First Steps in Acquiring Persian Morphology

Sarvenaz Moradi

Doctor of Philosophy

University of York

Language and Linguistic Science

September 2021

Abstract

Very few studies have looked at Persian morphology in children and none have so far dealt with the earliest production of inflectional morphemes, which occur as suffixes on nouns and both prefixes and suffixes on verbs. This study is based on four children whose data were recorded longitudinally in Iran on a biweekly basis, beginning when they were reportedly still in the single-word period. We address the following questions: What are the first steps of morphological development in Persian? Which morphemes do children pick up at the very start of morphology? And what other aspects of linguistic development relate to the early morphological development? Do children produce inflected forms before the onset of two-word stage or after?

For this purpose, we analyse extent of use of all the inflectional morphemes that occur in one or more of the four children's first forms. We identify and quantify these morphemes along with their first point of appearance and contrastive use. We test the relationship between each of three factors that may relate to morphological development: (1) phonological development (2) lexical development (3) syntactic development.

The results of this study show that a correlation between early morphological development and syntactic development exist while such correlation was not seen between morphological development and phonological and between morphological development and lexical development. The results of this study are consistent with the constructivist model of Pre-/Proto-morphology, in which it is believed that children start with rote-learnt inflected forms in the pre-morphology period and start using these inflectional morphemes contrastively which results in their productive use in the proto-morphology period.

Table of contents

List of Tables

List of Figures

To

Melica Rastegari,

Dr. Hamed Abdi,

and

Dr. Sonja Eissenbeiß

To Melica; for making me fall in love with learning a foreign language (English) when I was only thirteen.

To Hamed; for bringing out in me the joy of listening closely to the way people (especially children) speak.

And...

To Sonja; for making me excited about learning a whole new bunch of languages...computer programming ones.

After all, it was always about LANGUAGE!

Acknowledgments

It has been a very challenging four and a half years for me, my family, my country and more recently for the whole world. Doing a PhD in the middle of every crazy thing going on around me was not easy at all. But I was privileged enough to receive all kinds of help and support from the amazing people around me.

Firstly, I would like to say how grateful I am to have had Professor Marilyn Vihman as my supervisor. Without her persistent help, the goal of this research would not have been realized. She is and always will be an inspiration. I would like to also thank Dr. Tamar Keren-Portnoy for her valuable comments on this research.

The Humanities Research Centre, located within the beautiful Berrick Saul Building has been my second home for the last four years. So, I would like to acknowledge Professor Richard Ogden, the HRC director and Helen Jacobs, the HRC administrator for doing their best in providing us an amazing place to carry out research both pre- and post-pandemic.

I am indebted to Professor Matt Matravers, head of Goodricke College, and Dr. Pürnur Altay, Graduate Students' Association President for their successful attempts in making my PhD journey easier. I would also like to extend my sincere gratitude to Janet Malcolmson and Dr. Jonathan Fisk for their guidance and to Huma Anwaar, Dr. Serge Sagna, and Dr. Parastou Marvasti for sharing their knowledge with me.

I greatly appreciate my friends for their help, from moral support to finding participants to having writing sessions with me: Mohammadreza Hosseinpour, Neshat Naserinejad, Shadi Shirazi, Anna Zhenova, Pori Saikia, Lenia Galata, Sanaz Ramezani, Sahar Farzaneh, Sadaf Sohofi, Shahrzad Akbaroff, Hadi Janfada, Mohammad Fereidooni, Nastaran Afshar, Mahzad Mirmoradi, Vida Abbasi, Shaghayegh Raeesi, Mahdieh Torabi, and Sheila Tadayon.

I would like to say thank you from the bottom of my heart to the participating families who were kind and welcoming enough to allow me (with all my equipment) to be a part of such precious moments interacting with their lovely children.

Finally, I would like to thank my beloved family, without whom I would not have been able to complete this research:

My mum Mehrzad and my dad Hassan, for their constant love and support, and for making this dream of mine possible.

My amazing grandmother, Maman Manijeh, for being my biggest ally in life and for always defending me and my choices.

My brother, Sahand, for being a constant advocate in my life, for always being there for me, and for always brainstorming with me on how to overcome obstacles.

And of course,

My partner, my missing piece and home, Richard Harris, for his love, support, and understanding. For always holding my hand when life gets tough, for always believing in me, and for putting my needs and my happiness before his.

Also, a big thank you to my favourite kiddo, Nazanin, for being the amazing little person that she is and for allowing me to use her drawing for my online participant recruitment poster.

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Were there not in the world the means to speak,
How could the meaning bud bloom so sleek?!

Attar of Nishapur
Translated by Hamed Abdi, Satah

1. Introduction

'How do children acquire language?' is a question that linguists have been trying to answer for the last 100 years. It has been addressed from different perspectives, from infants' perception of sounds, words, patterns, meaning, etc to their production, appearance, and use. Yet the question remains, 'Where does the child's linguistic knowledge come from?'. Accounting for the developmental source of linguistic knowledge has been a central concern of the study of child language (Vihman et al., 2009). Theoretical approaches try to explain whether the child's ability to acquire language is innate, or whether some sort of cognitive ability is at play. Researchers often focus on acquisition of one aspect of linguistic sub-fields, for example sounds, words, phrases, or sentences. There has also been research done on the relation between pre-linguistic and linguistic domains of a child's development, for example how the child's babbling practice plays a role in the emergent phonological system. A wide range of studies have been conducted, from naturalistic data collection to designing controlled experiments to test a specific hypothesis that looks closely at one of the aspects of a linguistic sub-field.

Among the different subfields of linguistics, morphology is one that runs into other subfields. But what is morphology? Morphology is 'the grammar of words' as Audring and Masini (2018) put it, or 'the internal structure of words' (Carstairs-McCarthy, 2005). According to Audring and Masini (2018) morphology entails both the form and structure of words, but it also includes their meanings, their relations, and the way new words can be formed. While derivational morphology is concerned with word formation processes that create related words (Seidenberg & Gonnerman, 2000) inflectional morphology is concerned with ways in which morphemes are used to mark words for certain grammatical features (Ambridge & Lieven, 2011).

Going back to our 'How do children acquire language?' question, typically developing children slowly obtain some cues about the input language patterning at different levels (i.e., segments, syllables, accentual patterns, words, phrases, and clauses) (Vihman et al., 2009). While beginning to make sense of the speech around them, children pay attention to the relationships between the above-mentioned levels of language, for example between words, or parts of words (Allen & Behrens, 2019). At around their first birthday they produce their first

words. Around their second birthday, they produce their first word combinations by putting those words together. Inflected forms in some languages follow a few months after the first words, or at the same time as the child's first word combinations emerge.

Theoretical background

Theories of language acquisition

Two of the main approaches to language acquisition studies are: generative, and usage-based. Generative approach began with Chomsky (1965)'s generative grammar, arguing that children have an innate language endowment that enables them to learn language. The usage-based approach, however, denies assuming an innate language endowment and instead proposes that children are equipped with cognitive skills that enable them to acquire language. A usage-based model assumes that communication is the main motivation for language learning (Behrens, 2020), i.e., 'language use'. The usage-based approach holds far more diversity compared to generative approach (Newmeyer, 2021). According to Newmeyer, despite this diversity of opinion within usage-based approach there lies a common idea that the connection between language knowledge and language use is far more intimate than recognised by generative approach.

Piagetian approach to language acquisition is the predecessor to usage-based approach. As Behrens (2009: p 384) puts it "the usage-based approach to language acquisition relies on insights from cognitive linguistics, a nonmodular theory that assumes that linguistic structure is tied to the semantics and pragmatics it encodes". Cognitive development precedes linguistic development. It is argued that knowledge comes from experience and general learning principles. This is done through communication and social interactions (Allen & Behrens, 2019). The child uses cognitive processes of learning and categorization, for example different types of pattern finding (Tomasello, 2003). The child gradually moves from item-specific to abstract syntactic knowledge when analysing patterns (Allen & Behrens, 2019).

Theories of morphological acquisition

Children's ability to utilize inflectional morphology systems in a productive manner is a subject of significant theoretical interest (Plunkett & Marchman, 1993). According to Engelmann et al. (2019) inflectional morphology plays a central role in language acquisition research for two main reasons. Firstly, inflectional systems are frequently highly complex. This

makes their acquisition one of the most significant challenges for learners. Secondly, inflectional systems have often been used as a benchmark for evaluating various theories of language acquisition due to their complexity and the insights they can provide into the language acquisition process more broadly (Engelmann et al., 2019). Inflectional morphology has been a particularly useful domain for evaluating the debate between generativist and constructivist approaches to language acquisition due to the insights it can provide into the language acquisition process more broadly (Räsänen et al., 2016).

The earliest studies in the modern era of morphological acquisition goes back to the late 1950s (Allen & Behrens, 2019). These were the longitudinal case studies like Bloom (1970) and Brown (1973). These first studies were conducted based on data collected from monolingual children speaking European languages (Guijarro-Fuentes et al., 2008). More studies have been conducted on a greater variety of languages since. In the last 40 years, research on the acquisition of inflectional morphology in languages from various regions has been increasing. This has provided insight into the role that language-specific factors play in this process (Penke, 2012).

Generative approaches predict that inflectional morphemes are productive as soon as they are acquired, while constructive approaches assume that the early inflected forms are unanalysed and rote-learnt by the child (Ambridge & Lieven, 2011). Generative approaches predict that children will not make agreement errors, provided they have learnt the required morpheme while constructive approaches predict that in the absence of these rote-learnt forms, children make omission and commission errors (Ambridge & Lieven, 2011).

One constructivist approach to morphological acquisition is developed by Dressler (1997). In this approach early morphological development divides into three phases: 1. Pre-morphology 2. Proto-morphology 3. Morphology proper. The pre-/proto-morphological approach posits that children first learn basic forms, like the nominative singular, through rote learning in the pre-morphology stage before progressing to acquiring multiple forms of the same lemma (i.e., mini paradigms) in the proto-morphology stage (Savičiūtė & Ambridge, 2018).

During pre-morphology, word forms are rote-learnt and usually in base form. A limited number of lexically stored inflectional forms can be found in the child's speech. These words are basic, memorised language forms in short utterances (Laalo, 2009). The word type/lemma

ratio is usually one. In transition to the next phase, the number of word forms gradually exceeds the number of lemmas (Stephany & Voeikova, 2009).

In proto-morphology, the child generalises the rote-learnt forms. This generalisation leads to the detection of morphological principles of decomposition of both form and meaning word-internally. The child also starts to construct morphological patterns by analogy. As a result, grammatical oppositions emerge and they begin to create new words based on the morphological structure of words they already know (Laalo, 2009). These develop into mini-paradigms. Bittner et al. (2003: xvi) define a true mini-paradigm as a "non-isolated set of minimally three phonologically unambiguous and distinct inflectional forms of the same lemma produced spontaneously in contrasting syntactic or situative context in the same month of recordings". The emergence of these mini-paradigms can be taken as evidence of morphological productivity.

Upon the beginning of morphology proper, the child acquires the basic language-specific properties of the adult morphology. A noticeable increase in the productivity of morphological combinations, as well as the emergence of more frequent overgeneralisation errors marks the beginning of this stage.

Constructivist accounts of language development focus on the importance of word form frequency, while many generativist accounts take into account the concept of a default morphosyntactic structure (Savičiūtė & Ambridge, 2018). Constructivist accounts of language development only propose that children use storage, analogy, and competition to learn language. The pre-/proto-morphological approach, on the other hand, also suggests that children develop symbolic rules as they learn language (Savičiūtė & Ambridge, 2018). The pre-/proto-morphological approach assumes that multiple rules are used in language learning. However, this approach also assumes that these rules are developed through analogy, and that this process is influenced by the frequency of different surface forms (Granlund et al., 2019).

One important current debate is whether children acquire inflectional morphemes by forming on-the-fly analogies across multiple stored exemplars or by stored abstracting rules (Ambridge, 2020; Granlund et al., 2019). The exemplar theory originated as a model for explaining how similarity and classification work in the perception process (Pierrehumbert, 2001). Exemplar models attempt to capture the way that humans retain detailed memories of

language-based events through the use of exemplars, which are stored and used for comparison with new input (Bod & Cochran, 2007). The basic concept behind this idea is that all prior language experiences are recorded as exemplars and that linguistic behaviour is determined by the stored exemplars that are most similar to the current situation (Bod & Cochran, 2007). According to Pierrehumbert (2001) in the exemplar model, categories are represented in the memory by a large group of exemplars from that category, which are stored in a cognitive map. The ones that are most similar to each other are stored close together, while those that are less similar are stored further apart. This means that when a new token is encountered, it is classified based on how similar it is to the other stored tokens. According to Rytting (2002) the exemplar model suggests that speakers do not use a fixed set of rules to modify words, like attaching a suffix. Instead, speakers rely on their memory of specific exemplars of words with the suffix already attached to guide their language production. This approach assumes that speakers use these stored exemplars rather than rules to produce language. According to Ambridge (2020) based on this approach, learners store specific examples of language, including the words used, the meaning of the words, and the context in which they were used. When producing or understanding new utterance, learners draw on these stored exemplars and use them as a basis for creating new utterance based on similarities to the stored exemplars. This process of using stored exemplars to create new utterance happens in real time. What this means is that "... at all levels, novel combinations can be generated by analogy as soon as the learner has stored, in principle, a single relevant exemplar" (Ambridge, 2020: 513-514). According to Ambridge (2020) an exemplar account gives a unified explanation for how language is acquired across all domains and it also offers an explanation for the correlation observed between different domains.

Children's early errors

According to Räsänen et al. (2016) generativist accounts take the early error-free performance observed in language acquisition as evidence for innate abstract knowledge of inflection. In contrast, constructivist accounts predict not only that children will make errors, but also that the pattern of correct and incorrect use of inflections will closely mirror the input to which the child has been exposed.

Many generativist accounts propose that children's errors occur when they resort to an unmarked form as a substitute when the desired form is not available. In contrast, constructivist accounts suggest that children's errors occur when they default to the

inflectional form of the target word that they have been exposed to most frequently (Savičiūtė & Ambridge, 2018).

It is anticipated that a U-shaped learning trajectory will be observed, which is a common error pattern in development featuring an increasing rate of overgeneralizations that are eventually corrected. Children are able to effectively use morphological rules in their speech, following this U-shaped pattern of development. Initially, irregular forms are accurately reproduced through memorization. However, as the children over-apply rules, the irregular forms become regularized. Eventually, the children are able to accurately produce the irregular forms again. This has been seen in other languages (Gervain, 2022).

Recovery from errors

Acquiring a language that primarily follows regular patterns can still be difficult, particularly when multiple forms correspond to the same function. This leads to a challenging and error-prone acquisition process (Nakipoğlu et al., 2022). In a rule-based approach recovery from errors can occur through a process known as 'blocking', which posits that the gradual learning of the correct irregular form will prevent the use of an overregularized form over time (Nakipoğlu et al., 2022). However in an analogy-based approach the competition between correct and overregularized forms in a child's memory is determined by their strength, with repeated encounters with the correct form strengthening its representation while the incorrect form is eliminated over time (Nakipoğlu et al., 2022).

Morphology at the centre of linguistics

As Spencer and Zwicky (2001: p 1) point out "morphology is at the conceptual centre of linguistics". This is due to the nature of morphological studies. Morphology is the study of word structure and since words are at the interface between phonology, syntax, and semantics, it is no surprise to see morphology at the centre of linguistic studies. Inflection creates grammatical forms of words, and is a part of morphology, but it also has an impact on the syntax of a sentence by affecting other constituents in the construction (Penke, 2012). Gervain (2022) notes that the relationship between grammar and vocabulary is established through morphology, and the process of acquiring the language involves the simultaneous development of both elements. As well as affecting sentence structure, the choice between different inflectional allomorphs might be phonologically determined, and the resulting word

form must comply with the phonological rules of the language (Penke, 2012).

Different factors play a role in specifying the way in which morphemes are ordered. According to Deen (2009) these factors include properties of the sound system, properties of the morphemes themselves and properties of the syntax. In the course of a child's linguistic development, the first linguistic domain that develops is phonology, followed by morphological development which requires decomposition of the sound sequences into units from utterances (Bittner et al., 2003). According to Bittner et al. (2003) for morphology to be developed, all of these units (words, syllables, and phonemes) along with a sufficient number of their representatives and properties need to be categorised.

Morphological acquisition studies

Morphological development seems to be undissociable from a quantitative enrichment of lexical, syntactic, and morphological structures (Bittner et al., 2003). Overcoming the single-word stage seems to be an essential step in morphological development (Bittner et al., 2003). According to Bittner et al. (2003) even in morphologically rich languages some aspects of syntax need to be enriched before (or at the same time) qualitative changes in verbal inflection can be seen. This is while, in Turkish, a highly inflected language, in which inflectional morphology is very regular (Aksu-Koç & Ketrez, 2011), inflectional morphology appears quite early (Aksu-Koç & Slobin, 1985). Aksu-Koç & Ketrez (2011) reported the emergence of inflection in a Turkish-speaking child between 1;3 and 2;0. Their findings provide evidence that in Turkish inflected single words are produced before the onset of the two-word stage.

Thordardottir et al. (2002)'s research on 45 English-speaking and 51 Icelandic-speaking children, with age range of 22 to 26 months, showed a strong and systematic relationship between lexical development and grammatical development in English and Icelandic. Their result showed the emergence of inflectional morphology was significantly correlated to the children's vocabulary size. Icelandic is a highly inflected language. They also observed that the Icelandic-speaking children started producing regular inflectional morphology sooner with a smaller vocabulary size compared to the English-speaking children. Stolt et al. (2009) investigated the relation between the emergence of grammar and the lexical growth in 181 Finnish children at 2;0. They reported a strong relation between the two. However, the acquisition of nominal and verbal inflections showed different lexical dependency. According to their findings, case form types occurred when the nominal lexicon size was between 50 and

250. Verb inflection types occurred from the very beginning of the verb lexicon acquisition.

Mariscal & Gallego (2012) focused on the relationship between lexical development and grammatical complexity (including noun and verb morphology). They analysed the data from 593 Spanish-speaking children with an age range of 1;4 and 2;5. They found a strong correlation between the two.

Vihman et al. (2013) carried out intensive analyses of phonological, lexical, morphological, and syntactical development of late talkers at the end of the single-word period. Late talkers refers to children with an unusually small productive vocabulary size (for their age), in the absence of any other known neurological, sensory, or cognitive deficit. They demonstrated that a combination of high age and small phonological development can be a strong predicator of low accuracy in consonant use and relatively poor lexicon, morphology, and syntax later. Their study concluded that phonetic and phonological knowledge is a key foundation for later linguistic development.

Morphological studies in Persian

Research on Persian child language is not as extensive as studies on English child language (Kazemi et al., 2015). The majority of studies conducted within this country tend to focus on a clinical perspective. A limited number of research studies have been conducted to investigate the morphosyntactic aspects of acquiring the Persian language (Haresabadi et al., 2018). Even fewer studies have focused especially on the earliest stages of the morphological development in Persian. Pouladi and Khoddam (2003) compared the mean length of utterance (MLU) in two groups of children: 48-54 months and 54-60 months. They reported mean MLU scores of 7.09 and 7.5 respectively. A longitudinal analysis was conducted by (Meshkato-Dini (2004)[1] on the emergence of inflectional affixes in two Persian children between the ages of 23 and 42 months (seen in Kazemi et al., 2015). Ghelmanipour (2006) looked at some morpho-syntactical features in 1.5 to 2.5-year-olds. She concluded that the mean MLU for 27–30-month-olds is 2.7. Oryandi-Zanjani et al. (2006) reported the mean length of utterance in words (MLUw) for 580 children between the ages of 2 to 5. Kazemi et al. (2012) reported the mean length of utterance in morphemes (MLUm) of 171 children between the ages of 2.5 and 5.5 years. They found a positive correlation between MLUm and age.

[1] Full text of the paper was not available for review.

Samadi (1997) studied the early stages of morphological acquisition on Persian in three children between the ages of 1;8 and 3;4. Samadi developed an adaptation of the LARSP (Language Assessment, Remediation and Screening Procedure) profile for Persian: P-LARSP (also see Samadi & Perkins, 2012, 1998). In her 1997 PhD thesis, Samadi showed that MLUm could be used as a reliable and effective measure to estimate the level of language development. In Samadi (1997)'s study, MLUm was proved to be reliable up to MLUm 4 in Persian. Samadi concluded that MLUm can be used as a reliable measure of language development at the early stages while it loses its validity at the later stages of development.

Jalilevand et al. (2012)'s longitudinal study was based on two children from the age of 12 to 60 months. They concluded that Persian-speaking children start showing evidence of using grammatical morphemes before their MLU reaches 2. They documented a fast increase in the children's MLUm from the age of 24 to 29 months. The MLUm was observed to have a slower increase from the age of 42 to 60 months.

A longitudinal study by Marvasti (2014) focused on the order of acquisition of verbal morphology in Persian. She analysed the data collected from three monolingual children with an age range of 1;8 to 3;11. Marvasti (2014) determined the point of acquisition of Persian verbal inflections based on the productivity and contrastive use of inflections by Pizzuto & Caselli (1994) which she adjusted for Persian. Marvasti (2014) concluded that the acquisition of verbal morphology in Persian is of a gradual nature. She also reported the influence of input in terms of both type and token frequency of verbal morphemes on their order of development. She reported that the typological factors' role (i.e., transparency and perceptual salience) were not as clear as the frequency of the morphemes in the input. Marvasti's findings confirmed the interdependence of lexical development and morpho-syntactical development. However her research found that while the rate at which children use different forms of speech increases as they expand their vocabulary, this does not support 'critical mass hypothesis', as proposed by Marchman and Bates (1994). In other words, as children's language skills develop, they are able to use their grammatical inflections with a wider variety of verbs, but this is not linked to a specific increase in their vocabulary size.

Jalilevand et al. (2016) developed Persian developmental sentencing score (PDSS), as a clinical measurement tool, based on Lee (1974)'s developmental sentencing score (DSS) for Persian.

They collected language samples of 115 children during free play and picture description. The ages of the children were between 2;5 and 5;4. They reported a significant correlation between PDSS and MLUm, suggesting that PDSS could be used as a reliable numerical tool for assessing Persian-acquiring's morpho-syntactical development. They also demonstrated a significant correlation between age and grammatical sub-categories. Among the different sub-categories, the verb morphology, grammatical morpheme, sentence structure, and preposition and conjunction showed the highest correlation with age. Their study also revealed a positive correlation between the overall PDSS score and the following sub-categories' scores: verb morphology, grammatical morpheme, sentence structure, and preposition and conjunction. Question words, however, demonstrated no correlation with the overall PDSS score, and only a poor correlation with age. PDSS will be explained in more detail in chapter 3.

A longitudinal study was conducted by Zarei Mahmood Abadi et al. (2021) to examine the development of MLU and the emergence of simple compound sentences in a child, from 15 to 34 months of age. The study revealed that the child in question produced her first simple two-word sentence at 21 months and 13 days and a four-word compound sentence at 26 months and 29 days. The emergence of compound sentences led to a significant acceleration in the increase of the child's MLU. The researchers observed a frequent addition of a unit to the length of the sentence (word) within a short period of time. A strong correlation was found between the child's MLU, age, mean sentence length, and number of words in the Persian language, as well as the total number of words and age. The results of the study indicate that the appearance of compound sentences in the child's output also has a significant impact on increasing the MLU.

Research objectives, questions, and corresponding hypotheses

The aim of this study is to investigate the morphological development of Persian qualitatively and its relation with phonological development, lexical development, and syntactic development of the child, quantitatively.

This thesis has two key aims. Firstly, this thesis will attempt to provide a descriptive overview of the earliest inflectional morphology development in children. To be able to depict a clear picture of what child's linguistic development look like, sub-domains of lexical, phonological, syntactical, and morphological will be looked at as independent phenomena. Secondly, it will assess the relation between these sub-domains (i.e., phonological development, lexical

development, and syntactic development) and morphological development; specifically, during the end of the single word period and into the two-word period (see **Error! Reference source not found.**) across the age period we expect to see the first signs of morphology's emergence in the child's language. A thorough statistical analysis will be employed to achieve this. Therefore, three research questions are developed from the second research objective of this thesis. The three research questions are as follows:

1. Can the lexicon size of children influence their morphological acquisition?
2. Can the phonological abilities of children impact their morphological acquisition?
3. Can the child's syntactical development be used as a predictor for their morphological acquisition?

