdam.

Verhagen, J. 2011. 'Verb placement in second language acquisition: Experimental evidence for the different behavior of auxiliary and lexical verbs'. *Applied Psycholinguistics* 32: 821–858.

Verhagen, J. 2013. 'From dummy auxiliary to auxiliary in Moroccan adult learners' production and comprehension of Dutch'. In E. Blom, I. Van de Craats, and J. Verhagen (eds). *Dummy Auxiliaries in First and Second Language Acquisition*. Berlin: De Gruyter Mouton. 281–306.

Wagner-Gough, J. 1975. 'Comparative studies in second language learning'. In E. Hatch (ed.). *Second Language Acquisition: A Book of Readings*. Rowley, MA: Newbury House. 76–90.

White, L. 1989. *Universal Grammar and Second Language Acquisition*. Amsterdam: John Benjamins.

White, L. 1996. 'Clitics in child L2 French'. In R. Hawkins and H. Clahsen (eds). *Generative Approaches to First and Second Language Acquisition*. Amsterdam: John Benjamins. 335–68.

White, L. 2003. *Second Language Acquisition and Universal Grammar*. Cambridge: Cambridge University Press.

Wong-Fillmore, L. 1976. 'The second time around: Cognitive and social strategies in second language acquisition.' Unpublished PhD dissertation, Stanford University.

Wray, A. 2002. *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.

Wray, A. 2008. *Formulaic Language: Pushing the Boundaries*. Oxford: Oxford University Press.

Wulff, S. 2019. 'Acquisition of formulaic language from a usage-based perspective'. In A. Siyanova– Chanturia and A. Pellicer–Sanchez (eds). *Understanding Formulaic Language: A second Language Acquisition Perspective*. New York: Routledge. 19–37.

Wulff, S. 2008. *Rethinking Idiomaticity: A Usage-based Approach*. London: Continuum.

Yang, S. 2014. 'The functions of the nontarget be in the written interlanguage of Chinese learners of English'. *Language Acquisition* 21: 279–303.

Yang, Y. and R. Lyster. 2010. 'Effects of form-focused practice and feedback on Chinese EFL learners' acquisition of regular and irregular past tense forms'. *Studies in Second Language Acquisition* 32: 235–263.

Ying, Y.-W. 2001. 'Psychotherapy with traumatized Southeast Asian refugees'. *Clinical Social Work Journal* 29: 65–78.

Young-Scholten, M. 2013. 'Low-educated immigrants and the social relevance of second language acquisition research'. *Second Language Research* 29: 441–454.

Young-Scholten, M. 2015. 'Who are adolescents and adults who develop literacy for the first time in an L2, and why are they of research interest?' *Writing Systems Research* 7: 1–3.

Young-Scholten, M. and R. Naeb. 2020. 'Acquisition and Assessment of Morphosyntax'. In J. K. Peyton and M. Young-Scholten (eds). *Teaching Adult Immigrants with Limited Formal Education: Theory, Research and Practice*. Bristol: Multilingual Matters. 79–103.

Young-Scholten, M. and N. Strom. 2005. 'First-time L2 readers: Is there a critical period?' In I. van de Craats, J. Kurvers, and M. Young-Scholten (eds). *Low-educated Second Language and Literacy Acquisition*. Utrech: Landelijki Onderzoekschool Taalwetenschap (LOT). 45–68.

Yuan, B. 2001. 'The status of thematic verbs in the second language acquisition of Chinese: Against inevitability of thematic-verb raising in second language acquisition'. *Second Language Research* 17: 248–272.

Yuan, B. 2004. 'Negation in French–Chinese, German–Chinese and English–Chinese interlanguages'. *Transactions of the Philological Society* 102: 169–197.

Yule, G. 1997. *Referential Communication Tasks*. Cambridge: Cambridge University Press.

Zanuttini, R. 1991. 'Syntactic properties of sentential negation. A comparative study of Romance languages'. Unpublished PhD dissertation, University of Pennsylvania.

Zobl, H. and J. Liceras. 1994. 'Functional categories and acquisition orders'. *Language Learning* 44: 159–180.

Zuckerman, S. 2001. 'The acquisition of 'optional' movement'. Unpublished PhD dissertation, University of Groningen.