Based on the above questions the following hypotheses are formed:

1. There is a relation between lexical development and morphological development
2. There is a relation between phonological development and morphological development
3. There is a relation between syntactic development and morphological development

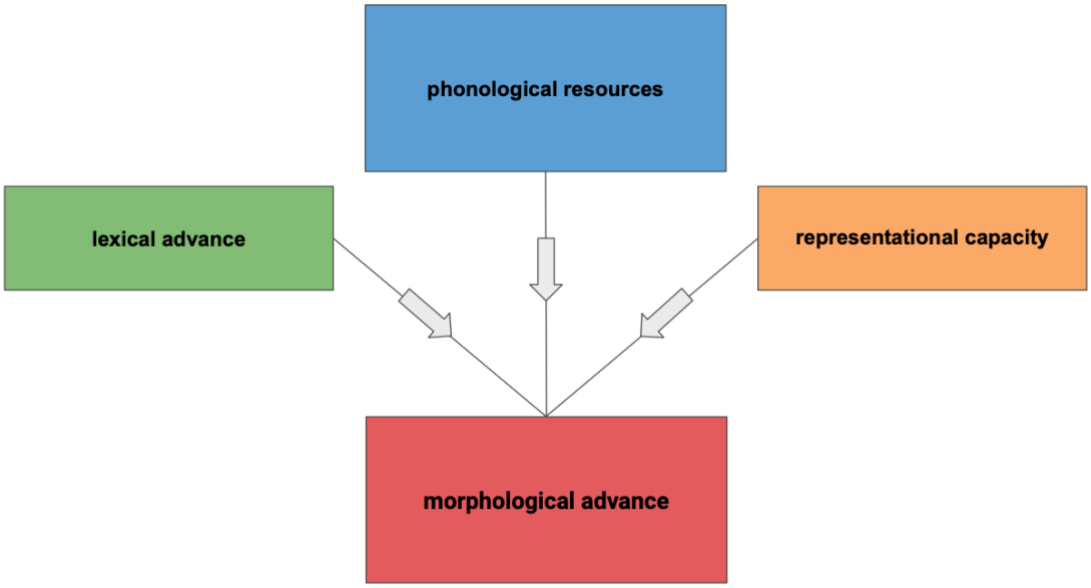

Figure 1 Model

To be able to investigave the above-mentioned hypotheses, naturlaistic data was collected from children who has not shown any signs of using inflectional morphomes and/or word combinations according to parental reports. The data was transcribed and coded. Each child's linguistic development was measures and described thoroughly using different measures. A statistical analysis was then used to asses the relationship between the children's lexical, phonological, syntacic and their morphological development.

In order to examine the relationships between these domains, a selection of various measures were employed. The research questions will then be addressed through the use of statistical analysis.

Learning trajectory of a morphologically rich language

The morphological structure of Persian is mostly of an agglutinating one, meaning that it involves the use of multiple affixes, and lies in between highly inflected languages like Arabic and Czech, and less inflected languages like English (Bijankhan et al., 2011). In agglutinating languages, grammatical relationships are shown through a collection of prefixes and suffixes. Usually, each grammatical function is represented by a single morpheme (Ladányi et al., 2020). In these languages morphology connects the grammar and vocabulary and acquisition necessarily proceeds in parallel in the two domains (Gervain, 2022). Gervain (2022) further suggests that syntax is often indicated through the use of suffixes therefore, the ability to recognize these suffixes is an important factor in learning the syntax of the language.

It has been suggested by some researchers that the complexity of the morphology in agglutinating languages is not the primary determinant of ease of acquisition. Rather, it is believed by these scholars that the richness, regularity, predictability, and transparency of the morphological system are the key factors that play a role in facilitating the acquisition of morphology in these languages. In this section, we will briefly mention the ideas of researchers who believe that the richness, regularity, predictability, and transparency of the morphological system plays a key role in facilitating the acquisition of morphology in agglutinating languages. we will not go into much detail about these theories, but rather provide a general overview of the different perspectives on this topic.

Kelly et al. (2014) argues that the complexity of the morphological structure of a language

does not necessarily mean that it will be more difficult to learn. There are some (e.g., Demuth, 1992; Xanthos et al., 2012) research that indicates that languages with more morphological forms may actually be easier to acquire. The fewer variations and irregularities present in a system, the more manageable it is to learn a language within that system (Nakipoğlu et al., 2022). According to Kelly et al. (2014) this suggests that regularity, rather than complexity, may be the key factor in determining how easily a language's morphology can be learned. This implies that having more morphological complexity does not necessarily make it harder for a child to acquire that language. In languages with rich morphology, it serves more purposes and there are more connections between forms and meanings making it more informative (Wijnen et al., 2001) than in languages with less morphology (Dressler, 2003). Children learn about the importance of morphology in the language they are acquiring, and they are more attuned to it if the language they are learning has a lot of morphology (Dressler, 2003). One example is Turkish. Turkish is an agglutinating language (Oflazer, 1994; Yalçin, 1996 among others). The regularity of conjugation in Turkish (Aksu-Koç & Ketrez, 2011) facilitates the learning process (Yalçin, 1996). In their paper, Kelly et al. (2014) argue that the acquisition of inflectional morphology in Turkish is more beneficial for children than in English. Kelly et al. (2014) cites the much greater usefulness of acquiring this aspect of the language and the significant difference in orderly variation available in the respective inputs as the reasons for this. Kelly et al. (2014) also note that for children, the usefulness of acquiring inflectional morphology is more important than the superficial simplicity of a language. Another notable example that can be mentioned is the Finnish language. Finnish-speaking children begin using case forms early on in their language development (Laalo, 2009). According to Laalo (2009) mastering the relationship between stem forms and suffixes in the inflectional system is a major challenge in learning Finnish due to the large number of inflectional classes and the rich allomorphy of suffixes.

One point of contention is the role that regularity and number of inflections play in shaping the acquisition process. It seems that children's ability to acquire morphological forms is influenced more by the predictability of the forms, rather than the sheer number of inflections present in a language.

Ackerman & Malouf (2013) proposed that there is a limit to the level of irregularity in an inflectional system. Cotterell et al. (2019) later refined this hypothesis by suggesting that systems that have a large number of forms per paradigm have even stricter limits on the level

of irregularity per distinct form. The two dimensions, form number and irregularity, have an interaction between them, meaning that a system cannot have high complexity in both at the same time. To put it simply they state that if a language requires its speakers to use many different forms, those forms must be relatively predictable.

This point of view aligns with the idea presented by Cotterell et al. (2019). They propose that morphological systems can be characterized by their degree of regularity or unpredictability, but not both. From this perspective, it is suggested that a language with a high degree of regularity, such as Finnish or Turkish, may facilitate the acquisition of morphology in children, as the forms are more predictable and consistent. This is in contrast to a language with a high degree of unpredictability, where the forms are more irregular and difficult to predict. This highlights the importance of regularity in shaping children's ability to acquire morphological forms.

It is crucial to note that inflectional morphology, both for verbs and nouns, can vary considerably among languages. The results presented in Dressler (2003) demonstrate the necessity for distinguishing various morphological systems within each language, as previously proposed. The complexity of verb morphology in Persian is significantly greater than that of noun morphology. This leads to the expectation that children's detection of verb morphology will occur earlier, as the abundance of productive patterns in verb morphology is significantly more extensive.

Moreover based on the pre-/proto-morphological approach, it is predicted that during the one-word stage (i.e., pre-morphological stage), there will be an observed presence of rote-learned non-productive single words with inflectional morphemes. Furthermore, during the proto-morphological stage, a productive use of inflectional morphemes is expected to be evident. Given that Persian is a pro-drop language, and shares similarities with languages such as Turkish in terms of its complex yet regular morphology, it is anticipated that the appearance of inflected single words shortly before the onset of the two-word stage will be observed in the present study.

With regards to the emergence of initial word combinations children enter the two-word stage generally around their second birthday. Vihman (2022) states first word combinations are generally observed in the same session or within the first month following a 25-word session.

A 25-word session is when 25 or more spontaneous words are produced in one session. Considering the complexity of the morphological system in Persian we expect to see the first inflected forms shortly before or around the 25-word stage.

Research on children suggests that grammar development is closely linked to vocabulary size as children move from single words to sentences and ultimately gain mastery over the morphosyntactic structures of their native language (Bates & Goodman, 1997). The onset of word combinations in a child's language development is typically observed when they have a vocabulary size between 50 and 100 words and subsequently, a consistent correlation between lexical and grammatical development is observed (Bates et al., 1991). The outcome of Bates and Goodman (1997)'s study indicated that the most accurate prediction of grammatical abilities at 28 months of age is the size of a child's vocabulary at 20 months of age.

As mentioned earlier Marvasti (2014)'s results did not support 'critical mass hypothesis'. It is possible that the lack of support for the critical mass hypothesis put forth by Marchman and Bates (1994) in Marvasti (2014)'s study was due to the fact that the children included in her research had already surpassed the point at which this effect would be observed. By expanding the research to include the final stages of one-word utterances and the onset of word combinations, a period during which the first uses of inflectional morphology are expected to occur, it may be possible to more clearly discern the interdependence between lexical development and morphological development. Hence one benefit of this study is that it seeks to extend Marvasti (2014)'s previous research by investigating the acquisition of inflectional morphology in younger children during the earlier stages of development that is its onset.

Based on the existing literature, it is proposed that the acquisition of inflectional morphemes in Persian by language learners will exhibit a U-shaped developmental pattern. Specifically, it is expected that initial progress in acquiring inflectional morphemes will be rapid, followed by a period of relative stagnation, and ultimately culminating in a second period of accelerated acquisition. It is important to note, however, that individual variations in the timing and duration of each stage may occur. It is expected that the acquisition of inflectional morphemes will follow a predictable sequence, where simpler forms will be acquired before more complex forms. In particular, children are expected to first demonstrate proficiency in using inflectional forms that are more frequent and have higher number of allomorphs, before gradually progressing to the productive use of less frequent and complex inflectional forms. It is hence

predicted that, in the process of language acquisition, individuals will acquire simpler inflectional morphemes before those with strong allomorphs.

Contributions of this research

There is a significant amount of data on the production of language that suggests that young learners are able to easily master complex morphology. However, the initial stages of this learning process, specifically how young infants begin to perceive and analyse complex morphological forms, is not as well understood (Gervain, 2022). The purpose of this study is to gain a deeper understanding of the earliest stages of learning inflectional morphology. By investigating this topic, it aims to help fill the current gap in knowledge about the subject.

Moreover in the present study, an alternative approach to assessing morphological development was utilized. Specifically, the focus was on comparing each individual child's development to their own, rather than comparing groups of different ages as has been done in some previous studies (e.g., Pouladi & Khoddam, 2003). This approach allows for a more individualized and accurate tracking of development and enables the creation of a more detailed picture of morphological development. It should be noted that while this approach may not be entirely novel, it is a valuable addition to the existing methodologies used in this field of research. Furthermore, the present study extends the age range of children under examination by including the final stages of the single-word period. This allows for a more comprehensive understanding of morphological development during this crucial stage. Additionally, the study also utilizes a more advanced statistical model than those used in previous studies. This allows for a more robust and accurate analysis of the data, leading to a deeper understanding of morphological development in children.

It is worth noting that derivational morphology comes later in the course of a child's linguistic development. Since here the focus is on the earliest stages of morphology, it is expected not to be able to study derivational morphology. Derivational morphology simply is not present in the child's production at this stage.

The theoretical approach of this thesis is a constructivist model of Pre- and Proto-morphology in analysing the early stages of inflectional morphology acquisition in Persian.

Outline of thesis

This thesis has been organised in the following way: Chapter 1 (current chapter) gives a background for the study. Chapter 2 presents an overview of adult Persian morphology. Chapter 3 is concerned with the methodology employed for this study. Chapters 4 to 7 report the analyses of the children's data. Chapter 8 analyses the results overall. Chapter 9 concludes the thesis with a discussion, summary, and limitations of this study followed by introducing directions for future research.

2. Persian

Introduction

This chapter provides a brief description of Persian, its phonology, and morphology so that a clear picture of morphological development in children can be provided later.

Persian is the official language of Iran. It belongs to the Indo-Iranian branch of Eastern Indo-European languages. Until 1935 Iran was known to the rest of the world as Persia, which is why the language is widely known as Persian. 110 million people in the world use or understand Persian, out of which 65 million are native speakers (Windfuhr, 2009). 70 million of these people live in Iran and refer to the language as *Farsi*. 33 million live in Afghanistan and call it *Dari*. 7.2 million live in Tajikistan, where the language is officially known as *Tajiki*. All three varieties are mutually intelligible to a fairly high degree (Beeman, 2005). The Persian spoken in Iran is mostly spoken as a native language in the central and eastern parts of the country.

Persian is believed to have two registers, one being the formal or written register and the other the colloquial register. The two varieties "are closely and systematically related but obey different rules and must be considered two separate systems" (Jasbi 2020, p. 135). Mahmoodi-Bakhtiari (2018b)'s study has established that Persian displays diglossic characteristics, with significant variations observed between its colloquial and written forms in terms of phonology, morphology, syntax, and semantics. Mahmoodi-Bakhtiari (2018b)'s research illustrates that the distinctions between the two registers of the Persian language are substantial, with colloquial Persian possessing specific morphemes that are unique to it. Additionally, the study demonstrates that personal enclitics and verb endings exhibit slight variations between colloquial and written Persian. Furthermore, the investigation reveals that personal enclitics can be added to prepositions in colloquial Persian, and certain words can only receive the plural suffix in colloquial forms of the language. This highlights the complexity of the Persian language and its variations between colloquial and written forms. Children acquire colloquial Persian as they are growing up and only learn formal Persian in school (Jasbi, 2020). The children in this study were exposed to colloquial Tehrani Persian, spoken in Tehran, Iran. Whenever in doubt about a specific target word form produced by a participant, the caregiver's speech was used as reference.

Persian is a pro-drop language or null subject language. In pro-drop languages, the pronoun subject can be optionally omitted. Persian is taken to have an SOV word order underlyingly. However, Colloquial Persian involves many rearrangements (Darzi & Boroujerdi, 2015) that imply a free word order (Karimi, 1994). Persian word order in declaratives is (S)(O)(PP)VI[2]. This means all the elements except for the verb are optional. Hence a verb on its own can constitute a sentence in Persian (Mahootian, 1997), as can be seen (2.1):

(2.1) goft-am
said-1S
'I said'

Persian phonology

Consonants and vowels

Persian has a total of 29 phonemes: six vowels and 23 consonants (see Figure 2 and Table 1). The vowels are traditionally divided into two categories of short /a, e, o/ and long /ɑ, u, i/, although vowel length is not considered phonologically significant in Persian (Mahootian, 1997; Samareh, 1977) .

[2] S = subject
O = object
PP = preposition phrase
V = verb
I = inflection

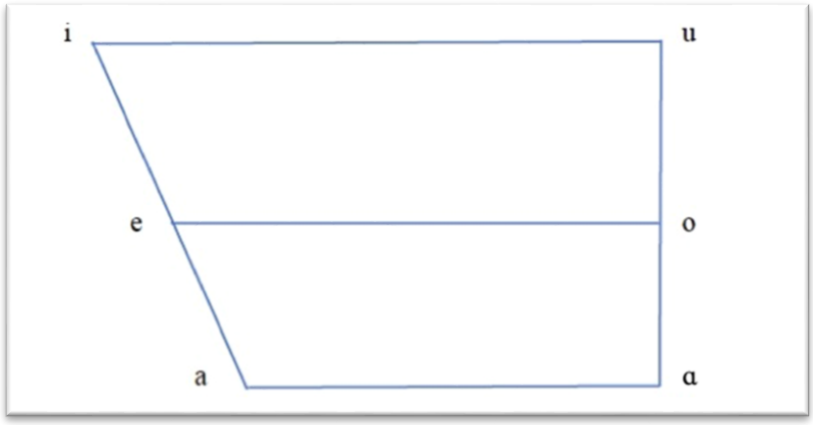

Figure 2 Persian vowels

Table 1 Persian consonants

	labio-labial/ labio-dental	apico-alevolar/ apico-dental	dorso-post-alveolar	dorso-palatal	dorso-prevelar	dorso-post-velar	dorso-uvular	glottal
stop	p b	t d			k g		ɢ	ʔ
affricate			ʧ ʤ					
fricative	f v	s z	ʃ ʒ			x		h
nasal	m	n						
liquid		l,r						
glide				j				

voiceless on left, voiced on right

(adapted from Mahootian 1997)

Persian syllable structure does not allow two vowels in one syllable. In other words, no word can start with a vowel on the surface structure in Persian (Dehghan & Kambuziya, 2012). According to Sadeghi (2002), insertion of a glottal stop at the beginning of a word that starts with a vowel is necessary to resolve hiatus. It is worth mentioning that not all linguists agree with this (see Navab Safavi, Fallahi, and Ghadimi Fomani 2020). One of the consequences of this alternative view is the assumption that Persian has six syllable types (V, VC, VCC, CV, CVC, CVCC) instead of the commonly accepted three (i.e., CV, CVC, and CVCC). There are also acoustic studies demonstrating that glottal stop may have other allomorphs (e.g. creaky voice or completely irregular vibration of vocal cords: Yazarlou, 2014), allowing the possible

omission of the glottal stop /ʔ/ and potentially allowing two adjacent vowels under certain circumstances (e.g. in rapid and continuous speech) (Staji et al., 2011).

In any sequence the syllable boundary is any consonant that immediately precedes the vowel. For example, the word *dast-am* 'hand-POSS.1SPC[3] (my hand)' is composed of the two syllables: *das* and *tam*. The most frequent syllable type is believed to be CVC at the word level, and the most frequent type is CV+CVC (Mahootian, 1997).

Persian does not allow initial consonant clusters. Consonant clusters are restricted to word-medial and final positions (e.g. *marg* 'death', *naɢʃ* 'role'). Consonant clusters can only consist of two consonants, or three consonants if across syllable boundaries.

Stress

Stress in Persian is predictable and generally non-phonemic. In nouns and noun phrases, it falls on the final syllable. Most Persian affixes change the stress in the word (Ghomeshi 2003, among others). In verbs, the stress falls either on the initial or the non-final syllable. The only exception is the verbs with no affixes, in which the word's stress falls on the final syllable (Ferguson, 1957; Kahnemuyipour, 2003):

(2.1) xórd
ate.3S
'He/she/it ate'

As mentioned earlier affixation can change the stress in a word. Among the affixes, object marker /-o/, the indefinite marker /-i/, pronominal enclitics, verb agreement affixes, and EZAFE /-e/ do not change the stress. Examples of these are provided in (2.2) to (2.6), respectively.

(2.2) dár-o be-band
door-OM SBJV-shut.2S

[3] A single dash '-' is used throughtout this thesis when there is a one-to-one correspondance between the Persian morphemes and the English glosses. Whereas, a single period '.' is used when there is not a one-to-one correspondance between the two (i.e., the parts of English glosses comprise more than two pieces).

'Shut the door'

(2.3) túp-i
ball-INDEF
'a ball'

(2.4) doxtár-am
daughter-POSS.1SPC
'my daughter'

(2.5) xórd-am
ate-1S
'I ate'

(2.6) daftár-e zard
notebook-EZ yellow
'yellow bag(s)'

The rest of the affixes change the stress in the word. Firstly, all the prefixes change the stress. The durative prefix /mi-/, the subjunctive/imperative prefix /be-/ and the negative prefix /na, ne-/, take the primary stress. If /na-, ne-/ and /mi-/ both appear as prefixes for a verb, /na-, ne-/ takes the stress. Examples of these are presented in (2.7) to (2.10).

(2.7) mí-xord
DUR-ate.3S
'He/she/it was eating'

(2.8) bó-xor-e
SBJV-eat-3S
'That he/she/it eats'

(2.9) ná-xor
IMPNEG-eat.2S

‘Don’t eat’

(2.10) né-mi-xor-am
NEG-DUR-eat-1S
‘I’m not eating’

Secondly, the definite /-e/ suffix (Ghomeshi, 2020) and the plural /-ɑ/ suffix (Ghomeshi, 2003) take the stress. This means that the stress remains on the final syllable of the word. Compare the pairs in (2.11) for the definite /-e/ and in (2.12) for the plural /-ɑ/:

(2.11) mɑʃín vs. mɑʃɪn-é
car car-DEF
‘car’ ‘the car’

(2.12) deráxt vs. deraxt-ɑ́
tree tree-PL
‘tree’ ‘trees’

Persian morphology

Languages are traditionally classified on the basis of their morphology. Languages can be agglutinating, synthetic or inflecting-fusional, or polysynthetic (Dressler, 2010). In agglutinating languages there is a one-to-one correspondence between the morpheme and the concept. The borders between morphemes are clear. In inflecting-fusional languages there is not a one-to-one correspondence between the morpheme and the concept, which means the morpheme borders are not clear. Finally, polysynthetic languages can combine a large number of morphemes into a single word without having clear morpheme borders.

Some argue that languages cannot be classified so ideally (Dressler, 2003; Dressler, 2010). According to Dressler (2003; 2010) some languages may be typologically distinct due to the different typological characters of their noun versus verb inflections. Persian language has been said to show primarily exhibits agglutinative characteristics, while also displaying subtle

fusional properties (Kalbasi, 1992; Samareh, 1990). The morphological typology of the Persian language has been a subject of debate among scholars in recent years, with varying opinions on its classification. This disagreement may be attributed to the fact that the language exhibits behaviours of multiple morphological types. It has been observed that Persian displays characteristics of both agglutinating and analytic languages, which may contribute to the difficulty in determining a clear morphological classification. Furthermore, the complexity of the language and the criteria used to evaluate its morphological characteristics may also play a role in the lack of agreement among scholars regarding its morphological typology. According to Majidi and Mirdehghan (2021), Persian exhibits characteristics of both an agglutinating language and an analytic language. When considering border transparency, Persian behaves as an agglutinating language. However, when considering the internal complexity of words, Persian exhibits analytic properties. While according to Kalbasi (2008) Persian appears to behave as an agglutinating type in noun morphology and inflecting-fusional in verb morphology (seen in Kazemi, 2013). This topic falls outside the parameters of the current study and will not be addressed further.

Persian morphemes are categorised into the following categories: lexical, functional, derivational, inflectional, and clitics. Lexical morphemes are the open class of simple words that can be used independently. These also include the main morpheme in a verb. These are the main word classes: nouns, verbs, adjectives, adverbs, and prepositions. Functional morphemes contain pronouns and conjunctions.[4] Derivational morphemes are involved in making new words. They outnumber inflectional morphemes in Persian. Inflectional morphemes are bound morphemes. They can appear as both prefixes and suffixes on verbs and only as suffixes on nouns. Finally, clitics are bound morphemes that cannot be used independently. However, unlike inflectional morphemes, they are not part of the word structure.

In this thesis, the focus will be on the last two categories (i.e., inflectional morphemes and clitics). Both will be referred to as 'inflectional' for the ease of analysis.

[4] In formal Persian, the /-ɑ/ variant of the object marker is also classified as a functional morpheme (see Meshkato-Dini, 2008).

Verbal morphology

Every verb has two stems: past and present. The verb conjugates in terms of three persons and two numbers. The past stem plus /-an/ form the infinitive. The regular verbs form their present stem by omitting the past stem ending /-id, -ɑd, -d/ (Mahmoodi-Bakhtiari, 2018a). Table 2 presents an example of each of these stem endings.

Table 2 Persian regular verbs

Present stem	Past stem	Infinitive	Gloss
bor	bor-id	bor-id-an	'to cut'
oft	oft-ɑd	oft-ɑd-an	'to fall'
kan	kan-d	kan-d-an	'to dig'

There are also irregular verbs, where there is no relationship between the two stems (see Table 3).

Table 3 Persian irregular verbs

Present stem	Past stem	Infinitive	Gloss
gu	goft	goft-an	'to say'
bin	did	did-an	'to see'
kon	kard	kard-an	'to do'

Verbs can be simple, prepositional, or compound. Compound verbs are made with a non-verbal part combined with a verb through either combination or incorporation processes (Dabir-Moghaddam, 1997). The non-verbal part can be a noun, an adjective, an adverbial phrase, a nominal phrase, or a prepositional phrase. The verbal part follows the non-verbal part.

(2.13) bɑz kardan (ADV + V)
open to-do
'to open'

(2.14) zamin xordan (N + V)
floor to-hit
'to fall'

(2.15) pejdɑ ʃodan (ADJ + V)
found to-become
'to be found'

(2.16) be donjɑ umadan (PP + V)
to world to-come
'to be born'

The inflectional structure of the Persian verb is as follows:

$$[(\text{NEG } /\text{na}-, \text{ne}-/)] + \begin{bmatrix} (\text{PRES } /\text{mi}-/) \\ (\text{SBJV/IMP } /\text{be}-, \text{bo}-/) \end{bmatrix} + [\text{verb}] + [(\text{PSPT } /-\text{e}/)] + [(\text{personal suffixes})] + [(\text{personal enclitics})]$$

(Eslami & Lemjiri, 2009)

It is worth noting that pragmatically NEG and SBJV/IMP cannot both appear (Eslami & Lemjiri, 2009).

Verbal prefixes

Verbal prefixes in Persian are as follows: NEG /na-, ne-/, DUR /mi-/, and IMP/SBJV /be-, bo-/. NEG /ne-/ occurs when followed by /mi-/. NEG and IMP/SBJV cannot co-occur. Attached to the past stem, /mi-/ forms past progressive, past habitual, and counterfactual conditional (Perry & Kaye, 2007). When attached to the present stem, it forms the general present, progressive present, habitual present, and future (Perry & Kaye, 2007). The imperative /be-, bo-/ prefix attaches to the present stem to form the imperative. The negative prefix /na-, ne-/ precedes the DUR /mi-/ prefix.

Personal suffixes

Personal suffixes mark subject-verb person and number agreement on the verb (Table 4).

Table 4 Personal suffixes (past and present)

person/number	singular	plural
first person	-am	-im
second person	-i	-in
third person:	-e (present tense)	-an
	∅ (past tense)	

Persian marks only two persons in imperatives: second person singular and plural (Table 5). The 2S appears in the base form. The 2P /-in/ is added to the base 2S to form the second person plural.

Table 5 Personal suffixes (imperative)

person/number	singular	plural
first person	-	-
second person	∅	-in
third person	-	-

All inflectional prefixes and suffixes attach to the verbal element of compound verbs, as in the following examples.

(2.17) dorost kard
fix did.3S
'(He/she/it) fixed (it)'

(2.18) dorost na-kard
fix NEG-did.3S
'(He/she/it) did not fix (it)'

(2.19) dast bezan
hand hit.2S
'Touch (it)!'

In some verbs, the IMP/SBJV prefix is often deleted in speech:

(2.20) harekat (be)kon
move (IMP)-do.2S
'Move!'

Copula 'to be'

The present tense of the verb *budan* 'to be' (Table 6) can be expressed through enclitics, which are attached to adjectives, pronouns, and nouns. These forms do not change the stress in the word.

Table 6 Copula *budan* 'to be'

person/number	singular	plural
first person	-am	-im
second person	-i	-in
third person	-e/-s	-an

Past participle suffix /-e/

In formal Persian the past participle suffix /-e/ followed by copula (Eslami, 2019), marks the present perfect form of the verb as in (2.21):

(2.21) did-e-am
saw-PSPT-be.1S
'I have seen (it)'

However, in colloquial Tehrani Persian PSPT /-e/ is omitted for 1S, 2S, 1P, 2P, and 3P. In 3S the whole inflected copula is omitted. There is also a shift in stress (Yousefi, 2018). This morphological stress contrast is the result of PSPT /-e/ being assimilated to the vowel of the

suffix or even disappearing altogether (Ferguson 1957). The stress, however, remains on the syllable (Table 7).

Table 7 Present perfect in colloquial Tehrani Persian

person/number	singular	plural
first person	did-ám	did-ím
	saw-PSPT.1S	saw-PSPT.1P
second person	did-í	did-ín
	saw-PSPT.2S	saw-PSPT.2P
third person	did-é	did-án
	saw-PSPT.3S	saw-PSPT.3P

This stress shift results in two very similar inflected types which differ only in terms of stress: /díd-am/ (simple past) vs. /did-ám/ (present perfect).

In the data collected for this study, only one present perfect inflected type (24 inflected tokens) was recorded. This inflected type is the present perfect 3S as in 2.11:

(2.22) CF[5]: xajid-e
AT[6]: xarid-e
bought-PSPT.(be.3S)

For ease of analysis this suffix will be marked as PSPT.3S.

Non-verbal morphology

The inflectional structure of the Persian noun is as follows:

[5] CF: Child Form
[6] AT: Adult Target

$$[\text{noun}] + [(\text{plural } /-\text{ɑ}/)] + \begin{bmatrix} (\text{indefinte } /-\text{i}/) \\ (\text{definite } /-\text{e}/) \\ (\text{relative clause suffix } /-\text{i}/) \\ (\text{personal suffix/pronominal clitics}) \\ (\text{Ezafe } /-\text{e}/) \end{bmatrix} + [(\text{copula to be clitic})]$$

(adapted from Eslami & Alizade, 2009)

The indefinite /-i/ and the relative clause suffix /-i/ have not been recorded in our data and so will not be discussed further. The definite /-e/ stands in complementary distribution to the plural suffix /-ɑ/ (Hincha, 1961, as cited in von Heusinger & Sadeghpoor, 2020).

Pronominal enclitics

Pronominal enclitics can attach to different word categories: noun, preposition, verb, adjective, question word, and demonstrative pronoun (Sorahi & Alinezhad, 2013). These enclitics are given in Table 8, followed by examples of the enclitics attached to the above-mentioned categories. In the verbal distributions, the enclitic can appear pre- or post-verbally. If they appear in the pre-verbal distribution, the verb is marked for third person singular (see examples below) (Rasekh-Mahand, 2011).

Table 8 Persian pronominal enclitics

person/number	singular	plural
first person	-am	-(e)mun
second person	-et/-at	-(e)tun
third person	-eʃ/-aʃ	-(e)ʃun

(2.23) mɑʃɪn-etun
car-2P
'your (PL) car'

(2.24) az-am gereft
from-1S took.3s
'(He/she/it) took (it) from me'

(2.25) ovord-am-eʃ
brough-1S-3S
'I brought it'

(2.26) dard-am umad
pain-1S came.3S
'(It) hurt me'

(2.27) koʤɑ-t dard mi-kon-e
where-2S pain DUR-do-3S
'Where does it hurt?'

(2.28) inʤɑ-m dard mikone
here-1S pain DUR-do-3S
'(It) hurts here'

EZAFE /-e/ enclitic

According to Karimi & Brame (2012), the EZAFE construction is a construction with the morpheme /-e/ (or /-je/ after vowels other than /i/) which links together the elements of a single constituent. The literal meaning of EZAFE is 'addition'. The EZAFE particle appears between any two items that have some kind of a connection, for example between a noun and its complement, an adjective and its complement or a preposition and its complement (Ghomeshi, 1997). It also generally appears between a first and last name. Examples are provided below, in (2.29) to (2.34):

(2.29) xordan-e ʔɑb
drinking-EZ water
'the drinking of water'

(2.30) montazer-e Ali
waiting-EZ Ali
'waiting for Ali'

(2.31) ʃahr-e tehrɑn
city-EZ Tehran
'the city of Tehran'

(2.32) poʃt-e madrese
behind-EZ school
'behind the school'

(2.33) pesar-e xub
boy-EZ good
'the good boy'

(2.34) Maryam-e Karimi
first name-EZ last name

Definite marker /-e/

Standard Persian has no overt definite article (Ghomeshi, 2003; Karimi, 1994). However, in colloquial Persian the suffix /-e/ marks definiteness. It appears as /-he/ when the word ends in /e/ (Kahnemuyipour, 2014). This suffix takes the stress. It marks 'determinedness' or 'fixedness' of the nominal (Jasbi, 2020) that is the referent is either known to the participants in the conversation or the referent is unknown but fixed. Two examples of DEF are provided in (2.35) and (2.36).

(2.35) xar-e
donkey-DEF
'the donkey'

(2.36) xar kuʧulu-e
donkey small-DEF
'the small donkey'

Plural /-ɑ/

Plural is most commonly marked with the suffix /-hɑ/ (in non-colloquial Persian) or /-ɑ/ (see example (2.37)). As mentioned earlier, the DEF suffix /-e/ and the PL suffix /-ɑ/ cannot appear together (Hincha, 1961 seen in von Heusinger & Sadeghpoor, 2020). It has been argued that /-ɑ/ marks both plurality and definiteness (Ghomeshi, 2003).

(2.37) ketɑb-ɑ
book-PL
'books'

Object marker /-o/

Suffix /-o, -ro/ is used in colloquial Persian as the object marker. /-ro/ is used when the word ends in non-high vowels (Hedberg et al., 2009) and /-o/ is used when words end in consonants (Mahootian, 1997). According to Ghomeshi (2003), non-referential bare nouns can be distinguished from definite bare nouns by the presence of the object marker /-o/ in direct object position (compare (2.38) with (2.39)).

(2.38) ketɑb xund-am
book read-1S
'I read books'

(2.39) ketɑb-o xund-am
book-OM read-1S
'I read the book'

Additive particle /-am/

The additive ADD particle /-am/ has two functions, the first being as a focus particle meaning 'also' or as a topic particle meaning 'as for' (Sato & Karimi, 2016). In the present study, we have only a few cases of this particle meaning 'also', as in (2.40):

(2.40) Ali-am mi-xɑ-d be-r-e
Ali-ADD DUR-want-3S SBJV-go-3S

‘Ali too wants to go’

Numerical classifier /-tɑ/

The default classifier /-tɑ/ marks count nouns (Ghomeshi, 2003; Mahootian, 1997), as in (2.41):

(2.41) do-tɑ ketɑb
two-CL book
‘two books’

The numeral must in most cases be followed by a classifier in colloquial Persian (Hamedani, 2011). The only exception is /je(k)/ ‘one’, which cannot be followed by /-tɑ/.

Allomorphy and regularity in Persian morphology

Agglutinating languages, such as Finnish, Turkish, Basque, and Hungarian, have affixes that can be easily distinguished and typically encode a single grammatical feature (Gervain, 2022). These affixes are added to the root of a word in a straightforward manner. Grammatical relationships in agglutinating languages are conveyed through the use of a large set of affixes, with each grammatical function being represented usually by a single morpheme (Ladányi et al., 2020).

Allomorphy refers to a situation in which a single lexical item, meaning, function, or morphosyntactic category can be expressed in different phonological forms depending on the context in which it is used (Paster, 2014). This can involve a single morpheme having multiple forms. Allomorphy is prevalent in Persian. This is evident in inflectional morphemes, which often have various allomorphs. For instance, it has been previously noted that for the object marker OM the allomorph /-ro/ is observed when words end in non-high vowels (Hedberg et al., 2009) while /-o/ is observed to be used when words end in consonants (Mahootian, 1997). Similarly, the imperative prefix /be-/ displays allomorphy with the forms /bi-/ and /bo-/. The IMP prefix /be-/ appears as /bi-/ before the vowels /a/ and /ɑ/, as in the imperative forms /bi-ɑr/ ‘IMP-bring-2S’ (= bring!). It appears as /bo-/ before the back vowels /o/ and /u/, as in /bo-xor/ ‘IMP-eat-2S’ (= eat!), and /bo-ro/ ‘IMP-go-2S’ (= go!). However, there are exceptions to this pattern, such as /be-bor/ ‘IMP-cut-2S’ (= cut!) (Mahmoodi-

Bakhtiari, 2018a). The following morphemes, which were discussed in the current chapter, exhibit allomorphy: NEG, IMP/SBJV, EZAFE, PL, OM, ADD, pronominal enclitics for 1st person singular, copular to be and personal suffixes for all person/numbers expect for 3S /-s/ (both past and present).

The verbal morphology of Persian is very regular (Mahmoodi-Bakhtiari, 2018a) yet complex. The non-verbal morphology of Persian can be characterized by its regularity and simplicity. Although the existence of allomorphy in Persian can present an additional challenge for children learning to acquire inflectional morphemes, the regular nature of inflectional morphology in Persian suggests that this process may be less difficult compared to languages with more complex inflectional systems such as Turkish.

Summary

The language under investigation in this study is Persian, and information about its adult phonology and morphology was provided. The typology of Persian was discussed, with a focus on the inflectional morphemes observed in the data of children. These included both verbal and non-verbal morphemes. Additionally, the topic of allomorphy and regularity in Persian morphology was addressed. This chapter aimed to provide a comprehensive overview of the linguistic features of Persian that are relevant to the study of the earliest steps of learning inflectional morphology. The discussion of these topics will serve as a foundation for the subsequent analysis of the data collected from children.

3. Methodology

Introduction

This thesis is based on the data collected from naturalistic speech samples collected longitudinally. Collecting naturalistic data on a longitudinal basis has been used as a big source of evidence for children's linguistic development since Brown (1973)'s seminal study. Brown's naturalistic data on Adam, Eve, and Sarah provided the qualitative and quantitative basis for new measures like MLU to be developed and used to asses children's linguistic development (Behrens, 2008). According to Behrens (2008), such naturalistic data provides a basis for qualitative analysis on the morphological development of children which can be used to list the appearance of morphemes and to assess their productivity.

The goal of a naturalist study is to provide a representative and varied sample of the child's everyday speech (Eisenbeiss, 2010). Despite the fact that naturalistic data collection and transcription can be very time-consuming, it can be used many times by the original researchers and others if made available publicly (Ambridge & Rowland, 2013). Another benefit of naturalistic studies is the close resemblance to the real-life situation (Eisenbeiss, 2010). With regards to studying morphological analysis, naturalistic data "provides very clear evidence for how children deal with morphological productivity" by sidestepping the methodological issues of behavioural methods (Lignos & Yang, 2016: p 776).

This does not mean that this method of data collection and analysis is without its potential problems. One potential problem can arise from analysing children's errors. Different results can be obtained by analysing the same type of errors in children (Rowland et al., 2008). According to Rowland et al. (2008), this might be either due to the impact of the sample size, or the choice of analysis technique. The former results in rare errors being missed in smaller samples. The latter results in different reliability of the error rate calculation. A second potential problem is estimating the child's knowledge incorrectly (Eisenbeiss, 2010). According to Eisenbeiss (2010) on the one hand, linguistically unchallenging routine activities might lead to underestimating the child's knowledge by not providing appropriate context for producing certain forms. On the other hand, producing frequent formulaic or semi-formulaic patterns might lead to overestimating the child's knowledge as the occurrence of an element cannot simply be regarded as evidence for its acquisition. Consequently, the findings of

naturalist studies need to be interpreted qualitatively and with caution.

Data collection

For the purpose of this study, ethical approval for this research was obtained from the L&LS Ethics Committee. Recruiting participants living in Tehran began through family, friends, and social media (Appendix 1 – social media participant recruitment poster) in two phases. Phase one took place in summer 2017 and phase two in summer 2018. Seven families (eight infants) were recruited to take part in the study in 2017 and seven more families (seven infants) in 2018.

The families were contacted by phone, given information about the study, and invited to take part. The information sheet (Appendix 2 – Information Sheet (English): data collection 1 and Appendix 3 – Information Sheet (English): data collection 2) was sent to them electronically. The families were given a few days to decide whether they wanted to take part. Written informed consent (Appendix 4 – Consent Form (English)) was obtained prior to the first recording session.

In 2017, five (out of seven) participating families dropped out at different stages of the data collection. Two of the children (in one family) were unwilling to speak in the presence of the researcher and the camera, despite having a few home visits without the equipment prior to the beginning of the data collection. Hence data collection with this family was discontinued. This means that in 2017 only one child's data was recorded, for five successful sessions. In 2018, however, all seven participating families were successfully recorded for at least six sessions.

For this research spontaneous speech samples were collected under naturalistic conditions. The interactions between the infants and their parents/caregivers were audio-video recorded in the infant's home (or in one case the grandparent's home). In one case, after audio-video recording the seven initial sessions, the parent of the infant audio-recorded five more sessions. The parent/caregiver was asked to engage the infant in an activity likely to lead them to produce some words (e.g., reading books, playing with toys, drawing pictures, etc.). Before the start of each session, the parent/caregiver was asked if there were any newly learnt words. If so, the parent/caregiver was asked to try to elicit those words from the infant during the usual

activities.

Prior to the first recording session each parent/caregiver was asked to fill in the Persian adaptation (Kazemi et al., 2016) of the MacArthur-Bates Communicative Development Inventories (CDI) (Fenson et al., 1993) and a linguistic background questionnaire prepared by the researcher (see Appendix 8 – Linguistics background form (Persian)) . The parent/caregiver was asked to mark the words they thought the infant understood and/or said. The CDI made it possible to identify the number of words the infant could say/understand based on the judgement of the parent/caregiver, while the linguistic background questionnaire consisted of questions about the parents' linguistic background and what languages the infants were being exposed to via their parents and any other family members or caregivers. The questionnaire also contained a few questions designed to ascertain how much time each parent/caregiver spends with the infant. Parental report is believed to be a good way to obtain a global measure of the infant's linguistic ability (Ambridge et al., 2013); it can also be used an as an extra clue in the word identification process.

Each recording session lasted between 30-58 minutes (mean = 41.5). The researcher recorded the sessions in phase one with two Sennheiser EW 100-ENG G2 Wireless microphones and a Canon XA30 camera. One microphone was attached to the infant's vest and one to the parent/caregiver's shirt. For phase two, Zoom Q8 handy video recorder and two Tascam DR-10L recorders with lavalier microphones were used to audio-video record the sessions. As mentioned above, in the case of one of the infants the parent continued audio-recording five more sessions with one Aqta at-1260 Wireless microphone and a Sony ICD-PX470 Digital Voice Recorder. In total, 47 hours of child-caregiver interactions were recorded, out of which 27 hours of child speech were transcribed.

Participants

The infants were selected based on the following criteria:

I. The infant is full term.
II. The infant has no known hearing/medical problems.
III. The infant is being brought up in a household where only Persian is spoken.
IV. The parents are both native speakers of Persian.

V. The infant does not yet produce word combinations.[7]

An exception was made with regards to the third criterion in one case. The Persian-German bilingual father attempted to expose the infant to German upon returning home from work, accounting for approximately half the time he spent with her. Figure 3 presents an age timeline of all the infants that participated in the study for a minimum number of five complete recording sessions. Each timeline shows the participant's session numbers and their age. Each data point has two numbers. The number above the data point is the session number. The number below is the participant's age at the time of that session.

[7] Initially this study required that each participant know approximately 50 words, but this criterion was eventually replaced with this last criterion in order to gather data from all the families willing to take part.

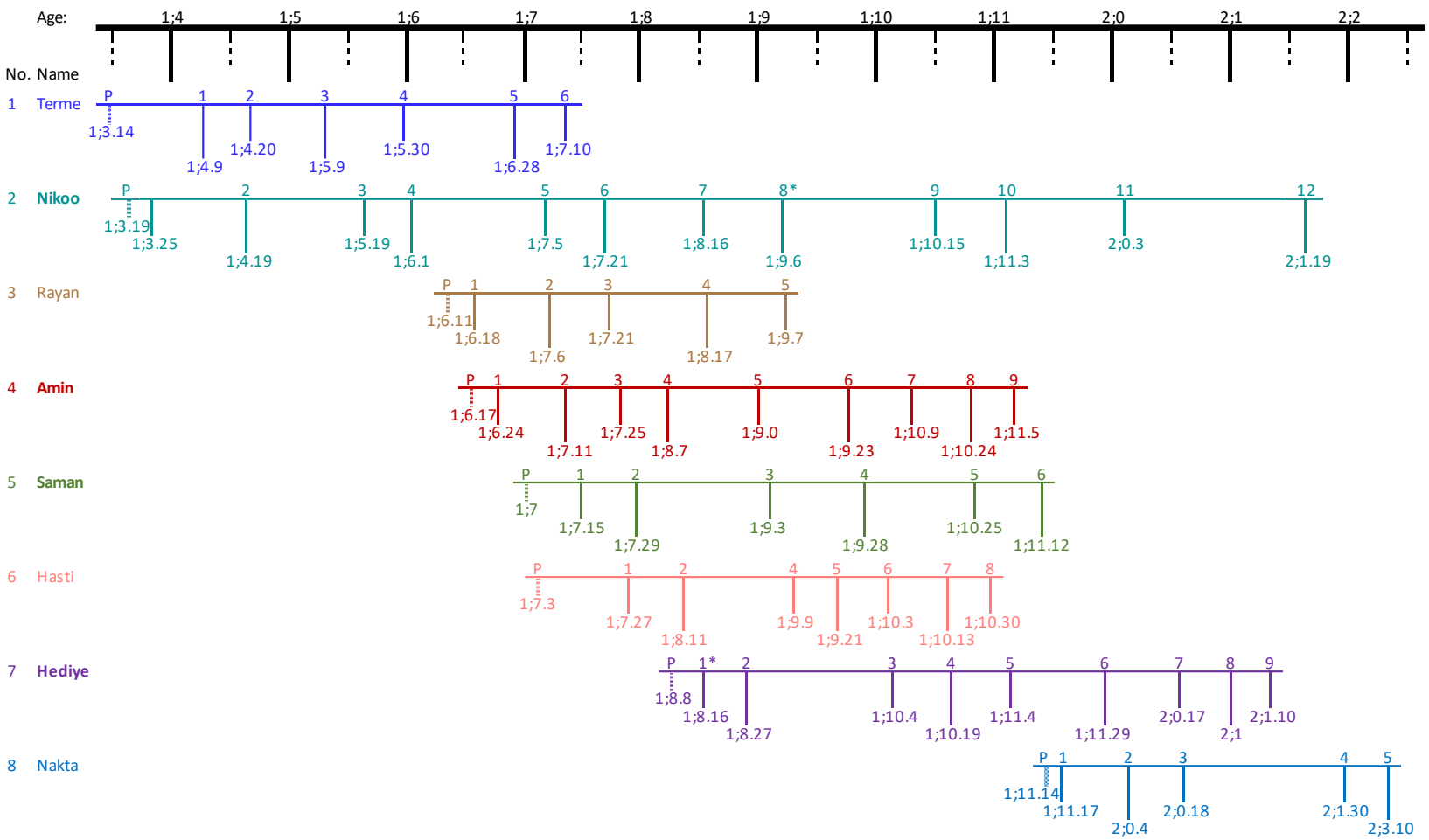

Figure 3 Participant timeline

The number above each data point represents the session number.
The number below each data point represents the infant's age at the time of that session.
The names of the infants whose data will be analysed in this study are in marked in **boldface**.
P: The pre-session when the CDI questionnaire was filled.
*: Cancelled session or a session with a poor sound quality.

The researcher chose to transcribe in full the sessions of four of the infants that seemed to be the most likely to exhibit developmental changes over the course of the study. The selection of these four participants for the study was based on the criterion that they demonstrated a development trajectory consistent with expected norms. Specifically, the remaining participants were excluded from the study as they did not exhibit sufficient evidence of development according to these norms. For instance, Terme did not begin using inflectional morphemes until the final recording session, while Hasti displayed no indication of having reached the two-word stage or using inflectional morphemes. Rayan was excluded from further analysis, despite the fact that his sessions were fully transcribed, as it was determined that his exceptional memory would impede the ability to determine if his longer utterances were the result of productive language use or simply repeated statements he had heard previously. This posed a significant challenge in determining his true language development. Ultimately, participant Nakta was excluded from the study due to a preliminary analysis indicating that she was more advanced than the other participants. Specifically, her ability to form long sentences was observed during the first recording session, which suggested that she was beyond the scope of the current study's focus.

Table 9 shows the participants' pseudonyms, number of sessions, gender, receptive (R) and productive (P) CDI scores and age on the day their parents/caregivers were asked to fill in the CDI list. As can be seen in the following table, a total of 36 sessions were recorded with these four infants.

Table 9 Infants' information

Number	Infant's Pseudonym	No. of Sessions	Gender	CDI Score: Receptive	CDI score: Productive	Pre-session: Age
2	Amin	9	male	311	36	1;6.17
1	Saman	6	male	157	23	1;7.0
4	Nikoo	12[8]	female	28	4	1;3.19
3	Hediye	9[9]	female	270	50	1;8.8

[8] Nikoo's 8th session was disregarded due to technical problems with the recording; hence only 11 of the 12 recorded sessions were used for analysis.
[9] Hediye's 1st session was disregarded due to the child being unwell.

All the infants were born and live in Tehran. Additionally, all the infants are from monolingual Persian-speaking families except for Hediye, as mentioned earlier.

Method and analysis

Transcribing and coding

Synchronising the recorded audio and video files was done with Adobe Premiere Pro CC 2017. Infants' vocalizations were phonetically transcribed with ELAN software (Sloetjes & Wittenburg, 2008), using the International Phonetic Alphabet (IPA). All infant productions that sounded like adult words were listed as word candidates. Other types of vocalizations (jargon/babble/grunt) were marked but excluded from the analysis. Vihman and McCune (1994)'s word identification criteria were employed to arrive at a decision as to whether a candidate could be counted as a word. These criteria include four based on contextual evidence, three based on vocalization shape, and four based on the word candidate's relation to other vocalizations. After rating the word candidates based on the criteria, all the infant vocalizations accepted as words/phrases were listed. Modelled words are words a child hears and produces after some intervening speech either by the adult or the child (Macken & Barton, 1979), whereas imitated words are those the infant imitates instantly upon hearing. Infants' modelled or imitated words, as well as words sung from nursery rhymes were listed. These were excluded from the tally of spontaneous word production.

All onomatopoeia were excluded from this analysis, with the exception of the words used to refer to either 'cat' [mio] (Saman & Hediye) for /gorbe/ or 'car' [bibib] (Saman) for /mɑʃin/. In the case of [bibib], the child started adding inflectional morphemes to his production: [bibib-ɑ] 'car-PL' (Saman 1;11.12). In this specific case, even the parents had adopted the child form over the course of data collection, especially in the earlier sessions.

Persian CDI

More than 680 words are listed in the Persian adaptation (toddler form) (Kazemi et al., 2016) of the CDI (Fenson et al., 1993). These were analysed to establish a basic understanding of what words 18- to 30-month-old children are likely to know or at least have been exposed to (Kazemi et al., 2008). CVCVC (130 words) is the most frequent word structure in the CDI.

This is consistent with descriptive studies of the adult language. The next most frequent structures are CVC (93 words), CVCV (77 words), and CVCCVC (65 words). These structures together account for around 50% of what an infant might hear up to 30 months of age.

Linguistic measures

In this study, various measures are employed to examine the relative significance of lexical, phonological, and syntactic development in relation to infants' morphological development. To conduct in-depth analyses of the children's phonological, lexical, representational, and morphological development, it is necessary to utilize certain measures. These measures will allow us to quantify the child's morphological development and evaluate its relation with other aspects of the child's linguistic development. These measures are outlined below (Figure 4).

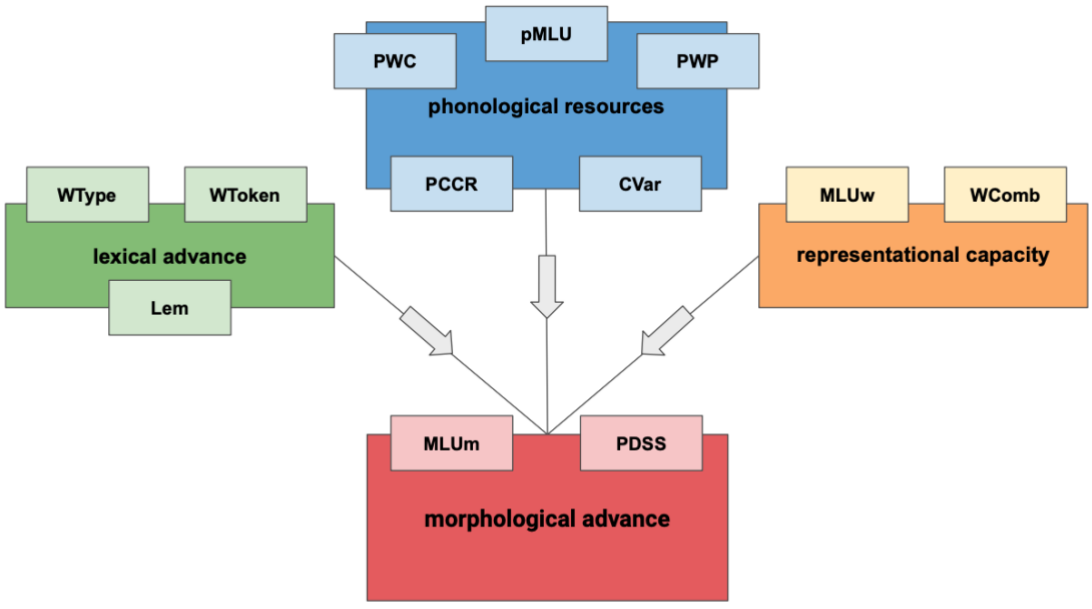

Figure 4 Model: measures

Lexical development

In order to investigate the lexical development of the children certain measures were calculated. These measures are lemmas, word types, and word tokens. Lemma refers to "the abstract base of a lexical entry (often called lexeme), i.e., to the correlation of (specific) lexical meaning with (specific) phonological material, which creates the lexical sign" (Bittner et al. 2003, p. XXXIX). Lemma is the uninflected base of a word. Verb lemmas in Persian are presented in the infinitive form, for example *raftan* 'to go'. These forms end in either *-tan* or *-dan*. Word type refers to an inflected form of a lemma. Finally, word token refers to the individual occurrence of an (un)inflected form of a lemma in the child's speech.

The child's word type and token counts are each categorized by two measures: 1. all word types (including imitated and modelled production) and 2. spontaneous (excluding imitated and modelled production). This results in five *count* measures: 1. Spontaneous word tokens (SpnWToken) 2. Spontaneous word types (SpnWType) 3. All word tokens (AllWToken) 4. All word types (AllWType) 5. Lemmas (Lem).

Wherever relevant, certain 'word points' will be used as a basis for comparing the children's development in a more meaningful way (i.e., rather than based on age), following Vihman et al. (1986). The points chosen here are 5-, 25-, 50-, and 100-words. The 5-word point (5wp) represents the session in which the child used five or more spontaneous words in one recording session.[10] At the 25-word point the infant could be expected to have a cumulative vocabulary of fifty words or more (Vihman & Miller, 1988). We can expect the infant to produce their first word combinations within a month after the infant has arrived at the 25-word point, which marks the end of the single-word period (Vihman, 2019).

First identifiable words

According to Ferguson and Farwell (1975), children's first identifiable words are rather accurate despite the fact that both their syllable length and syllable structure are limited. The infants' first identifiable recorded words will be presented.

[10] It is worth noting that in Vihman et al. (1986), where the 4wp is used instead, the recording sessions were a half-hour long, while here they were around 40 minutes long.

Syllable structures

Syllable structures refer to the range of overall shape and length in syllables of the infant's forms (Vihman, 2016). Following Vihman (2016), each structure was counted as a variant word shape in the analysis if it accounted for at least 10% of the words identified in that session. Variants are different shapes of a single word type (i.e., word shapes). Based on the procedure outlined in Vihman (2019), first all the words produced in the session were categorized based on their word shape to arrive at the total number of syllable structures. Minor differences in vowel quality were disregarded, as were voicing differences. Different word shapes of a single word type were included as long as they gave rise to distinct syllable structures. An instance of this is Amin's [za:], [da:], and [za:d] for *zard* 'yellow'. In this case [za:] and [da:] are counted as CV and [za:d] is counted separately, as CVC, because of the presence of a coda.

Phonological development

Firstly, to measure the child's articulatory resources and planning, their consonant inventory size, and initial syllable structure will be analysed. Secondly, to be able to follow their phonological development, certain measures will be used. Percentage consonant correct-revised (PCC-R), as a segmental measure on the one hand, and phonological mean length of utterance (pMLU), percentage whole-word proximity (PWP), and percentage of whole-word correctness (PWC) as whole-word measures on the other hand, will give us an insight into the child's phonological development and its development.

According to Stoel-Gammon & Stone (1991) in order to arrive at a thorough assessment of a child's phonological development, both relational and independent analyses are needed. Relational analyses are carried out with reference to the adult form, while independent analyses are carried out without comparing the child forms to the adult targets. To be able to paint a complete picture of the children's phonological development both types of analyses will be done in this study. The child's consonant inventory size, initial syllable structure, and consonant variegation (CVar) measure help us to describe the child's phonological development and system independently. While PCC-R, pMLU, PWP, and PWC will be used with reference to the adult target forms. In relational analyses, a comparison is made between the child form and the adult form to be able to identify correct versus incorrect productions (Saaristo-Helin, 2009).

Consonant inventory

All the consonants that the infant produces are identified in each session, with at least two examples in each of the two positions (onset and coda), following Vihman (2019). Voicing is considered only in assessing consonant inventory size. Non-target consonants were included only if a target use was already present in the child's consonant inventory. Non-Persian consonants are included if produced by the infant. If the consonant in question was not present in the target word, it is not counted. For example, /r/ in [ɑrdɑms] for *adams* 'gum' (Hediye, 1;10.19) is not counted. If the consonant in question is produced in different positions in the adult target as compared with the child form, it is counted for the position in which the infant produced it. For example, /r/ in [ɑr] for *are* 'yes' (Amin, 1;9.23), is in coda position in the infant form and onset position in the adult target form. In addition to those used in Vihman (2019), one criterion was added for Persian: Geminated consonants were counted only once at onset. The quantification of the consonant inventory is based on word types rather than lemmas. For instance, /m/ is counted for both [man] 'I' and [man-am] 'I-be.1s' (Amin, 1;9.0).

Consonant variegation score (CVar)

Producing two different supraglottal consonants in a word might be the biggest challenge a child faces in the single-word period (Vihman, 2014). Following Vihman et al. (2013), a within-word consonant variegation (CVar) measure is used to measure the child's ability to produce variegated consonants. This is an extension of Stoel-Gammon (1989)'s mean babbling level.

Each infant form is given points on a scale of one to three. If the word produced contains only vowels, glides, /h/, or /ʔ/ (no true consonants), the form gets one point. If the word contains one and only one 'true' (supraglottal) consonant (no more than one consonant), the form receives two points. Finally, if the word has two or more different consonants (more than one true consonant), the form receives three points. Based on this scoring system, *aj* 'ouch', *mahi* 'fish', and *dom* 'tail' get one, two, and three points respectively. Voicing differences are disregarded, since a change in voicing only does not constitute consonant variegation (Vihman, 2019). For example [da:tuh] for *xarguf* 'rabbit' receives two points.

Variegated child forms receive variegation scores even if the adult form is not variegated (Vihman, 2019). For example [dini] for *nini* 'baby' gets a score of two. That is to say, the child's consonant accuracy is not taken into consideration in this scoring system. Stray consonants are not included if not present in the adult target, for example [ben] for *in* 'this' receives only two points.

Percentage consonant correct – revised (PCC-R)

PCC-R (Shriberg et al., 1997) is used to calculate the phonological accuracy of the consonants the infant produces. The PCC-R index is calculated by dividing the number of correct consonants in the child form by the total number of consonants, multiplied by 100. The only difference between PCC and PCC-R is that only omissions and substitutions are counted as incorrect and all other sound distortions, including addition, are considered correct (Gruber, 1999).

In the adult Persian, there are some final consonants in word-final consonant clusters that are often (but not always) deleted. For the purpose of calculating PCC-R in this study, both the presence and deletion of such consonants were counted as correct, since both are correct in the adult language. For example, [got] (Amin 1;10.24) for *gof(t)* 'said.3S' received a score of two; one point for /g/ and one for /t/ and [dɑ:ft] (Amin 1;8.7) for *raf(t)* 'went.3S' received a score of two, one for /f /and one for /t/.

Misplaced consonants were considered correct if there was enough evidence of their misplacement, e.g. in [bi-ja-dɑ:-nih] (Nikoo, 1;11.3) for *dar-bi-jɑr-im* 'off-SBJV-take-1P' (= so that I take (it) off), [dɑ:] for /dar/ was produced as a whole after [bi-ja] for /bi-jɑr/.

(3. 1) AF: dar - bi - jɑr - im

CF: bi - ja - dɑ: - nih

If there was not enough evidence for consonant misplacement, they were considered incorrect. For instance, there was no evidence that /m/ was misplaced in [dededʃɑm] (Nikoo,

1;11.3) for *medɑdrangi* 'colour-pencil'.

Wrongly placed consonants were not counted as correct if there was no evidence, for example [dɑ:ʃi] (Hediye 1;11.29), for *xorʃid* 'sun' receives one point out of four, while [a:ʃid] (for the same target) receives two points. This is also true for [taxa:] (Hediye 2;1.10), for *deraxt-e* 'tree-be.3s', which means it receives one point out of four. In cases where two different consonants are accepted in the adult target, the point is given, for example [babaji] (Hediye 2;1.10) and [babaʔi] (Hediye 2;0.17) are accepted for *babaʔi* 'sheep (CD[11])'. Extra final /h/ and /ʔ/ consonants are counted as correct but are not counted towards the total number of consonants.

Phonological mean length of utterance (pMLU)

pMLU (Ingram, 2002) calculates the accuracy of the child form as a whole. pMLU is calculated for both the adult words and the child forms. One point is assigned for each correct segment (consonant or vowel) in each word. An extra point is assigned for each correctly positioned consonant. For calculating pMLU the rules summarized in Saaristo-Helin (2009) were used, with some modifications (to be explained later). Six of these rules were introduced by Ingram (2002) and two were added by Bónová, Slancová, and Mikulajová (2005). Ingram (2002)'s rules concern the sample size, the lexical class of the words, compound words, variability of the words, production, and finally the consonant correct rules. Bónová et al. (2005)'s rule concern position and input.

Not all of Ingram (2002)'s rules can be easily applied to languages with different structures (Saaristo-Helin, 2009). Following two of Ingram (2002)'s rules was particularly challenging in this research. The first challenging rule was his sample size which suggests having a selection of at least 25 words, and preferably 50, depending on length of the recording. Here, however, to be able to include all the sessions in the data analysis as well as to follow the other pMLU rules, the minimum sample size rule was relaxed. As a result, all available words that fit the other rules were analysed.

The second challenging rule was Ingram (2002)'s third rule: compound (noun) rule. This rule has only been partially applied (which is based on the spelling of the word) because this rule

[11] CD: Child Directed word

cannot be fully applied to Persian. Ingram's criterion for compound nouns is based on spelling the word. This rule states that words that are spelled as a single word should be included while those spelled as two words should be excluded from the count (*cowboy* vs. *teddy bear*). This spelling criterion simply cannot be applied to Persian. Thus, all compound nouns were excluded. Excluding these compound nouns does not affect the score since the number of compound nouns in the data was very limited.

The main issue regarding the compound rule is Persian's compound verbs. The vast majority of Persian verbs are compound (Mahmoodi-Bakhtiari, 2018a). Since verbs (as opposed to nouns) are heavily inflected in Persian, excluding them would also considerably limit the morphological analysis. Moreover, Taelman et al. (2005) revealed that pMLU also partly reflects the morphosyntactic proficiency level of the child in languages like Dutch. Thus, keeping the compound verbs seemed crucial when studying the effect of the children's phonological advance on their morphological development.

Since Persian has geminated words, Kunnari et al. (2012)'s criteria were used to calculate the scores for adult geminated words. For these words, geminated consonant is given three points, two points in the segment count and one point for the correct consonant count. For example, *ab:azi* 'water game' is awarded eight points; six for its segments and two more for /b/ and /z/.

In calculating a pMLU score for words with optional final-consonant deletion (e.g. *nist* 'NEG.be.3S' and *raft* 'went.3S'), the forms with the deleted consonant are considered as the adult target (e.g, *nis* and *raf*), so that *nist* receives four points for /n, i, s/ and two more for /n/ and /s/; final /t/ receives no points. When pronominal pronouns are added to these base forms however, /t/ is always pronounced. For example, in *nist-am* 'NEG.be-1S' /t/ receives two points.

Taelman et al. (2005) discusses the lack of clarity in Ingram (2002)'s strict ordering constraint. They provide an example regarding the way to measure pMLU in the child forms' [lip] and [pil] for the adult target word *lip*. If no strict ordering constraint is applied, then both of these child forms would get the same score. For this research, a strict ordering constraint was applied in shorter words like *tup* 'ball'. But in longer words where one whole morpheme was misplaced, the strict ordering constraint was relaxed.

Percentage whole-word proximity (PWP)

PWP (Ingram, 2002) is the degree of accuracy in producing the words. It is calculated by dividing the infant's pMLU by the pMLU of the adult form. This measure provides an insight into the relation between the child's production to the adult target.

Proportion of whole-word correctness (PWC)

PWC (Ingram, 2002) measure is calculated by the number of entirely correctly produced words in a session divided by the total number of words produced in that session. This simple measure can tell us what proportion of the child's production is produced correctly as a whole.

Syntactic development

The two measures for determining the syntactic development of the child are the number of word combinations (WComb) produced in each session, and an analysis of MLUm.

Mean length of utterance in words (MLUw)

Following Brown (1973), mean length of utterance in words (MLUw) is calculated for each infant on a session by session basis. MLUw is the average number of words per utterance; this serves as another index of expressive language development.

Morphological development

Investigated morphemes

We analyse the extent of use of all of the inflectional morphemes that occur in one or more of the four children's first forms that potentially include an inflection marker. These inflection markers consist of five verbal prefixes [negative, imperative negative, durative, imperative, subjunctive], seven verbal agreements (pronominal suffixes), seven clitics of the copula 'to be', six pronominal enclitics, past participle marker, and six non-verbal morphemes [EZAFE particle, plural suffix, object marker, additive marker, classifier marker, and the definite marker]. A list of these inflectional markers can be seen in Table 10.

Table 10 Investigated morphemes

Morpheme	Type	Structures used in	Person/Number	Verbal/Non-verbal	Suffix/Prefix
/na-, ne-/	imperative negation	+ V	NA	verbal	prefix
/na-, ne-/	negation	+ V	NA	verbal	prefix
/mi-/	durative	+ V	NA	verbal	prefix
/be-, bo-/	imperative	+ V	NA	verbal	prefix
/be-, bo-/	subjunctive	+ V	NA	verbal	prefix
/-am/	V agreement	V +	1S	verbal	suffix
/-i/	V agreement	V +	2S	verbal	suffix
null*	V agreement (imperative only)	V base form	2S	verbal	suffix
/-e/	V agreement	V +	3S	verbal	suffix
null*	V agreement (past tense only)	V base form	3S	verbal	suffix
/-d/	V agreement (present tense V stem ending with vowel)	V +	3S	verbal	suffix
/-im/	V agreement	V +	1P	verbal	suffix
/-in/	V agreement	V +	2P	verbal	suffix
/-an/	V agreement	V +	3P	verbal	suffix
/-am/	copula	ADJ +, P +, N +	1S	verbal	suffix
/-i/	copula	ADJ +, P +, N +	2S	verbal	suffix
/-e/	copula	ADJ +, P +, N +	3S	verbal	suffix
/-im/	copula	ADJ +, P +, N +	1P	verbal	suffix
/-in/	copula	ADJ +, P +, N +	2P	verbal	suffix
/-an/	copula	ADJ +, P +, N +	3P	verbal	suffix
/-as, -s/	copula	ADJ +, P +, N +	3S	verbal	suffix
/-am/	pronominal enclitic	N +, P +, V +, ADJ +, QW +, Dem P +	1S	verbal/non-verbal	suffix

/-et/	pronominal enclitic	N +, P +, V +, ADJ +, QW +, Dem P +	2S	verbal/non-verbal	suffix
/-eʃ/	pronominal enclitic	N +, P +, V +, ADJ +, QW +, Dem P +	3S	verbal/non-verbal	suffix
/-emun/	pronominal enclitic	N +, P +, V +, ADJ +, QW +, Dem P +	1P	verbal/non-verbal	suffix
/-etun/	pronominal enclitic	N +, P +, V +, ADJ +, QW +, Dem P +	2P	verbal/non-verbal	suffix
/-eʃun/	pronominal enclitic	N +, P +, V +, ADJ +, QW +, Dem P +	3P	verbal/non-verbal	suffix
/-e/	past participle	V +	NA	verbal/non-verbal	suffix
/-ɑ/	plural	N +	NA	non-verbal	suffix
/-o, -ro/	object marker	direct object	NA	non-verbal	suffix
/-e/	Ezafe enclitic	N + ~ + compliment, ADJ + ~ + compliment, Prep + ~ + compliment	NA	non-verbal	suffix
/-e/	definite	N +	NA	non-verbal	suffix
/-am/	additive	N +, ADJ +, Dem P +	NA	non-verbal	suffix
/-tɑ/	number classifier	N +	NA	non-verbal	suffix

* Null included as it seems to be the base form for the children

Obligatory context is established for these morphemes (whenever possible), based on Cazden's (1968) criteria.

We identify the first appearance of inflectional morphemes as well as productivity of each of them, using the criteria introduced by Pizzuto & Caselli (1994), modified for Persian by Marvasti (2014). Pizzuto & Caselli's productivity criteria require the stem to appear in at least

two distinct forms and the inflection to appear with at least two different verbs (i.e., contrastive use of verb type and inflection, respectively). Given the structure of Persian verbs, which can simultaneously have a prefix and a suffix attached to the stem, Marvasti's modified criterion states that for a given prefix or suffix to be productive "the verb stem should appear in at least two distinct forms in either prefix or suffix position, not in both" (Marvasti, 2014: p 78). Marvasti's criteria are expanded here to apply to non-verbal morphemes as well. This means that the contrastive use of non-verbal stems is also established.

Morphological measures

The main measure that will be used to analyse the child's morphological development is: PDSS. However, since PDSS is not purely a morphological measure MLUm will be used a second measure.

Persian developmental sentencing score (PDSS)

As mentioned before in chapter 1, PDSS is a tool used in clinical settings to evaluate the morphosyntactic development of Persian-speaking children (Jalilevand et al., 2016). PDSS was calculated based on the criteria in Jalilevand (2017). Jalilevand's PDSS system is an adaptation of Lee (1974). Lee (1974)'s DSS is a tool for evaluating morphology and syntax. It was normalised on 200 children. Similarly, PDSS can be used to evaluate Persian-speaking children's performance on morphosyntax (Jalilevand et al., 2016). It is based on analysing the data collected from 115 children between 30 and 65 months (Jalilevand, 2017; Jalilevand et al., 2016).

This score is calculated for eight sub-categories[12]:

I. Verb morphology
II. Verb structure (modal and compound verbs)
III. Prepositions and conjunctions
IV. Pronouns
V. Question words
VI. Grammatical morphemes

[12] As mentioned in chapter 2, the focus of this thesis is inflectional morphemes and clitics. However, as can be seen here, functional morphemes are also measured in PDSS.

VII. Sentence type
VIII. Sentence structure

Any correctly produced item belonging to any one of these categories can be scored between one and six. Score one is given to items that appear early in the child's speech and score six is given to items that appear very late (Jalilevand, 2017). For example, NEG prefix /na-, ne-/ is given a score of one since it appears earlier than the plural suffix /-ɑ/ which is given a score of two. Another example is the pronoun *man* 'I' with a score of one compared to the pronoun *unɑ* 'they' with a score of three.

Jalilevand (2017) excluded all one-word utterances from her analysis. Here, however, following Miyata et al. (2013) and their criteria for Japanese, all complete utterances (including single-word utterances) have been included in the study. As mentioned earlier, Persian, like Japanese, is a pro-drop language. Sentences consisting of a single verb are considered complete sentences. In this analysis such single verbs receive no sentential scores, but they do receive the other relevant scores, if available. This is to ensure that short single sentences do not receive the same weight as longer and more complex sentences. This means that single noun utterances are also included in the analysis. If they do not show any evidence of the child's morphological knowledge, they do not contribute to the score. If they display any evidence of the child's morphological knowledge, they will contribute only to the score for the relevant sub-categories. Again, these receive no sentential scores.

Mean length of utterance in morpheme (MLUm)

MLUm (Brown, 1973) is the average number of morphemes per utterance. MLUm is calculated by dividing the total number of morphemes in a given utterance by the total number of utterances produced in a session. Only correct uses of morphemes are counted towards MLUm.

Following Ghaderniya et al. (2019), derivational morphemes are excluded from the calculation. That is, only free morphemes and bound inflectional morphemes are counted. No Ø morphemes in the base forms are counted. For example, in /bo-xor/ 'IMP-eat.2S' the unexpressed person/number (Ø) morpheme is not counted. This means /bo-xor/ receives a score of two, that is one for the prefix /mi-/ and one for the stem. One-morpheme utterances

are also included.

Comparing MLUm and PDSS as indicators of morphological development

Both PDSS and MLUm are measures that are commonly used to assess different aspects of language development in children. While PDSS is a measure of syntactic complexity and is used more in clinical settings and MLUm is a measure of morphological complexity, it is important to consider whether one of these measures is a more reliable indicator of a child's morphological development.

There are several arguments in favour of using MLUm as the primary indicator of a child's morphological development. First, MLUm is specifically designed to measure the average number of morphemes per utterance, making it more directly relevant to morphological development. Additionally, MLUm has been widely used in research on child language development and has been found to be a reliable predictor of a child's language abilities.

On the other hand, PDSS may also be a useful indicator of a child's morphological development. This is because morphological development is closely related to syntactic development and is a measure of syntactic complexity. Therefore, a child who is producing more complex sentences, as measured by PDSS, may also be demonstrating advanced morphological skills.

One potential limitation of MLUm is that it does not take into account the complexity of inflectional morphemes, which has been noted in the literature. This can potentially lead to an incomplete understanding of a child's overall morphological development. That is to say MLUm may not provide a comprehensive picture of a child's morphological skills, as it only measures the mean number of morphemes produced regardless of complexity. In contrast, PDSS may not be as useful for comparing the development of children from different language backgrounds, as it is based on the structure of Persian. This can limit its generalizability to children from other language backgrounds.

Several researchers have identified certain limitations with the use of MLU (Jalilevand et al., 2016). For instance, Miller and Chapman (1981) have reported that children of the same age can exhibit variations in their MLU. Klee and Fitzgerald (1985) also noted that while a strong

correlation between MLU and chronological age exists among typically developing children, MLU is only considered a reliable indicator of development up until around 3.0 morphemes in the stage II of the Brown's model (Brown, 1973) highlighting the need for a different assessment in language development research.

While there is some overlap between PDSS and MLUm, it is worth noting that they are measuring somewhat different aspects of language development. DSS focuses on the syntactic complexity of sentences (which also includes morphological development as well as factors like sentence type, use of question words and pronouns), while MLUm focuses on the morphological complexity of individual words in each utterance. In a language like Persian that is rich on verb morphology, it is possible for a child's MLUm and PDSS scores to be elevated due to the use of morphologically complex verb forms. However, it is important to note that the use of more advanced pronouns and question words, for example, may also lead to a higher PDSS score that does not accurately reflect the child's morphological development. In these cases, the PDSS score may not be a reliable indicator of the child's progress in morphological development. Therefore, it may be useful to use both PDSS and MLUm together in order to gain a more comprehensive understanding of a child's language development. These limitations should be taken into consideration when utilizing PDSS and MLUm as measures of morphological development. It is important to note that these measures might be best to be used in a complementary manner, rather than as substitutes for one another where possible as it may be most useful to use both measures together in order to gain a more comprehensive understanding of a child's language abilities.

Inflectional error analysis

The first step in the process of analysing children's erroneous production of a morpheme in a required context is to determine the obligatory context for that morpheme. Obligatory contexts can be established from the recordings based on Cazden (1968) and the concept of obligatory context introduced by Brown (1973). Two types of error are identified: errors of omission and errors of commission. If the morpheme in question is missing in the child form, this counts as omission. For example, [be-be] for /be-bin-am/ 'SBJV-see-1S' (Saman 1;9.3) is an omission error since Saman omitted 1S /-am/. If the child form contains the wrong morpheme, this counts as commission. An example for this type of error is [zad-as] 'yellow-be.3s' for /zard-e/ 'yellow-be.3s' (Amin 1;10.9) in which Amin used 3S copula /-as/, instead

of the other form of 3S copula /-e/.

The obligatory context cannot be determined for all the morphemes studied here. Among verbal prefixes, /be-/ is often omitted in imperative and subjunctive forms of verb *kardan* 'to do'. This means /bu bekon/ 'smell IMP-do.2s (smell!)', as in the example (3. 2), is often produced as /bu kon/. Hence the obligatory use of this morpheme cannot be determined in this case (Marvasti, 2014). However, when used with verb *dadan* 'to give', the obligatory context can be determined (see example (3. 3)).

(3. 2)	CF:	mɑmɑh	bu	don
	AT:	mɑmɑn	bu	kon
		mum	smell	(IMP) do.2S
		'Mum! Smell (it)!'		

Amin 1;10.9

(3. 3)	CF:	hol-am	deh
		push-1SPC	give.2S
	AT:	hol-am	be-de
		push-1SPC	IMP-give.2S
		'Push me!'	

Amin 1;10.24

Another example is the object marker /o-/. Determining the obligatory context for this morpheme is more straightforward in some cases than in others. For example, in the two examples (3. 4) and (3. 5) below, both verbs require direct objects and hence both child forms can be regarded as having obligatory context for OM /-o/.

(3. 4)	CF:	hɑpuʔ-e	mi-jɑ-t	man	xoj-e
		doggie-DET	DUR-come-3S	I	eat-3S
	AT:	hɑpuʔ-e	mi-jɑ-d	man-o	bo-xor-e
		doggie-DET	DUR-come-3S	I-OM	SBJV-eat-3S
		'The doggie would come to eat me'			

Amin 1;10.24

(3. 5) CF: da-ʃɪ bi-de
lid-POSS.3SPC IMP-give.2S
AD: dar-eʃ-o be-de
lid-POSS.3SPC-OM IMP-give.2S
'Give its lid to me!'

Nikoo 2;0.3

If we compare (3. 4) and (3. 5) to (3. 6) it is apparent that two very close possible adult forms (AT1 and AT2 in the following example) exist.

(3. 6) CF: nɑrega be-de
tangerine IMP-give.2S
AT 1: nɑrengi be-de
tangerine IMP-give.2S
'Give (me) tangerine/tangerines!'
AT 2: nɑrengi-o be-de
tangerine-OM IMP-give.2S
'Give (me) the tangerine!'

Nikoo 2,0.3

In this study, error analysis was not carried out on child data produced prior to the onset of inflectional morphology use. That is, potential omission errors were not identified before the onset of morphology. This means the identification of obligatory contexts began only when there was evidence for the relevant morphological development. For example, (3. 7) was not counted as an obligatory context for EZAFE particle, compared to (3. 8). As can be seen in (3. 7), the N + N[13] order is not yet adult-like, which would require switching the order of the two nouns. Although, in (3. 8), despite the omission of both the EZ particle and 3S copula, the word order is adult-like.

[13] N: Noun

(3. 7) CF: babaji bɑst
sheep clothes
AT: lebɑs-e babaji
clothes-EZ sheep
'sheep's clothes'

Amin 1;9

(3. 8) CF: ɑʃɑ mah
proper name I
AT: ɑʃraf-e man-e
proper name-EZ I-be.3S
'She's my Ashraf (referring to her grandmother)'

Nikoo 1;10.15

Engelmann et al., (2019) proposed a more detailed classification of the errors to enhance the analysis of the language acquisition process. According to Engelmann et al., (2019) errors usually fall into one of three categories. The first one is frequently-based substitutions (i.e., when a low-frequent target form is replaced with a higher-frequency form of the same word. Near-misses or one-feature errors are those that differ from the target form by only one feature (Leonard et al., 2002) are the second category. The third category is when the inflectional ending is correct regarding to the number as and person but is from a different class.

Summary

In this chapter, the methodology for the study was outlined and explained in detail. The participants and their relevant characteristics were also described. The process of transcribing and coding the data, as well as the method of analysis, were discussed. Additionally, the linguistic measures that were used in the study were also presented and explained.

4. Amin

Introduction

Amin is the third child of a family living in Tehran. He has two sisters aged over 10. At the time of the data collection his mother was working from home and was his primary caregiver. Amin's receptive and productive CDI scores were 311 and 36 respectively at the time of the pre-session at the age of 1;6.17. A total of nine sessions were recorded from the age of 1;6.24 to 1;11.5.

Amin's lexical development

Amin's lexical measures are presented in Figure 5 and Table 11. Amin was already at the 5-word point when the recording started, producing five spontaneous words in one session. As shown in Table 11, Amin produced eight identifiable word types (five spontaneously) and 37 tokens during his first recording session. By 1;7.25, Amin has almost reached the 25-word point. As we can see here, roughly a month later, Amin produced his first combinations, at 1;9. By 1;9.23 he had produced 57 different combined words/phrases so his use of combinations had gone up rapidly. The last recording session (1;11.5) clearly shows a decrease in production, probably because Amin turned the TV on for a while and spent some time watching it in silence before his mother turned it back off. The first point where his lemma count and his word types start to differ is at 1;9, which marks the onset of his inflectional morphology development and the onset of his two-word stage. Session five marks the two-word utterance phase, where he produced eight word combinations. This figure reveals a sharp increase in the number of word combinations produced at 1;9.23. After that, the number of word combinations vary less from session to session.

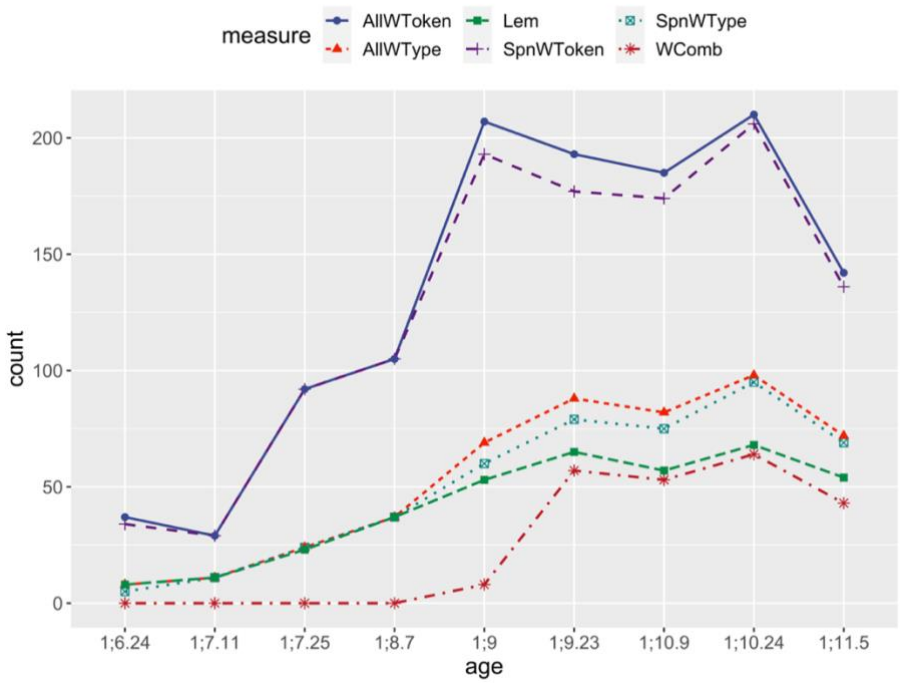

Figure 5 Amin's lexical development

Table 11 Amin's lexical development

session number	1	2	3	4	5	6	7	8	9
age	1;6.24	1;7.11	1;7.25	1;8.7	1;9.0	1;9.23	1;10.9	1;10.24	1;11.5
session length (min)	40	40	41	42	41	41	41	40	40
SpnWToken	34	29	92	105	193	177	174	206	136
SpnWType	5	11	24	37	60	79	75	95	69
AllWToken	37	29	92	105	207	193	185	210	142
AllWType	8	11	24	37	69	88	82	98	72
lemmas	8	11	23	37	53	65	57	68	54
word combinations	0	0	0	0	8	57	53	64	43

Amin's first recorded words

Table 12 lists the words identifiable in Amin's first session. Amin's spontaneous words are mostly accurate, whereas two of his imitated forms, [av:] and [uvev] for *ambulans* 'ambulance' and *utubus* 'bus', respectively, are less so. Amin's words are mostly one syllable long, with one disyllabic word.

Table 12 Amin's first recorded words

Target Word	Gloss	Child forms
/ɑb/	water	[oh]
/ɑmbulɑns/	ambulance	[av:] IM
/ɑx/	ouch	[ɑ:x], [ɑ::], [a:ʔ]
/eh/	oh	[eh], [e:], [ɑh], [ɑ:h], [ɑ:], [oh], [a:h]
/ku/	where	[gu] IM
/tup/	ball	[tub], [du::b]
/utubus/	bus	[uvev] IM
/ʔe/	oh	[ʔe]

IM = imitated and modelled words

Amin's consonant inventory

This section gives an overview of Amin's consonant inventory (CI) size at 5-, 25-, 50, and 100wps. Table 13 presents Amin's consonants at 1;6.24. As can be seen, Amin had very few consonants at the 5-word point. He produced five stops /b, t, d, g, ʔ/ and three fricatives /v, x, h/. He produced the labiodental fricative /v/ only in imitated words and only as a replacement for a target word.

Table 13 Amins CI at 1;6.24

	labio-labial/ labio-dental	apico-alevolar/ apico-dental	dorso-post-alveolar	dorso-palatal	dorso-prevelar	dorso-post-velar	dorso-uvular	glottal
stop	**b**	t d						ʔ
affricate								
fricative	**[v]**					**x**		h
nasal								
liquid								
glide								

The consonants produced at least twice, either in onset or coda positions, are marked in **bold face**.

The consonant produced in imitated words and as a replacement for a target word is inside brackets [ ].

Amin produced 14 consonants at the 25-word point at 1;7.25 (Table 14). Of the eight Persian stops, he produced six. The stops /k, ɢ/ had not yet appeared in his speech. The nasal /m/

occurred in both onset and coda, /n/ only in coda position. The fricative /x/ was produced in both onset and coda positions and / ʃ/ was produced only once, in coda position. Amin also produced the liquid /l/ (both positions) and the glide /j/ (only onset position). Amin produced six consonants at least twice in either onset or coda positions.

Table 14 Amin's CI at 1;7.25

	labio-labial/ labio-dental	apico-alevolar/ apico-dental	dorso-post-alveolar	dorso-palatal	dorso-prevelar	dorso-post-velar	dorso-uvular	glottal
stop	p **b**	**t** d			**g**			**ʔ**
affricate			**ʤ**					
fricative			ʃ			x		**h**
nasal	**m**	n						
liquid		l						
glide				j				

The consonants produced at least twice, either in onset or coda positions, are marked in **bold face**.

At 1;9, his consonant inventory had grown larger (Table 15). He produced 18 consonants out of the total number of 23 consonants found in adult Persian, 16 at least twice in either onset or coda position.

Table 15 Amin's CI at 1;9

	labio-labial/ labio-dental	apico-alevolar/ apico-dental	dorso-post-alveolar	dorso-palatal	dorso-prevelar	dorso-post-velar	dorso-uvular	glottal
stop	**p** **b**	**t** **d**			**k** **g**			**ʔ**
affricate			**ʧ** **ʤ**					
fricative	**f**	**s**	**ʃ**			**x**		**h**
nasal	**m**	**n**						
liquid		l						
glide				**j**				

The consonants produced at least twice, either in onset or coda positions, are marked in **bold face**.

Table 16 provides an overview of Amin's consonant inventory at 1;10.24 (100-word point). He produced 19 consonants, 18 at least twice in one of the onset/coda positions. At this stage the only missing consonants are the affricates /ʧ, ʤ/, which Amin had produced earlier, and the fricative /ʒ/, which infants produce much later in development (Jalilevand, 2011).

Table 16 Amin's CI at 1;10.24

	labio-labial/ labio-dental	apico-alevolar/ apico-dental	dorso-post-alveolar	dorso-palatal	dorso-prevelar	dorso-post-velar	dorso-uvular	glottal
stop	**p b**	**t d**			**k g**		**ɢ**	ʔ
affricate								
fricative	**f v**	**s z**	ʃ			**x**		**h**
nasal	**m**	**n**						
liquid		**r**						
glide				**j**				

The consonants produced at least twice, either in onset or coda positions, are marked in **bold face**.

Amin's syllable structure

Figure 6 shows Amin's most used syllable structures for all sessions. These syllable structures together account for more than 50% of the total number of productions in each session. His top syllable structure at the beginning of data collection was VC. At 1;7.11 he mostly preferred producing monosyllabic structures, but we also see an emergent use of disyllabic CVC:V. The figure shows that Amin's second most used syllable structure is the disyllabic structure CVC(:)VC (with or without gemination of the second consonant) at 1;9. His first use of verbal endings also falls at this 1;9.

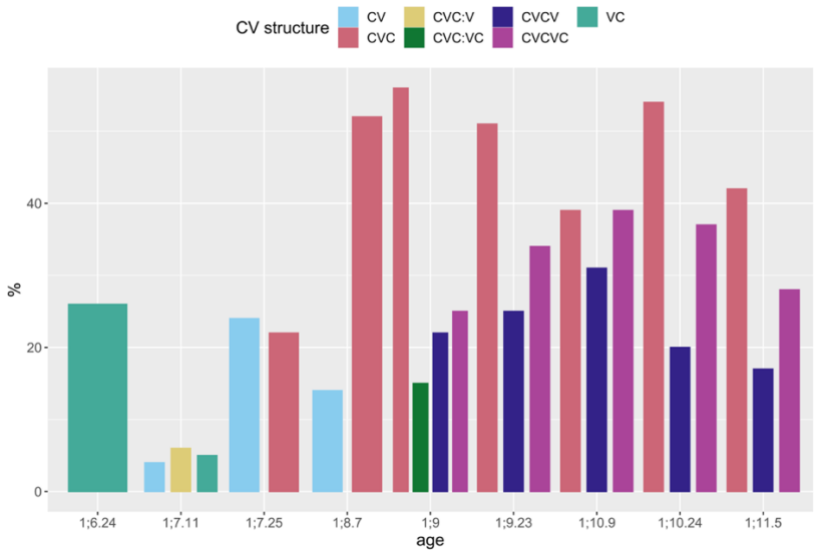

Figure 6 Amin's top CV structures

Amin's MLUw, MLUm, and pMLU scores can be seen in Figure 7. He produced only single-word utterances in the first four sessions, which is reflected in the MLUw graph (top left). At 1;9 a change in his MLUw score was observed, as he started producing word combinations. Session six (1;9.23) shows a more dramatic increase in the number of longer utterances (MLUw) and words having more morphemes (MLUm). From 1;9.23 onwards the increase seems to be more stable. Amin's pMLU keeps a steady increase throughout the data collection period, with a score of 2.8 in the first session and a score of 6.33 in the last session.

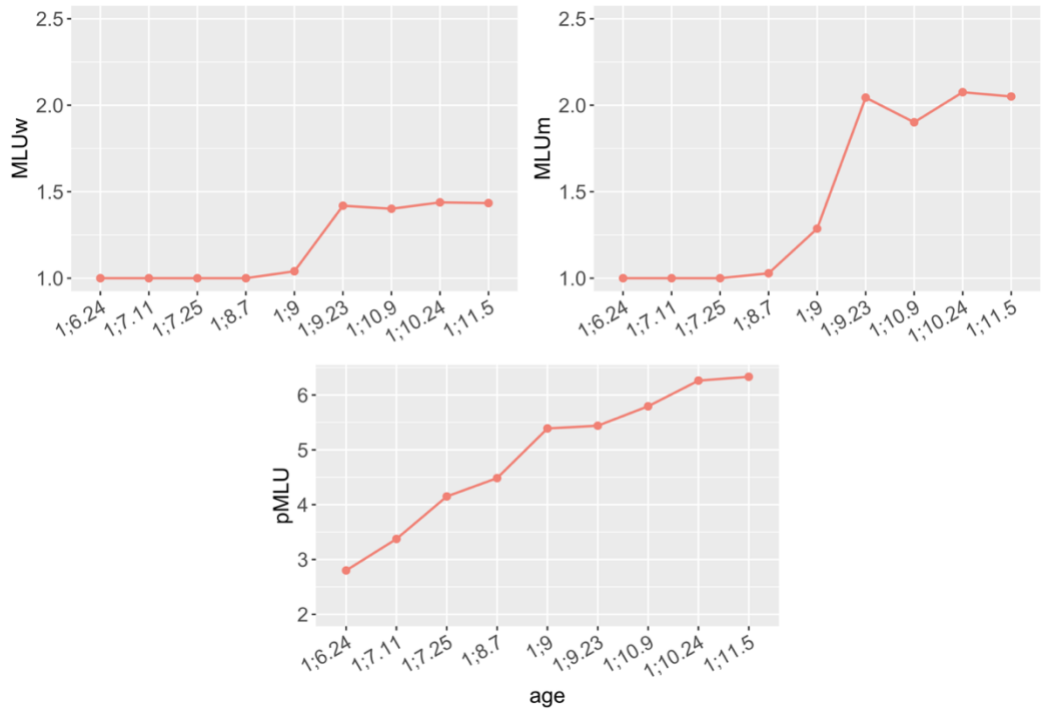

Figure 7 Amin's MLU scores

Table 17 presents Amin's first recorded combinations along with their target adult form, gloss, and the combination type. His first combinations are of three main types: 1. N + N, 2. 1S-pronoun + verb-1S, and 3. (possessed) noun-EZ + (possessor) 1S pronoun-verb.1S. The first type (N + N) does not occur in adult Persian, but the second two types are adult-like.

Table 17 Amin's first combinations

Child form	Adult target	Gloss	Combination type
[babaji bɑst]	/babaji lebɑs/	sheep clothes	noun + noun
[babaji nas:]	/babaji lebɑs/	sheep clothes	noun + noun
[man nista:m]	/man nist-am/	I NEG.be-1S	1S pronoun + verb-1S
[man hastam]	/man hast-am/	I be-1S	1S pronoun + verb-1S
[man aga:m]	/man hast-am/	I be-1S	1S pronoun + verb-1S
[man gam:ɑm]	/man hast-am/	I be-1S	1S pronoun + verb-1S

[man dada:m]	/man raft-am/	I went-1S	1S pronoun + verb-1S
[mam:ani manam]	/maman-e man-e/	mum-EZ I-be.3S	(possessed) noun-EZ + (possessor) 1S pronoun-verb.1S

Amin's phonological development

Figure 8 presents an overview of Amin's CVar, PWC, PCC-R and PWP scores over the course of data collection. There is a rising trend in his CVar and PWP scores. Amin's CVar score saw the most increase. His initial score of 1.43 increased to the high of 2.55 at 1;10.24. Amin's PCC-R started at 0.65. It seems that he is the least accurate at 1;7.11. At 1;7.25 there is a massive increase in this score. The score seems to decrease slightly at 1;9 (50wp), when the first phrases and inflectional morphemes appear in his speech. He starts with a score of 0.42, which increases to the high of 0.79 in Session 7 (1;10.19). There are two main points at which we observe a decrease in PWP. The first coincides with his first attempts at producing longer utterances and his first attempt at producing inflectional morphology (age 1;9). The second decrease falls in the last two sessions: at 1;10.24 and 1;11.5. As can be seen, PWC score peaks at 1;7.25 (25wp) and again at 1;10.24 (100wp).

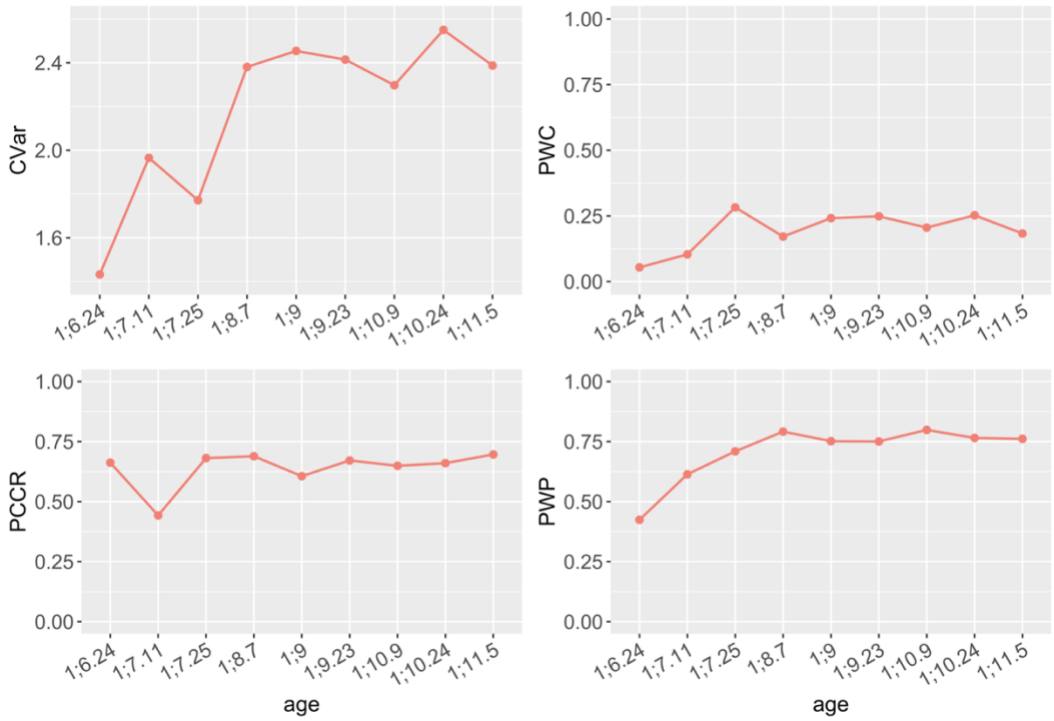

Figure 8 Amin's phonological development

Amin's morphological development

Amin's PDSS score

It is important to stress that Amin produced one question word [gu] for *ku* 'where' in his first session, which resulted in a PDSS score of one. Since this score can give a false sense of Amin's developmental status, this session will be excluded from further PDSS analysis.

The next two sessions included no word/morpheme that would count towards the PDSS scores. At 1;8.7 Amin produced three different verb types, all 3S simple past. This form carries no pronominal suffixes to mark person and number on the verb (as seen earlier in Table 10). Only one of these verbs includes a prefix, namely, the NEG morpheme /na-, ne-/. The verb forms are as follows: [not] for /mord/ 'died.3S', [daf] /raf(t)/ 'went.3S', and [nis] /nis(t)/ 'NEG.be.3S'. The DEF morpheme was also first seen in this session: [do:de] for /dozd-e/ 'thief-DEF'. This results in a score of 1.05 at 1;8.7, which also marks the end of his one-word stage. Session five (1;9) marks Amin's first productive use of inflectional morphemes.

Table 18 and Figure 9 present Amin's PDSS information. As can be seen here, there is a

steady rise in Amin's mean PDSS score. His initial score was 1.05. The highest score of 3.95 has been recorded at 1;10.24 which also marked his 100wp.

Table 18 Amins PDSS score

Session number	Session length	age	PDSS_mean	PDSS_sd	PDSS_min	PDSS_max	PDSS_total
1	40	1;6.24	1	0	1	1	1
2	40	1;7.11	0	0	0	0	0
3	41	1;7.25	0	0	0	0	0
4	42	1;8.7	1.05	0.24	1	2	19
5	41	1;9.0	1.99	1.19	1	6	147
6	41	1;9.23	3.31	1.57	1	7	281
7	41	1;10.9	3.94	1.82	1	9	256
8	40	1;10.24	3.95	2.48	1	11	336
9	40	1;11.5	3.79	1.84	1	8	212

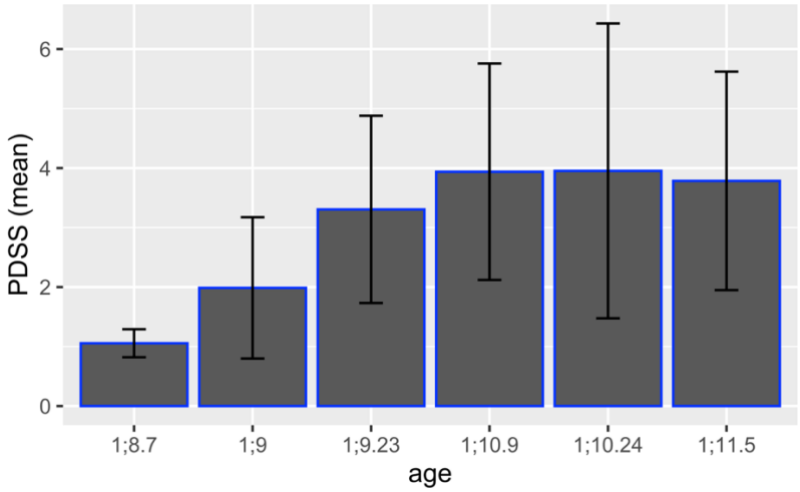

Figure 9 Amin's PDSS score

Figure 11 presents PDSS score for each of the analysed sub-categories. Verbal morphology shows a steady increase overall. Amin started producing compound verbs from 1;9 (25wp). His use of more advanced word structures is significantly higher in the last two sessions. Prepositions and conjunctions do not seem to show a steady rise. His score for this subcategory peaks at 1;7.11 and 1;11.5 (0.28), with the lowest recorded at 1;10.24. Amin's use of pronouns peaks at 1;10.9 and then declines slightly once again. He rarely used question

words. His score for the question words subcategory peaks at 1;10.24 (0.02). This subcategory score seems remarkably lower than the other subcategory scores. The inflectional morpheme category shows a steady increase with the lowest score recorded for 1;8.7 (0.22) and the highest score recorded for 1;10.24 (1.28). His initial score of 0.08 rises to 0.72 on 1;10.9 and then slightly decreases in the following two sessions. His sentence structure score follows the same trend.

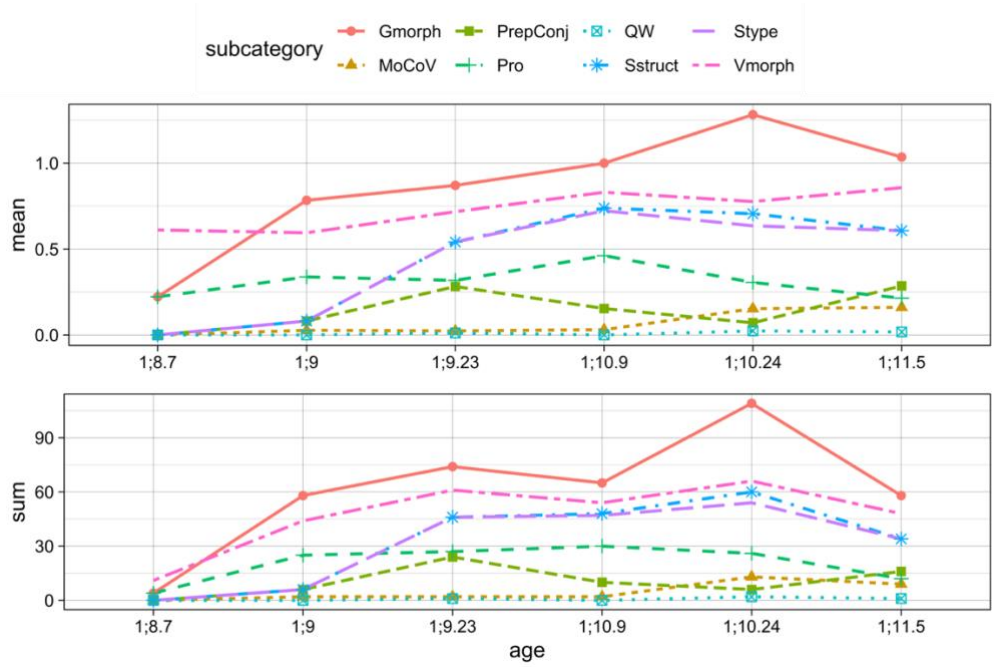

Figure 10 Amin's sub-categorical PDSS scores

Amin's morphemes

Amin's correctly used morphemes can be seen in Figure 11.

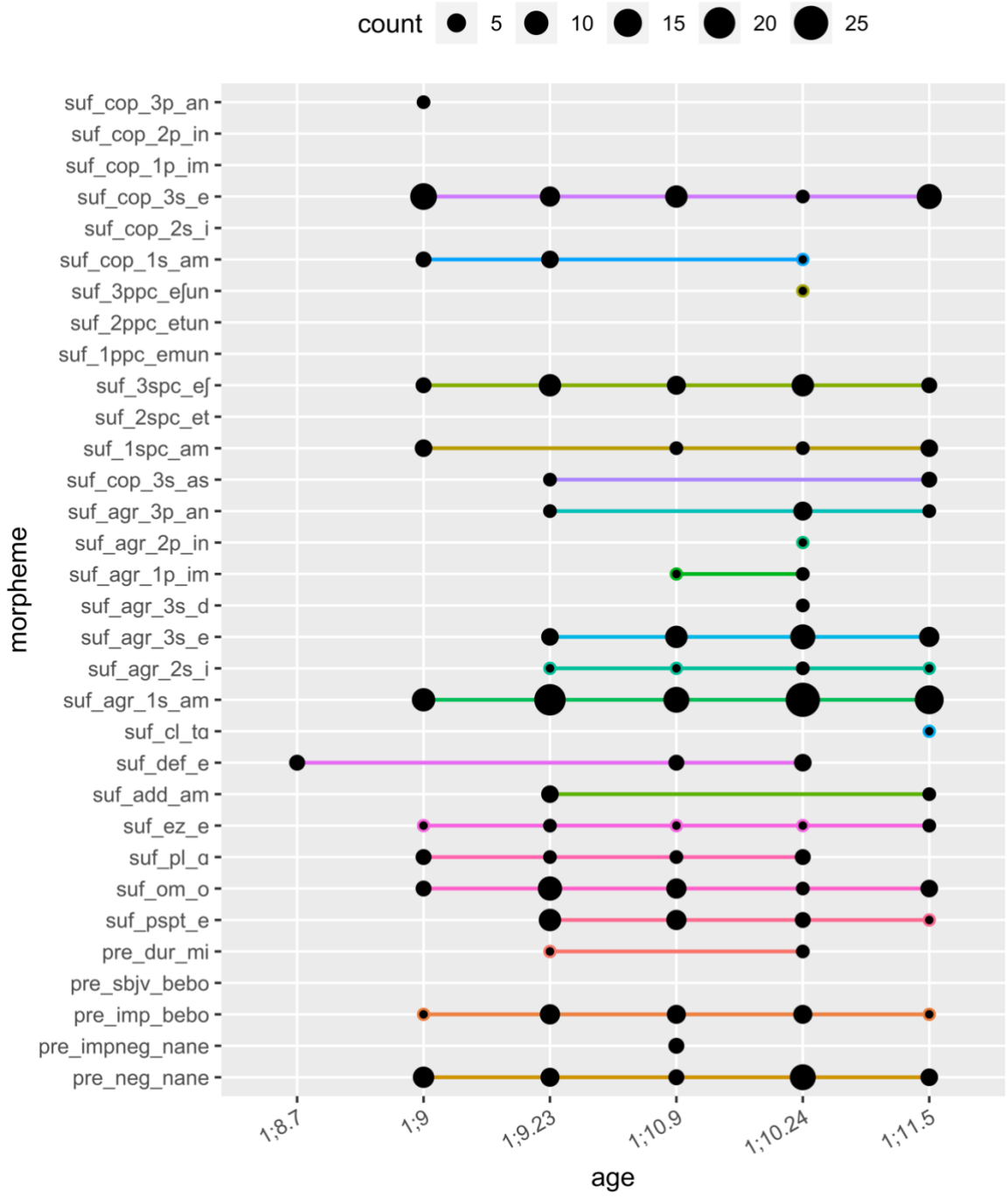

Figure 11 Amin's correctly used morphemes

As can be seen in Figure 11, Amin had no inflectional morphemes in the first three sessions. At 1;8.7 he produced his first inflectional morpheme. The first inflectional morpheme that appeared in his speech was DEF /-e/, on the noun /dozd/ as in *dozd-e* 'thief-DEF'. However, his use of this morpheme was not productive. He did not use this noun without this morpheme attached to it or with any other morphemes.

At age 1;9, many morphemes first appeared in Amin's speech. NEG /na-,ne/ appeared seven times but only with one lemma. He showed contrastive use of this prefix: [nist-am] 'NEG-be-1S' vs. [hast-am] 'be-1S'.

Only one case of IMP /be-,be-/ was recorded. OM /o-/ was also recorded attached to one stem: [in:-o] for /in-o/ 'this-OM'. He contrastively used PL /-ɑ/ in [in:-ɑ] for /in-ɑ/ 'this-PL'. EZ particle also appeared first at 1;9 in the utterance below:

(4.1)	CF:	mɑm:ɑn-i	mɑn-ɑm
		mum-EZ	I-be.1S
	AT:	mɑmɑn-e	mɑn-e
		mum-EZ	I-be.3S
		'That's my mother'	

He used pronominal suffix 1S /-am/ with four different verbs, for example:

(4.2)	CF:	dad-a:m	vs.	gald-am
	AT:	raft-am	vs.	kand-am
		went-1S		plucked-1S
		'I went'		'I plucked (it)'

As can be seen in 4.3, the copula 1S /-am/ also appeared with two different stems. He also produced one of the stems in its bare form.

(4.3)	CF:	man-am	vs.	gam:-ɑm
	AT:	man-am	vs.	xɑb-am
		I-be.1S		asleep-be.3S
		'It's me'		'I'm asleep'

1SPC /-am/ was also appeared with two different stems marking the possession: [ba-m] for /pɑm/ 'leg-POSS.1SPC' and [mɑm:ɑn-ɑm] for /mɑmɑn-ɑm/ 'mum-POSS.1SPC'. 3SPC /-eʃ/ was recorded with only one stem. Two imitated/modelled cases of 3SPC were recorded as well. Finally, the morpheme that Amin used most at this age was copula 3S /-e/. It attached to seven different word types, for example in:

(4.4)	CF:	gag:uʧ-e	vs.	ɑbu-e:
	AT:	xarguʃ-e	vs.	hɑpu-e
		rabbit-be.3S		doggie-be.3S
		'It's a rabbit'		'It's a doggie'

Both of these lemmas also appeared uninflected in Amin's speech.

At 1;9.23, Amin used six new inflectional morphemes. Prefix DUR /mi-/ appeared attached to one verb [mi-d-am] 'DUR-give-1S' while being omitted in other cases. Amin used PSPT.3S /-e/ with only one verb: [zad-e] 'hit- PSPT.3S'. He used ADD /-am/ in four cases, of which only one was spontaneous, the rest imitated or modelled. Amin produced one case of verbal agreement 2S '-i'. Verbal agreement 3S /-e/ appeared attached to two different verbs:

(4.5)	CF:	dɑj-e	vs.	nah-gɑj-e
	AT:	dɑr-e	vs.	na-xor-e
		have-3S		NEG-eat-3S
		'He/she/it has (it)'		'He/she/it doesn't eat (it)'

Amin had two cases of copula 3S /-as/ in one word type. First productive cases of NEG /na-,ne-/ were seen at this age:

(4.6)	CF:	nah-xoj-e	vs.	na-dar-am
	AT:	na-xor-e	vs.	na-dar-am
		NEG-eat-3S		NEG-have-1S
		'(so that) he/she/it doesn't eat'		'I don't have (it)'

IMP /-be,bo-/ was recorded with two different verbs, although his use of this prefix does not show any evidence of being used productively yet. Amin produced OM /-o/ with two new word types in the following two utterances:

(4.7)	CF:	maman-:o	dus	
	AT:	mɑmɑn-o	dus(t)	(dɑr-am)
		mum-OM	love	(have-1S)

'I love mum'

(4.8) CF: dudu-jɑ-jo gɑh
AT: ʤuʤu-hɑ-ro negɑh (kon)
birdie-PL-OM look (do.2s)
'Look at the birds!'

Amin produced PL /-ɑ/ with two different noun types, one already presented above. [dudu] for /ʤuʤu/ 'birdie' appeared both in two inflected forms and in uninflected forms, suggesting that PL has become productive at this stage:

(4.9) CF: dudu vs. dudu-jɑ-jo vs. ʤuʤuje
AT: ʤuʤu vs. ʤuʤu-hɑ-ro vs. ʤuʤue
Birdie vs. birdie-PL-OM vs. birdie-be.3S

Two more examples of EZ were recorded in the following utterances:

(4.10) CF: dɑ-jeh dudu uma
AT: sedɑ-je ʤuʤu umad
sound-EZ birdie came.3S
'There was the birdie's sound'

(4.11) CF: guʤi-e jɑst
AT: kuʃe-je rɑst
alley-EZ right
'the right-hand side alley'

Since Amin produced the uninflected form of the noun /kuʃe/ as well, it can be concluded that EZ has become productive.

At the age of 1;10.9 only three new morphemes were recorded. Perhaps the most important of all was IMPNEG /na-,ne/. This prefix was attached to three different verb types:

(4.12)	CF:	nah#[14]keʃ	vs.	nah#kon	vs.	nah#goj
	AT:	na-keʃ	vs.	na-kon	vs.	na-xor
		IMPNEG-draw.2S		IMPNEG-do.2S		IMPNEG-eat.2S

These verbs were also recorded without this prefix at the same age. Examples of these verbs are presented below in (4.13), (4.14), and (4.15):

(4.13)	CF:	babaji-je	raf	baba	xoj-e
	AT:	babaji-je	raf(t)	baba	bo-xor-e
		sheep-DEF	went.3S	food	SBJV-eat-3S
		'The sheep went to eat food'			

(4.14)	CF:	gɑdeh	kon
	AT:	gerje	kon
		cry	do.2S
		'Cry!'	

(4.15)	CF:	gɑji	keʃ
	AT:	xɑle	be-keʃ
		aunt	IMP-draw.2S
		'Draw a picture of my aunt!'	

For the first time Amin produced the same noun with and without DEF /-e/ at 1;10.9: [babaji] 'sheep' and [babaji-je] for /babaji-e/ 'sheep-DEF'. He also used DEF /-e/ with three different nouns.

At 1;10.24, there were two instances of verbal agreement 1P /-im/: [dɑd-im] 'gave-1P' and [kad-im] for /kard-im/ 'did-1P'. [dɑd-im] 'gave-1P' was produced in contrast to [dɑd-am] 'gave-1S'. [kad-im] was produced in contrast to [mih-kad-am] for /mi-kard-am/ 'DUR-did-1S'. Amin's contrastive use of 1P suggests that this morpheme was used productively.

His contrastive use of DEF /-e/ continued:

[14] Hashtag (#) represents a short pause.

(4.16)	CF:	hɑpu	vs.	hɑpu-ʔe	and	mɑn:et	vs.	mɑnet-e
	AT:	hɑpu	vs.	hɑpu-ʔe		vɑnet	vs.	vɑnet-e
		doggie		doggie-DEF		pickup truck		pickup truck-DEF

Amin only produced verbal agreement 2S /-i/ with one verb type: [nah#dɑj-i] for /na-dɑr-i/ 'NEG-have-2S'. Verbal agreement 3S /-e/ appeared specifically with the three verb types: [dɑre] for /dɑr-e/ 'have-3S', [xoj-e] for /bo-xor-e/ 'SBJV-eat-3S', and [doj-e] for /kon-e/ '(SBJV)do-3S'. As seen above, the verb *dɑʃtan* was produced for both 2S and 3S. The verb *xordan* was recorded with 1S, 3S, 1P, and 2P verbal agreements.

(4.17)	CF:	xor-am	vs.	xoj-e
	AT:	bo-xor-am	vs.	bo-xor-e
		SBJV-eat-1S		SBJV-eat-3S

(4.18)	CF:	nah-gor-i:d [15]	vs.	nah-xord-an
	AT:	na-xor-id	vs.	na-xor-an
		IMPNEG-eat-2P		NEG-eat-3P

The first appearance of verbal agreement 3S /-d/ was recorded twice with one verb type: [jɑ-t] and [mi-jɑ-t] both for /mi-jɑ-d/ 'DUR-come-3S'. 3PPC /-eʃun/ was also first emerged at this age in only one instance: [xuna-ʃun] 'house-POSS.3PPC'.

At the age of 1;11.5, the only case of CL /-tɑ/ first emerged in utterance below:

(4.19)	CF:	doh-tɑ	daj	dɑj-am
	AT:	dotɑ	daftar	dɑr-am
		two-CL	notebook	have-1S
		'I have two notebooks'		

Copula 3S /-as/ started being used with different lemmas for the first time at this age:

[15] /-id/ is the formal form of 2P /-in/

(4.20)	CF:	muʤa-s	vs.	hab:ɑmɑ-s
	AT:	murʃa-s	vs.	havɑpejmɑ-s
		ant-be.3S		airplane-be.3S

Amin's inflectional errors

Looking at Figure 12, it is apparent that Amin did not have inflectional morphemes up until 1;8.7 hence zero errors. At 1;8.7, as mentioned in the previous section, he only had one non-productive morpheme.

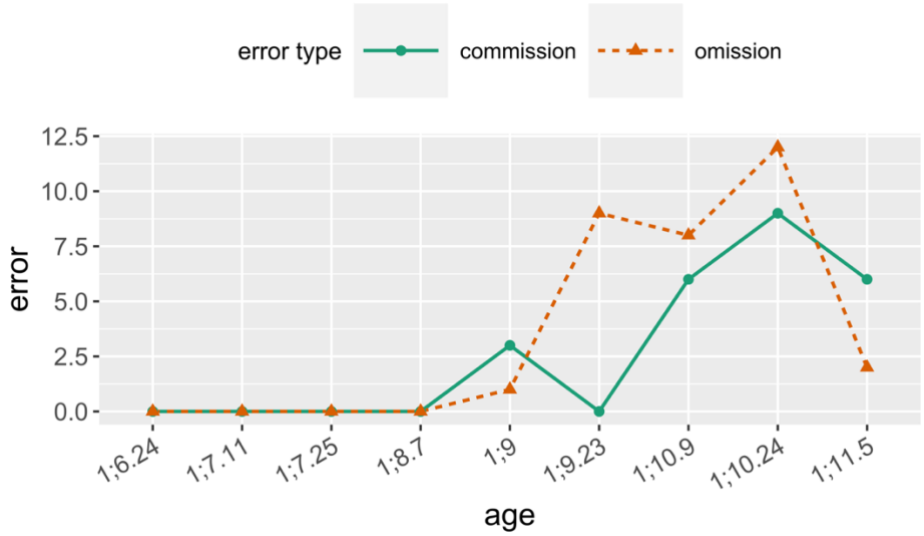

Figure 12 Amin's omission vs commission errors

Amin's erroneous use of morphemes can be seen in Figure 13. The first time he produced a variety of inflectional morphemes was at 1;9. Three out of four errors he made at this age were errors of commission. Two of his errors were agreement errors. For instance, in the following utterance the copula is in agreement with *man* (first person singular), instead of *maman* (third person singular).

(4.21)	CF:	mɑm:ɑn-i	mɑn-ɑm
		mum-EZ	I-be.1S
	AT:	mɑmɑn-e	man-e
		mum-EZ	I-be.3S
		'It's my mum'	

The second instance was wrong agreement between /to/ 'you' and 2S /-i/ as in:

(4.22)	CF:		to-ji:
			you-be.2S
	AT:	(mal-e)	to-e
		(possession-EZ)	you-be.3S
		'It belongs to you!'	

It can be concluded from these two errors that Amin was forming the person-number relation between the free morphemes and their bound counterparts within the copula paradigm. The other commission error was using copula 3S /-as/ instead of copula 3S /-e/. At 1;9.23 Amin made no commission errors and nine omission errors. He had four prefix type errors and two suffix type errors.

The prefix errors were as follows. The IMP /be-,bo-/ omissions were seen in: [gi:r] for /be-gir/ 'IMP-take.2S' and [bin] for /be-bin/ 'IMP-look.2S'. The SBJV /be-,bo-/ omission errors were recorded in [xorim] for [bo-xor-im] 'SBJV-eat-1P'. Finally, the DUR /mi-/ omissions were seen in [goj-am] for /mi-xor-am/ 'DUR-eat-1S' and [xoj-e] for /mi-xor-e/ 'DUR-eat-3S'.

Regarding his suffix errors, there was one case of agreement 1S /-am/ in [dɑj] for /dɑr-am/ 'have-1S' and one missing copula 1S /-am/ in the following utterance:

(4.23)	CF:	man	hɑpu
	AT:	man	hɑpu-am
		I	doggie-be.1S
		'I'm a dog'	

At 1;10.9, prefix omission errors continued. Missing IMP /be-,bo-/ was recorded in: [deh] for /be-de/ 'IMP-give.2s', [gij-as] for /be-gir-eʃ/ 'IMP-take.2S-3SPC (= take it)', and [keʃ] for /be-keʃ/ 'IMP-draw.2S'. There was one case of missing SBJV /be-,bo-/. There was also missing DUR /mi-/ with two verb types: [bar-e] for /mi-bar-e/ 'DUR-take-3S' and [ded-am] for /mi-keʃ-am/ 'DUR-draw-1S'. One case of wrong copula 3S (/-e/ instead of /-as/) was recorded in

[zad-as][16] for /zard-e/ 'yellow-be-3S'.

One commission error of OM /-o/ occurred while Amin was playing with a toy gun with his mother. His mother pretended to shoot him. Amin then pretended to fall on the floor and produced the following utterance below:

(4.24)	CF:	mano	mord-om
		I-OM	died-1S
	AT 1:	man	mord-am
		I	died-1S
		'I died'	
	AT 2:	man-o	koʃt
		'I-OM	killed.3S'
		'he/she/it killed me'	

He used the intransitive verb *mordan* 'to die' and attached the object marker to the subject of the sentence. It is unclear whether he used the only verb he knew that was semantically related to his intended meaning or he actually meant to use the verb *mordan* and that he simply made an error with attaching the object marker to the subject. It appears that there is an overgeneralization of either the object marker or the verb *mordan*.

There was one EZ particle omission error in:

(4.25)	CF:	daɢad-ɑ	man-ɑ
		pencil-PL	I-be.3S
	AT:	medɑd-ɑ-je	man-e
		pencil-PL-EZ	I-be.3S
		'These are my pencils'	

There was also one case of missing verbal agreement 3P /-an/ in the below utterance:

(4.26)	CF:	in-ɑh	tup-a	bud

[16] Not to be mistaken with /zard-as/ (zard-e-as) 'yellow-DEF.be.3S' which translates to 'it's the yellow one (as opposed to one with a different colour)'

	this-PL	ball-PL	was.3S
AT 1:	in-ɑ	tup	bud-an
	this-PL	ball	was.3P
AT 2:		tup-ɑ	bud-an
		ball-PL	was.3P'

At 1;10.24 Amin's both omission and commission errors peak. The majority of them are OM /-o/ and the verbal prefixes DUR /mi-/ and SBJV /be-,bo-/.

Amin had four cases of missing OM /-o/ all required being attached to pronouns /man/ 'I' and /mɑ/ 'we' as in the utterances below:

(4.27)	CF:	mɑnet:-ɑh	mɑh	nah-xord-an	
	AT:	vɑnet-ɑ	mɑ-ro	na-xord-an	
		pickup truck-PL	we-OM	NEG-eat-3P	
		'The pickup trucks didn't eat us'			
(4.28)	CF:	hɑpuʔ-e	mi-jɑ-t	man	xoj-e
	AT:	hɑpu-e	mi-jɑ-d	man-o	bo-xoj-e
		doggie-DEF	DUR-come-3S	I-OM	SBJV-eat-3S
		'The doggie is coming to eat me'			

Amin produced [nah#dɑj-i] 'NEG-have-2S' for /na-dɑr-e/ 'NEG-have-3S'. He continued using copula 1S /-am/ when attached to the pronoun /man/ 'I':

(4.29)	CF:	bɑnuneh	man-ah
		breakfast	I-be.1S
	AT:	sobune(-je)	man-e
		breakfast(-EZ)	I-be.3S

At 1;11.5 there was a sharp drop in both types of errors. Two cases of 3S /-e/ instead of 1S agreement /-am/ were observed.

(4.30)	CF:	man	daj-o	bɑz	kon-e
		I	door-OM	open	do-3S
	AT:	man	dar-o	bɑz	kon-am
		I	door-OM	open	do-1S'
		'I'd open the door'			

There was one case of verbal agreement 2S /-i/ instead of 1S /-am/:

(4.31)	CF:	kaftaji	tɑj-i
		pigeon	have-2S
	AT:	kaftar	dɑr-am
		pigeon	have-1S
		'I have a pigeon/pigeons'	

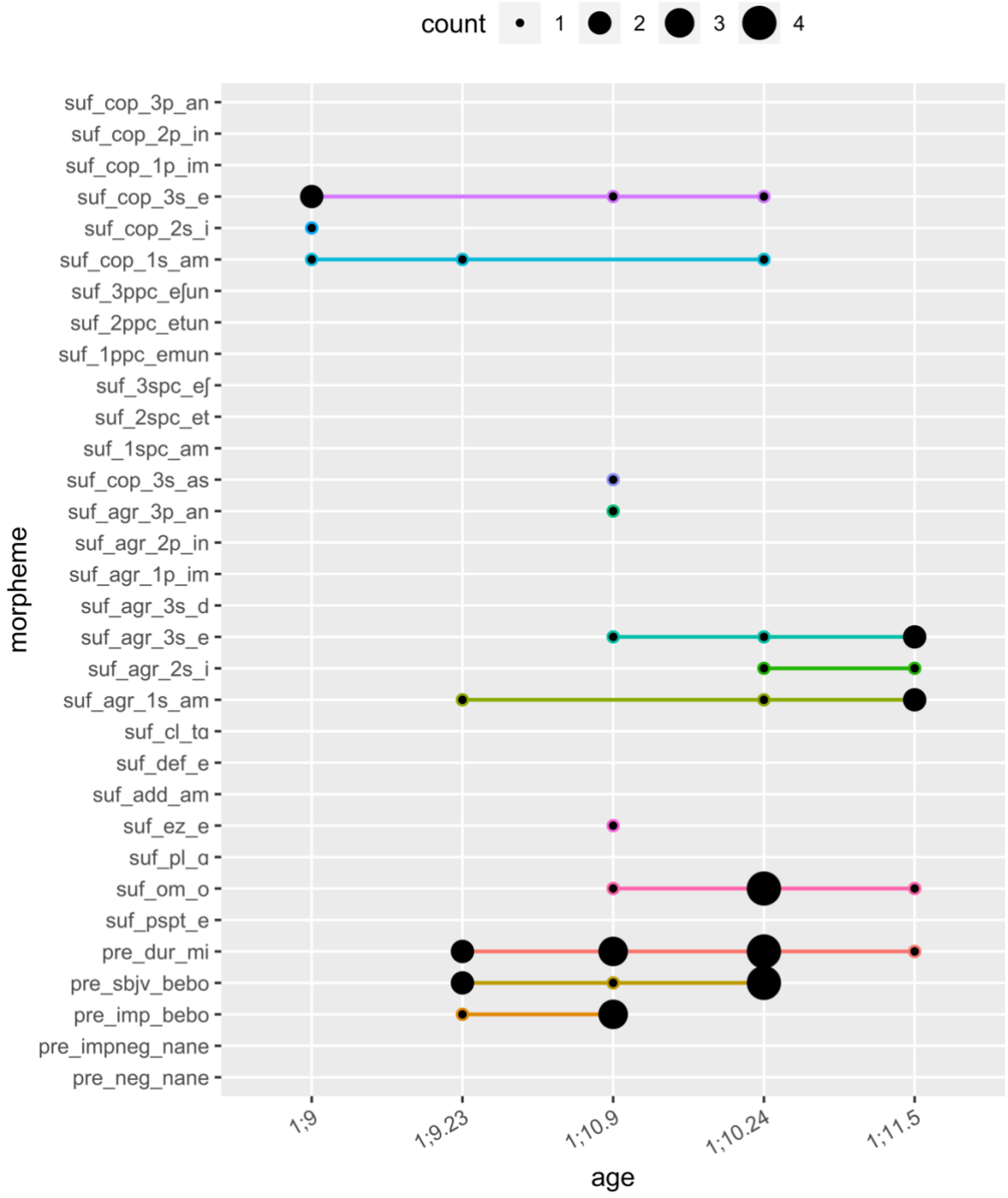

Figure 13 Amin's erroneous use of morphemes

First apprearance vs. contrastive use

To conclude this chapter, we look at Amin's inflectional morphemes when they first appeared versus when he first used those morphemes contrastively. Figure 14 shows most of Amin's inflectional morphemes became partially productive with a delay after they first appeared. Although some first appeared at the same age they were used contrastively. It is important to

note that this cannot be taken as evidence for productive use of these morphemes from the first time they were recorded.

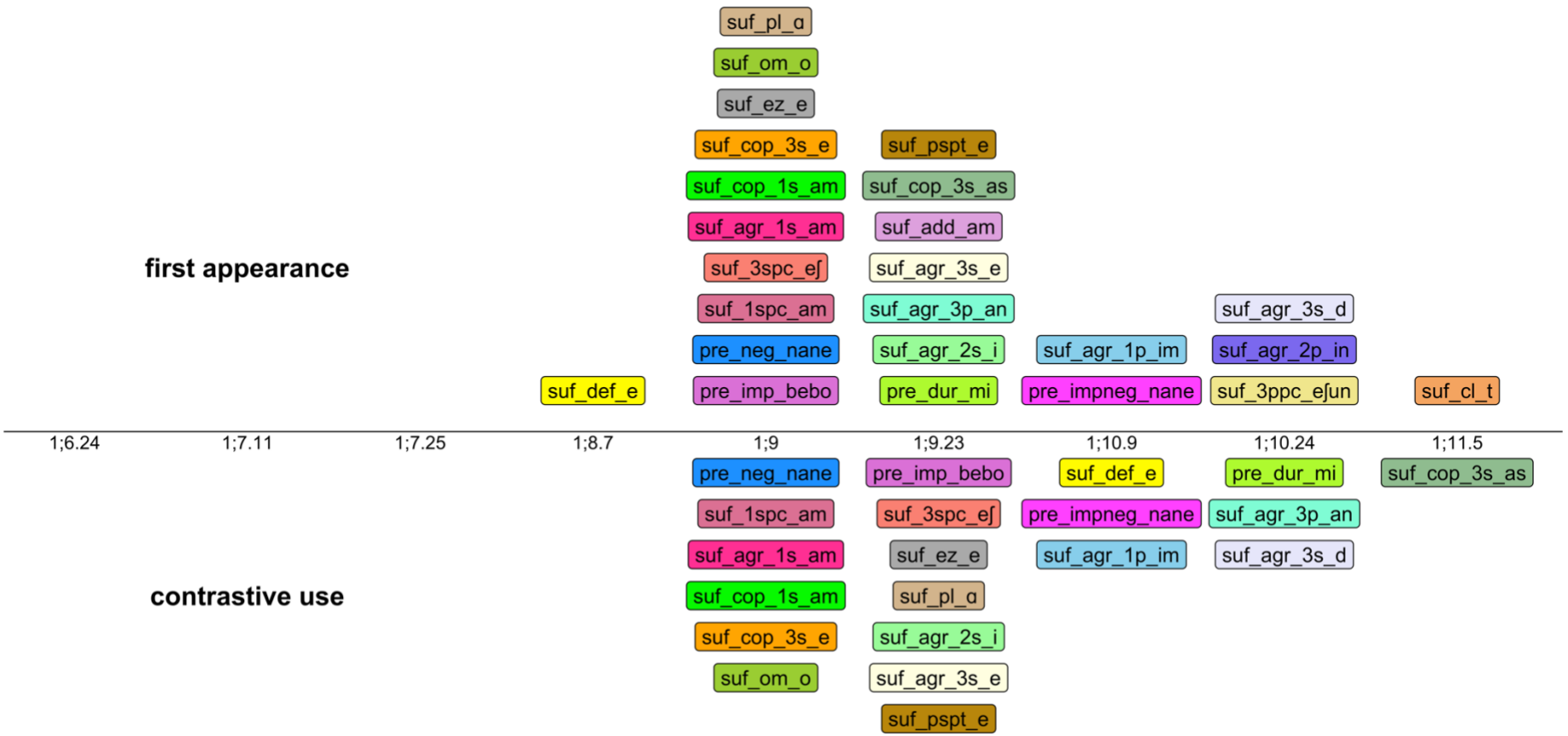

Figure 14 Amin's morphemes: first appearance vs. contrastive use

Summary

In this chapter, the linguistic development of Amin was examined. Amin was recorded from the age of 1;6.24 to 1;11.5. He did not produce any word combinations until 1;9.0. One of the most notable observations in Amin's speech was the use of the first-person singular pronoun, despite the fact that Persian is a pro-drop language, and the use of pronouns is not necessary. This was an intriguing discovery and may indicate the influence of frequency in the input on language acquisition, which is an area that warrants further exploration in future research. It is possible that the high frequency of the first-person singular pronoun in the input that Amin received may have contributed to his early and consistent use of this pronoun. This is an important area of study as it will give us a better understanding of how children's language development is shaped by the input they receive and the ways in which frequency of certain forms and structures may impact their acquisition of language. Additionally, examples of overgeneralization were observed in Amin's speech, as seen in example (4.25) on page 92. This phenomenon, in which a child applies a rule or pattern to a broader set of words or situations than is appropriate, is a common phenomenon seen in other languages as well. Overall, this chapter provides insight into the unique language development patterns of Amin and highlights the importance of considering individual differences in children's language acquisition.

5. Saman

Introduction

Saman is a second child of a monolingual family living in Tehran. He has an older sister. His mother was not working at the time and was the primary caregiver. His initial CDI scores were 157 and 23 for receptive and productive, respectively. His CDI scores were recorded at the age 1;7.0. Six sessions were recorded in a period of almost six months.

Saman's lexical development

Figure 15 and Table 19 depict Saman's lexical development. From the data in Table 19, it is apparent that Amin was already passed the 25-word point at the beginning of the data collection. He had already started combining words at 1;7.15. His word combination number rises steadily up to 1;9.28. There is then a massive increase in his number of word combinations.

Table 19 Saman's lexical development

session number	1	2	3	4	5	6
age	1;7.15	1;7.29	1;9.3	1;9.28	1;10.25	1;11.12
session length (min)	50	41	45	42	43	40
SpnWToken	82	87	142	132	326	291
SpnWType	31	30	34	33	60	69
AllWToken	82	91	143	140	331	299
AllWType	31	31	35	39	62	72
lemmas	23	28	33	36	49	57
word combinations	1	1	5	11	76	65

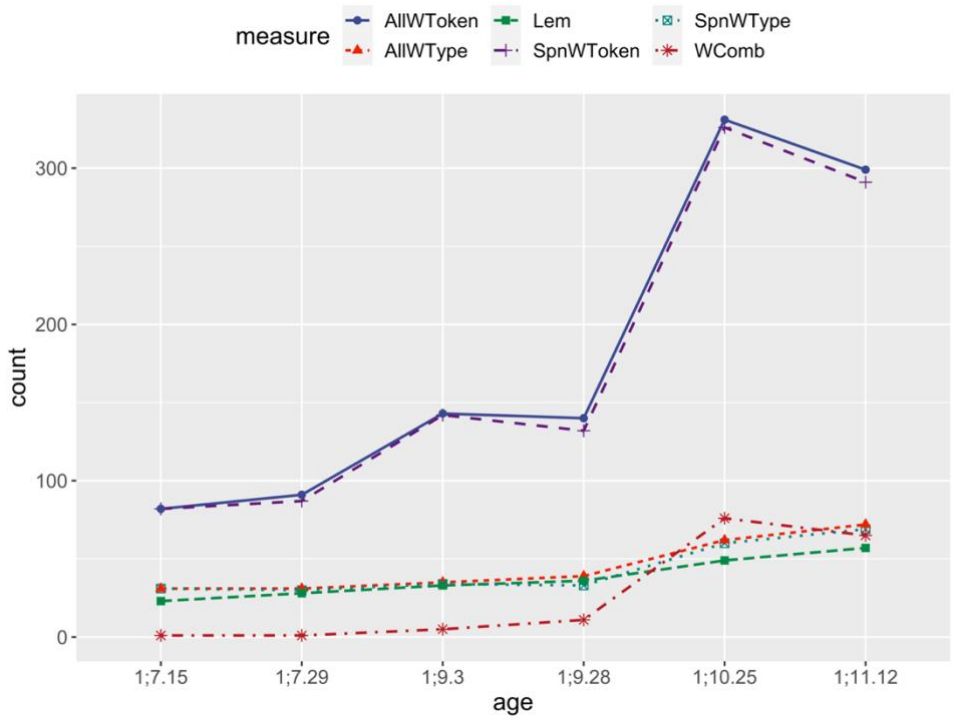

Figure 15 Saman's lexical development

Saman's first recorded words

As can be seen in Table 20, Saman already produced copula to be 3S /-e/ with different words. As stated earlier he also had one word combination.

Table 20 Saman's first recorded words

Target Word	Gloss	Child forms
/ah/	oh	[ajih], [e:h], [eh], [oh]
/ɑbi/	blue	[ɑbi]
/ɑbije/	blue-be.3S	[abije], [abije:], [ɑb:ijeh], [ɑbije], [eb:ije:h]
/ɑj/	ouch	[ɑhj]
/ɑm:ije/	food-be.3S	[ɑm:je]
/ɑx/	ouch	[o:x], [ox]
/ɑzitɑ/	proper name	[adidɑ], [adidɑ:j], [ɑ:dede], [ɑdid:ɑ], [ɑdidɑ], [ɑdijɑ:]
/bɑbɑ/	dad	[bɑb:ɑ], [bɑv:a]
/bɑbɑʤi/	grandpa	[bɑb:ɑʤ]
/bɑbɑe/	dad-be.3S	[bɑbi:je]

/dadar/	outside	[da:daʔ]
/dadare/	outside-be.3S	[dadadi:], [dadɑ:ʔe]
/dar/	door	[dadu]
/do/	two	[do:]
/ɢatkard/	hung up.3S	[iɢat#ha]
/hamum/	bath	[am:i]
/hamume/	bath-be.3S	[hamue:]
/ino bede/	this-OM IMP-give.2S	[ino bedej]
/mɑmɑn/	mum	[mɑm:ɑ], [mɑm:ɑn], [meɑm:ɑn], [umɑm:ɑn]
/mɑmɑni/	mum	[mɑm:ɑ:ni], [mɑm:ɑni], [mɑmɑni:]
/mɑmɑnie/	mum-be.3S	[mɑmɑne]
/na/	no	[n:a], [n:a:h], [na], [na:]
/nini/	baby	[n:ini]
/nis(t)/	NEG.be.3S	[niʃ], [nit], [nus]
/oftɑd/	fell.3S	[of]
/oftɑdam/	fell-1S	[hedɑ#am]
/omid/	proper name	[ɑm:i], [m:i], [mi], [mi:], [mi:s]
/omide/	proper name-be.3S	[om:ie], omi:je]
/vɑj/	oops	[u:j], [vijvij]
/ʔe/	oh	[e:], [ʔe]

Saman's consonant inventory

Table 21 presents Saman's consonant inventory at the beginning of the data collection, at the age of 1;7.15. He produced nine consonants at least twice in either the onset or the coda positions. His consonant inventory is missing the liquids and the dorso-prevelar stops.

Table 21 Samans CI at 1;7.15

	labio-labial/ labio-dental	apico-alevolar/ apico-dental	dorso-post-alveolar	dorso-palatal	dorso-prevelar	dorso-post-velar	dorso-uvular	glottal
stop	**b**	**t** **d**					ɢ	**ʔ**
affricate			ʤ					
fricative	f **v**	**s**	ʃ			x		**h**
nasal	**m**	**n**						
liquid								
glide				**j**				

The consonants produced at least twice, either in onset or coda positions, are marked in **bold face**.

At 1;9.28 the stop /p/ and the fricative /z/ appeared, as well as the dorso-prevelar /k, g/ and the liquid /l/. Saman produced 12 consonants at least in two work tokens in either onset or coda positions (see Table 22).

Table 22 Saman's CI at 1;9.28

	labio-labial/ labio-dental	apico-alevolar/ apico-dental	dorso-post-alveolar	dorso-palatal	dorso-prevelar	dorso-post-velar	dorso-uvular	glottal
stop	p **b**	**t** **d**			k g			**ʔ**
affricate			**ʧ** **ʤ**					
fricative	**f** **v**	s **z**				x		**h**
nasal	**m**	**n**						
liquid		**l**						
glide				**j**				

The consonants produced at least twice, either in onset or coda positions, are marked in **bold face**.

Table 23 shows Saman's consonant inventory at 1;11.12. He produced a total number of 19 consonants. He produced 14 consonants at least twice in either onset or coda positions.

Table 23 Saman's CI at 1;11.12

	labio-labial/ labio-dental	apico-alevolar/ apico-dental	dorso-post-alveolar	dorso-palatal	dorso-prevelar	dorso-post-velar	dorso-uvular	glottal
stop	**b**	**t** **d**			**g**		ɢ	**ʔ**
affricate			ʧ ʤ					
fricative	**v**	**s** **z**	**ʃ**			x		**h**
nasal	**m**	**n**						
liquid		**l** r						
glide				**j**				

The consonants produced at least twice, either in onset or coda positions, are marked in **bold face**.

Saman's syllable structures

Saman's top syllable structures are presented in Figure 16. These structures account for at least 50% of the total number of his production in each session. At the beginning of the data collection, at 1;7.15 his four top preferred structures were CV, VC, VCVCV, and CVCVCV. His use of inflectional morphemes from the beginning of the data collection is reflected in his three-syllable production, for example the use of copula 3S /-e/ in words like [omid-e] 'proper name-be.3S' (= it's Omid). At the age of 1;10.25 there is a rise in the use of VC. This coincides with Saman's highest MLUw score and the highest number of produced word combinations in a session. These combinations are structured as *in/un* + *noun* 'this/that + noun'.

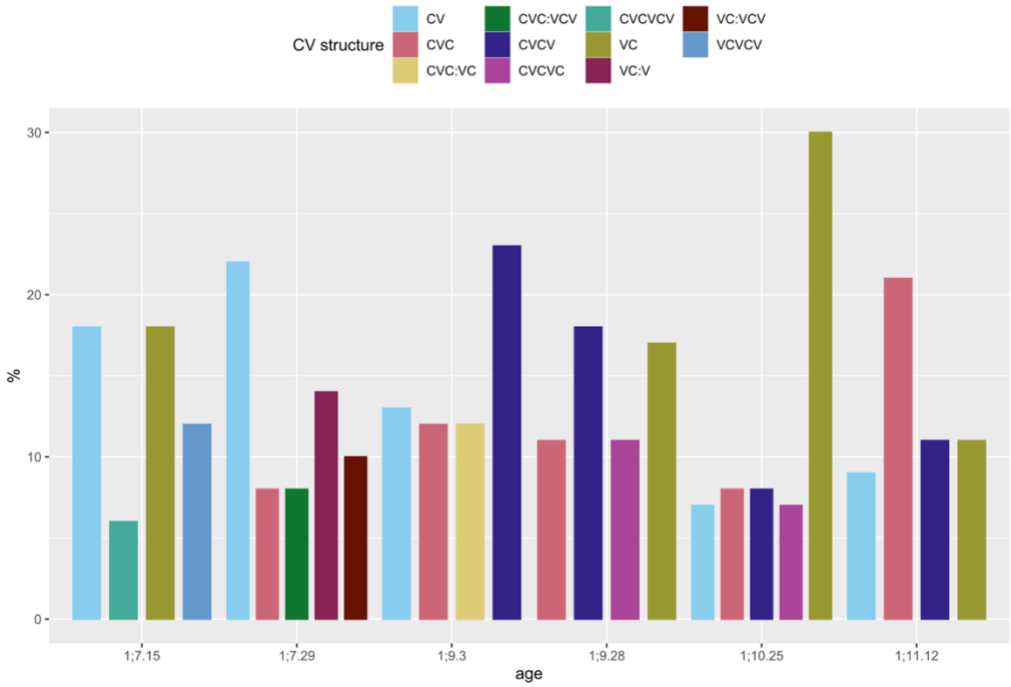

Figure 16 Saman's top syllable structures

Saman's MLU scores can be seen in Figure 17. Both his MLUw and MLUm show a general rising pattern. His pMLU score drops over the first four session with the highest recorded on the first session at 1;7.15. This high score might be due to his attempt at producing shorter words which resulted in producing more adult-like words. The decline in this score can be an indication of his attempt at producing longer words including inflectional morphemes. Saman's pMLU drops to the lowest of 4.35 at 1;9.28.

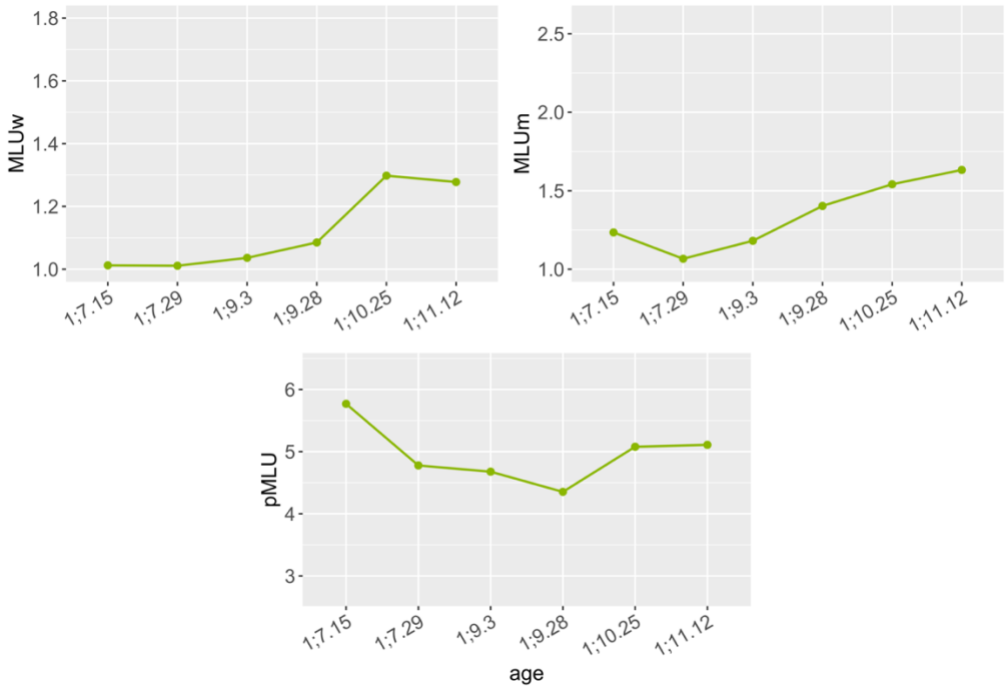

Figure 17 Saman's MLU scores

Saman's syntactic development

Saman's first word combinations can be found in Table 24. What is interesting in this table is Saman's surprisingly adult-like first combination recorded. The word combination [ino bedej] for /in-o be-de/ 'this-OM IMP-give.2S' (= Give (it) to me!) is too adult-like especially when compared to his later combinations at 1;9.3. His word combinations at 1;9.3 are the type of word combinations that can be expected to be produced productively in the child's telegraphic phase. This suggested that [ino bedej] has been rote-learnt by Saman as an unanalysed chunk.

Table 24 Saman's first combinations

Child form	Adult target (by word)	Gloss	Combination type	Age
[in-o be-dej]	ino bede	this-OM IMP-give.2S	object-OM + verb	1;7.15
[di d:a]	ino dar	this off	this + action	1;7.29
[bida na]	bita na	proper name no	noun + no	1;9.3
[baba aʰb]	baba ab	dad water	noun + noun	1;9.3
[alila ah]	ali ab	proper name water	noun + noun	1;9.3
[bida lal:a]	bita lala	proper name sleep	noun + noun	1;9.3
[ana nah]	ana raf(t)	proper name went.3S	noun + verb	1;9.3

Saman's phonological development

Saman's highest CVar score was recorded on the last session, at the age of 1;11.12 (Figure 18). This is also the same session with the highest recorded MLUm (Figure 17). Saman's highest CVar and MLU scores coincide with the highest number of produced word types (both SpnWType and AllWType) and lemmas.

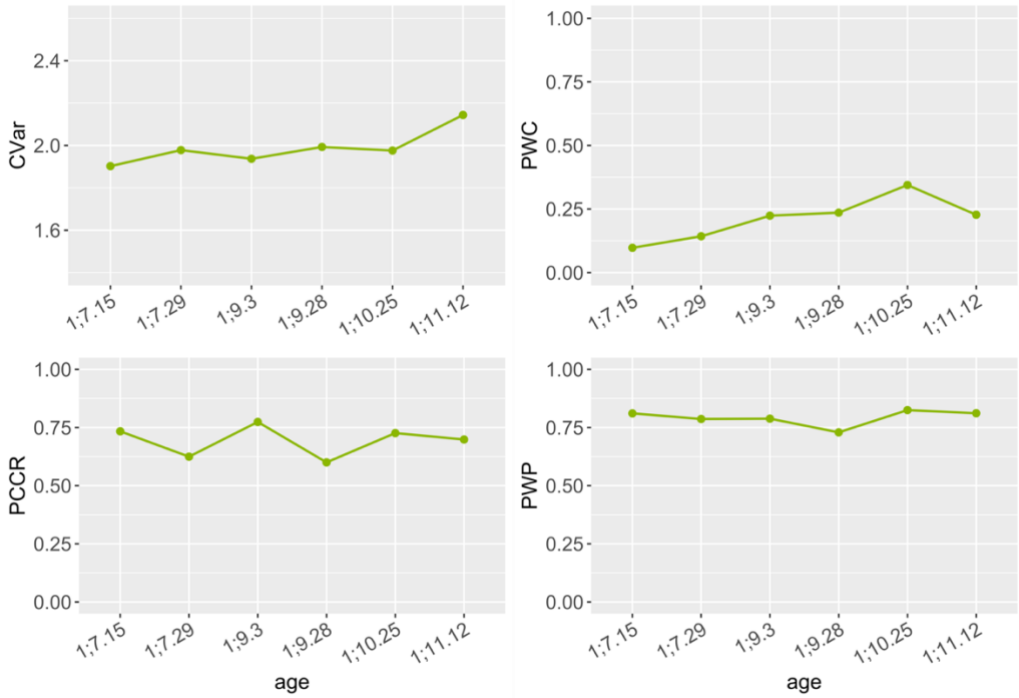

Figure 18 Saman's phonological development

Saman's morphological development

Saman's PDSS score

Saman's mean PDSS score, standard deviation, the minimum, and the maximum are presented in Table 25. His highest score was recorded at 1;11.12. Figure 19 shows Saman's mean PDSS score with the lowest score (1.33) at the age of 1;7.29 and the highest score at the age of 1;10.25 (3.1).

Table 25 Saman's PDSS score

Session number	Session length	Age	PDSS_mean	PDSS_sd	PDSS_min	PDSS_max	PDSS_total
1	50	1;7.15	2.15	0.93	1	6	43
2	41	1;7.29	1.33	0.49	1	2	20
3	45	1;9.3	1.81	0.56	1	3	49
4	42	1;9.28	2.78	1.72	1	6	125
5	43	1;10.25	3.1	2.11	1	6	322
6	40	1;11.12	2.76	1.54	1	8	323

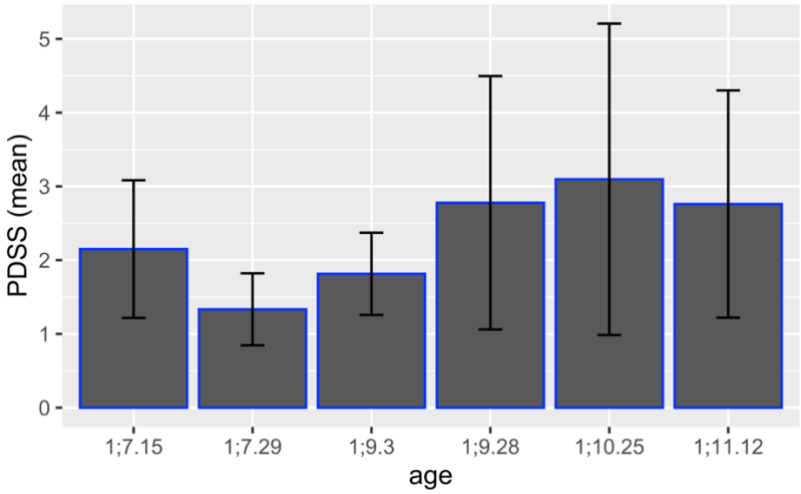

Figure 19 Saman's PDSS score

Figure 20 shows the PDSS sub-categorical mean and total scores. The figure reveals that the biggest increase was in the Gmorph (sum) and Vmorph (sum) categories.

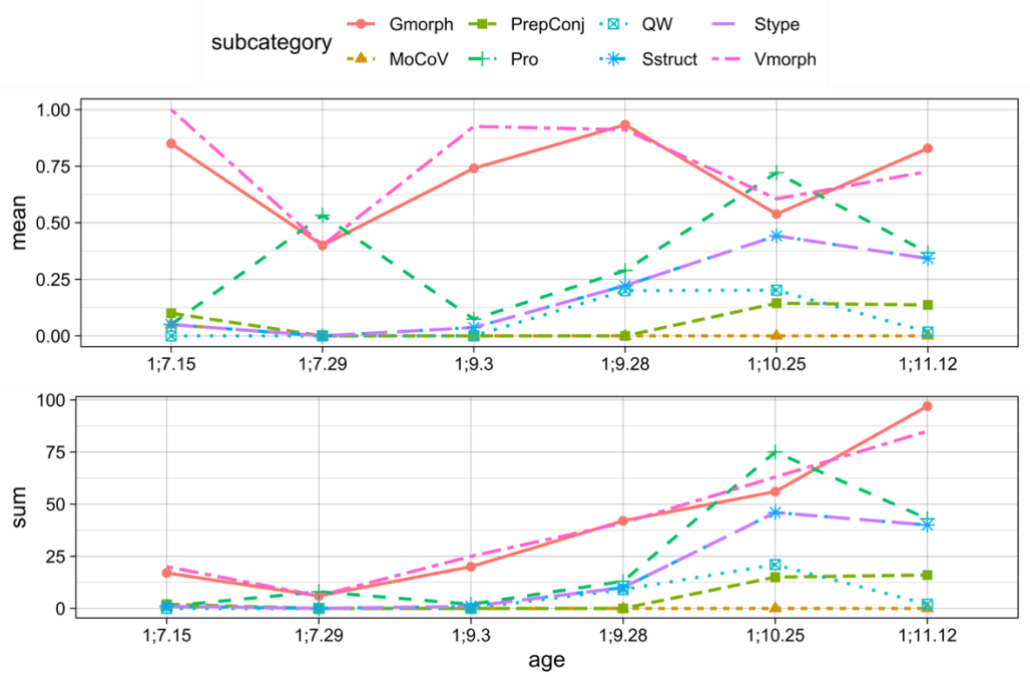

Figure 20 Saman's sub-categorical PDSS scores

Saman's morphemes

Saman's use of morphemes is presented in Figure 21. He productively used copula 3S /-e/ with 8 different lemmas at 1;7.15. Some of these lemmas were also used without this morpheme. One example is presented in the following:

(5.1)	CF:	ɑbi	vs.	abi-je
	AT:	ɑbi	vs.	ɑbi-je
		blue		blue-be.3S

He only had one case of verbal agreement 1S /-am/: [hedɑ#am] for /oftɑdam/ 'fell-1S'. Saman also produced IMP prefix /be-,bo-/ and OM /-o/ once in the following utterance:

(5.2)	CF:	in-o	be-dej
	AT:	in-o	be-de
		this-OM	IMP-give.2S
		'Give this to me!'	

NEG /na-,ne-/ was only used in what can be regarded as the base form: [nit] for /nis(t)/ 'NEG-be.3S'.

At 1;7.29 only two new morphemes emerged: copula 1S /-am/ and plural /-ɑ/ in [man-am] 'I-be.3S' and [in:ɑ] 'this-PL', respectively. [in] 'this' was recorded without any morphemes for the first time at this stage. This means that he used [in-o] 'this-OM' and [in:ɑ] 'this-PL' contrastively to [in] 'this' in the same month.

At 1;9 copula 3S /-ɑs/ appeared attached to one stem. Verbal agreement 1S /-am/ was also recorded once attached to a verb. Saman produced the same verb type six times without it:

(5.3)	CF:	be-bi:	vs.	be-b:i-a
	AT:	be-bin-am		be-bin-am
		SBJV-look-1s		SBJV-look-1s

SBJV /be-,bo-/ appeared with only one verb type, as can be seen in the above example. His first use of IMP /be-,bo-/ was recorded in [be-b:in] 'IMP-look.2S' and [bijɑ] 'IMP-come.2S'. The first appearance of PSPT 3S /-e/ was also seen in [id-e] for /raft-e/ 'went-PSPT.3S'.

At age 1;9.28, no new morphemes emerged. NEG /na-,ne-/ was again only used in the verb base form. Verbal agreement 1S /-am/ continued to only appear with the same word type as recorded at 1;9.

At 1;10.25, Saman produced the first recorded verbal agreement 1S /-e/ and the first recorded case of DUR /mi-/. These can be seen in the following utterance which he produced as a single word:

(5.4)	CF:	a-m:i-j-e	
	AT:	ɑb	mi-r-e
		water	DUR-go-3S
		'The water runs'	

He also attached SBJV /be-,bo-/ to two new verbs. These are presented in the following:

(5.5)	CF:	be-band-e	vs.	be-h
		SBJV-close-3S		SBJV-go
	AT:	be-band-am	vs.	be-ɾ-e
		SBJV-close-1S		SBJV-go-3S

NEG /na-,ne-/ still does not show any evidence of being productively produced. Plural /-ɑ/ and OM /-o/ were only recorded with [in] 'this'.

At 1;11.12, still no productive use of NEG /na-,ne-/ was recorded. Saman produced the first case of plural /-ɑ/ with a different stem produced in [bibib-ɑ] 'car-PL'. He also used the same lemma without this morpheme and with a different morpheme: [bibib] 'car' vs. [bib:ib-e] 'car-be.3S'. He also produced one more case of PSPT 3S /-e/ in [uman-e] for /umad-e/ 'came-PSPT.3s'. Finally, the first case of 3SPC /-eʃ/ appeared attached only to one stem.

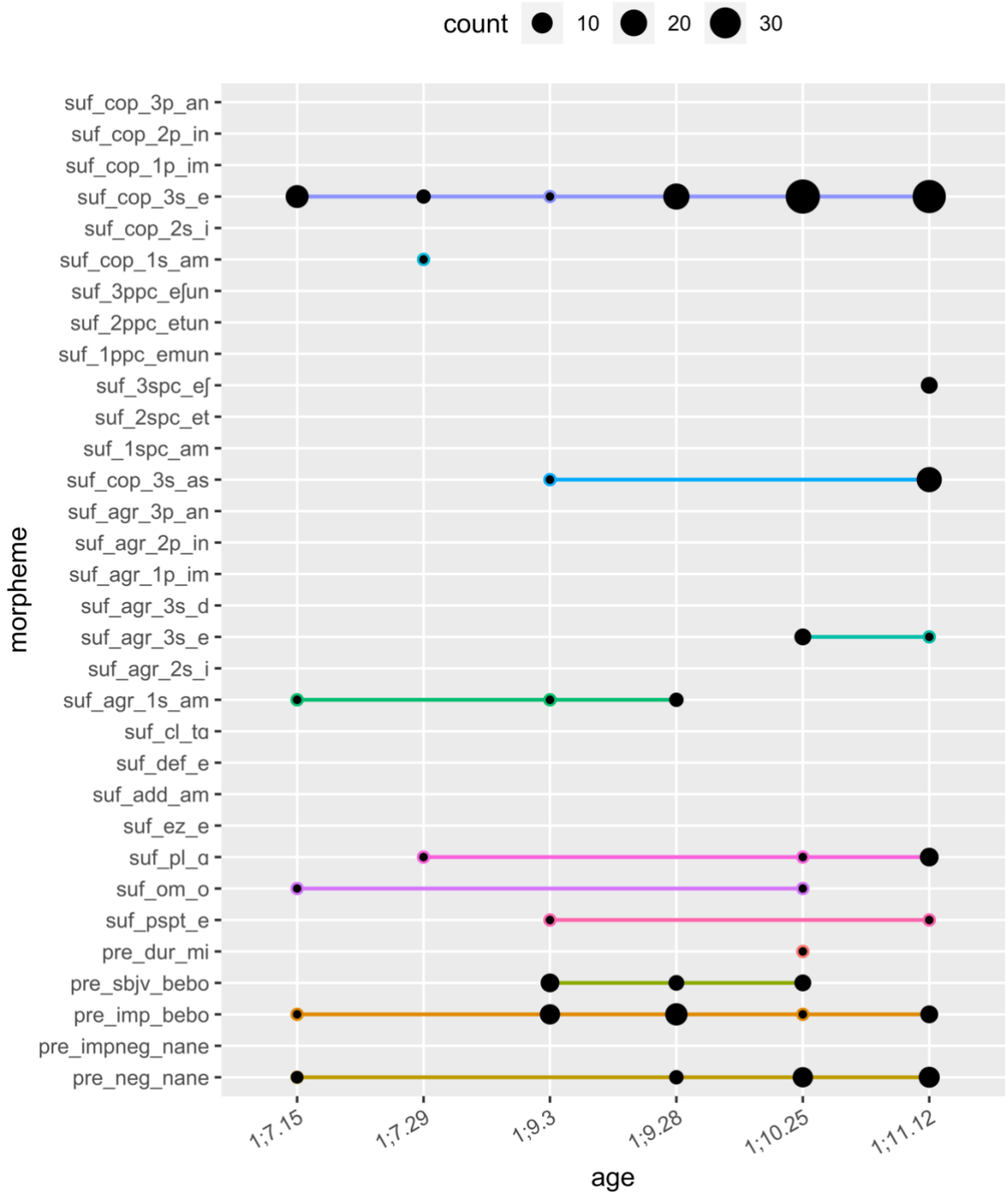

Figure 21 Saman's correctly used morphemes

Saman's inflectional errors

Saman made more omission errors than commission errors in general during the period of data collection (Figure 22). Amin's erroneous morphemes are presented in Figure 23. His errors first emerged at 1;9.3. His error was limited to omission of verbal agreement 1S /-am/ on the verb *didan* 'to see', for instance: [be-be] for /be-bin-am/ 'SBJV-see-1S'. At 1;9.28 one

more case of 1S /-am/ was recorded. He omitted copula 3S /-e/ twice, for instance in the following utterance:

(5.6) CF: in di
AT: in ʧi-je
this what-be.3S
'What is this?'

At 1;10.25 Saman used verbal agreement 3S, instead of 1S in [be-band-e] 'SBJV-close-3S' for /be-band-am/ 'SBJV-close-1S'. There was one case of erroneous use of copula 3S /-as/, instead of copula 3S /-e/ in:

(5.7) CF: in ʧi-s
AT: in ʧi-je
this what-be.3S
'What is this?'

At 1;11.12 Saman used copula 3S /-as/ instead of copula 1P /-an/ in:

(5.8) CF: bib:ib-ɑ indɑ-s
car-PL here-be.3S
AT: bibib-ɑ indʒɑ-n
car-PL here-be.3P
'The cars are here'

Two omission errors of copula 3S /-as/ were recorded: [indɑ] for /indʒɑ-s/ 'here-be.3S'. Finally, Saman produced verbal agreement 2S, instead of 3S: [be-d-i] 'SBJV-give-2S' for /be-d-e/ 'SBJV-give-3S'.

Figure 22 Saman's commission vs omission errors

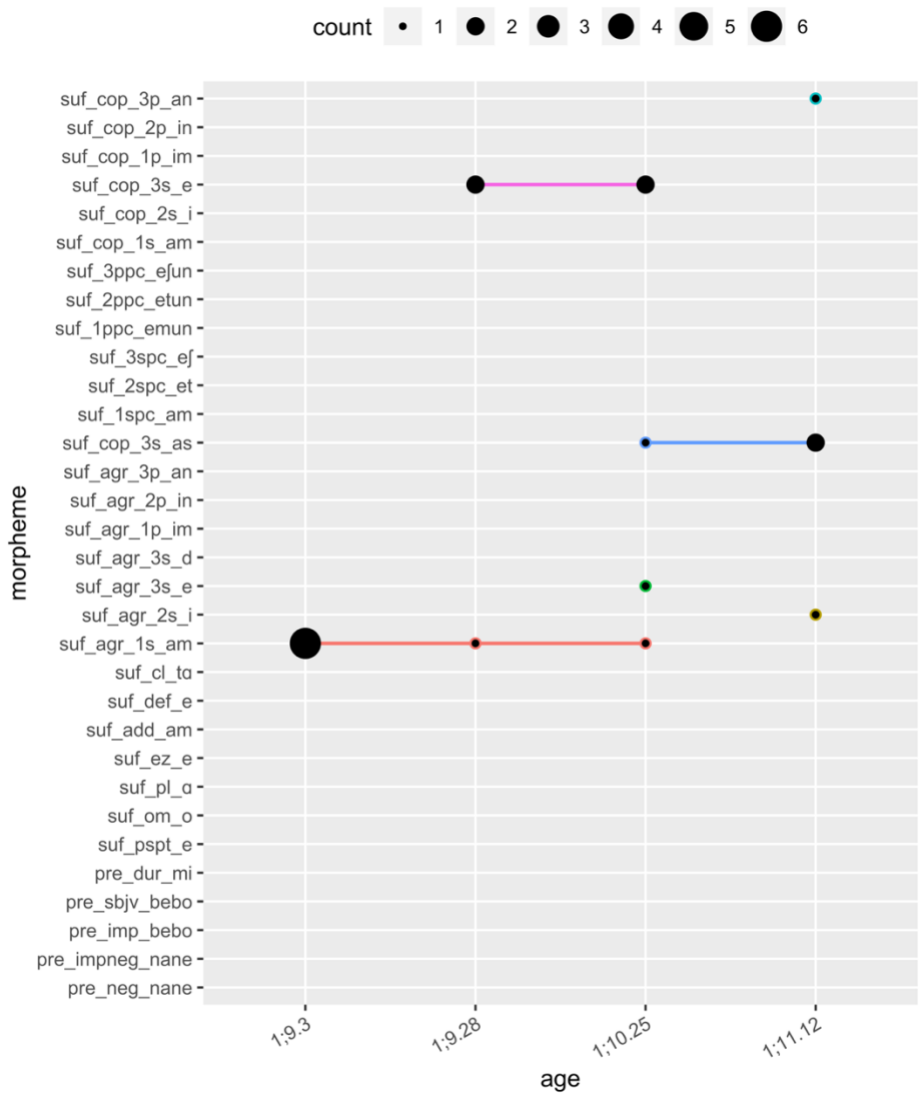

Figure 23 Saman's erroneous use of morphemes

First appearance vs. contrastive use

In order to conclude this chapter, the first appearance of his inflectional morphemes and the first contrastive use is reviewed here (Figure 24). During the period of the data collection, there were only five inflectional morphemes used contrastively by Saman. He used two of these contrastively at the same age they were first recorded.

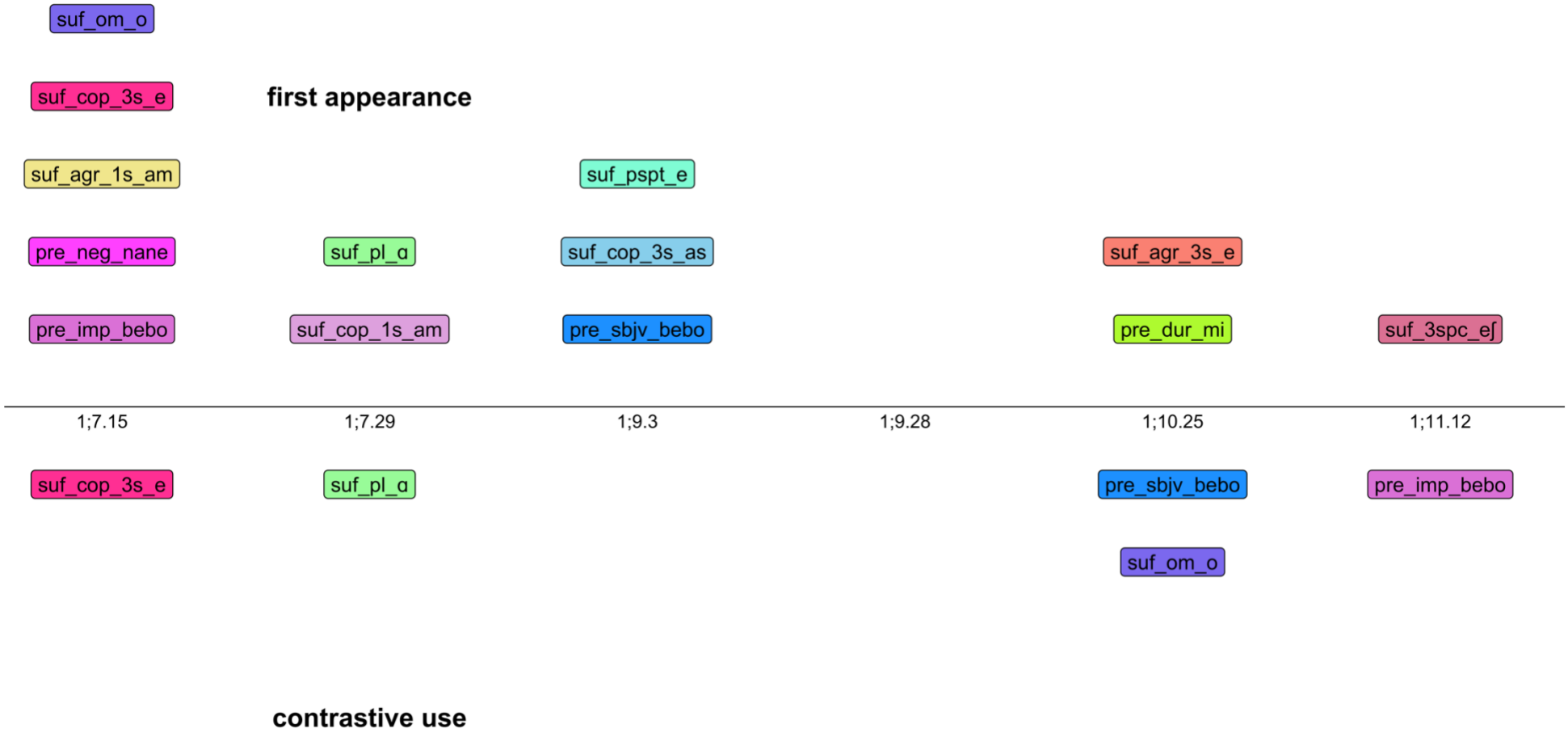

Figure 24 Saman's morphemes: first appearance vs. contrastive use

Summary

In this chapter, the linguistic development Saman was examined. Saman was recorded from the age of 1;7.15 to 1;11.12. The study found that Saman started with producing one word combination in the first session and progressed to 65 word combinations in the last session. The first word combination produced by Saman, which also contained inflectional morphology, was found to be accurate. However, Saman avoided using the object marker until later in the data collection. Additionally, it seems like Saman is showing evidence of pretty accurate usage of inflectional morphology following a U-shaped trajectory, where inflectional morphemes become less accurate before finally becoming productive.

6. Nikoo

Introduction

Nikoo is the second child of a monolingual family living in Tehran. She has an older sister. The mother was the primary caregiver and was not working at the time of the data collection. Her receptive and productive CDI scores were 28 and 4, respectively. Her CDI questionnaire was filled at the age of 1;3.19. A total number of 12 sessions were recorded. One session at the age of 1;9.6 was disregarded due to its bad sound quality. The first seven sessions were video-recorded by the researcher. The remaining sessions are audio-only, recorded by the mother of the infant. Her data collection lasted 10 months.

Nikoo's lexical development

The information about Nikoo's word tokens, types, lemmas, word combinations on each session, along with her age and the session lengths are presented in Table 26 and Figure 25. Nikoo was roughly at the 5-word point when the data collection started. She produced six spontaneous word types at 1;3.25. The 25-word point seems to have happened between the ages of 1;8.16 and 1;10.15 as there is a sudden rise in her word type production from 20 to 44 at the age of 1;9. As we can see here, she does not reach the 50-word point during the data collection period. Nikoo started combining words at the age of 1;4.19. The number of her word combinations increases sharply at 1;8.16. Nikoo had non-productive recording session at 1;7.21 which reflects in all of the measures. She was reluctant to talk to her mother for the majority of the session and was happily engaged playing on her own instead.

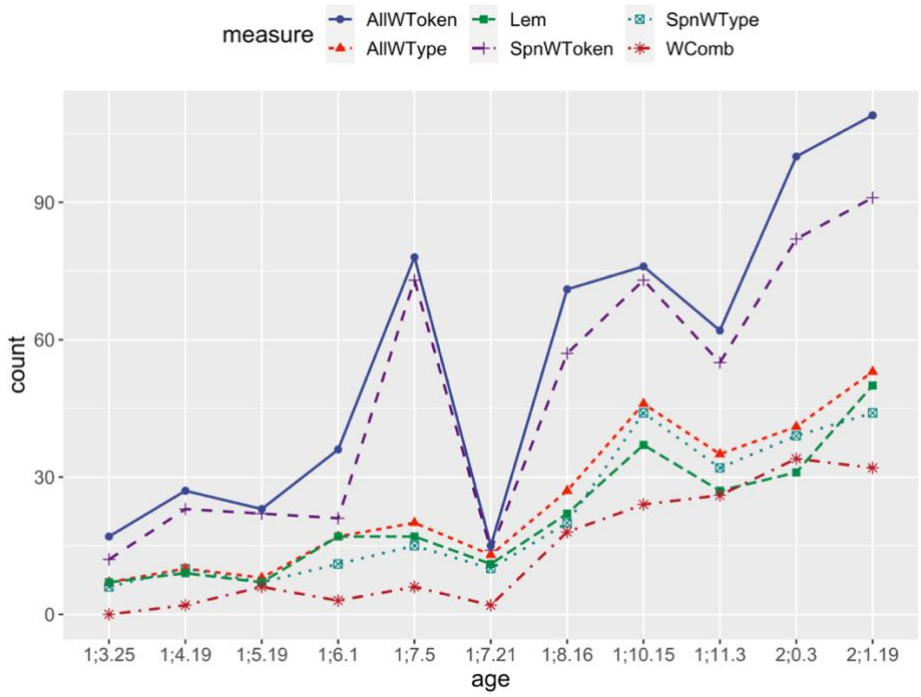

Figure 25 Nikoo's lexical development

Table 26 Nikoo's lexical development

session number	1	2	3	4	5	6	7	8	9	10	11
age	1;3.25	1;4.19	1;5.19	1;6.1	1;7.5	1;7.21	1;8.16	1;10.15	1;11.3	2;0.3	2;1.19
session length (min)	38	58	47	41	41	43	40	36	30	56	41
SpnWToken	12	23	22	21	73	14	57	73	55	82	91
SpnWType	6	10	7	11	15	10	20	44	32	39	44
AllWToken	17	27	23	36	78	15	71	76	62	100	109
AllWType	7	10	8	17	20	13	27	46	35	41	53
Lem	7	9	7	17	17	11	22	37	27	31	50
WComb	0	2	6	3	6	2	18	24	26	34	32

Nikoo's first recorded words

What is interesting is Nikoo's production of verbs from early on (Table 27). Her words are mostly, but not limited to, one syllable in length.

Table 27 Nikoo's first recorded words

Target Word	Gloss	Child forms
/mire/	DUR-go-3S	[ie:]
/oftɑd/	fell.3S	[otɑd]
/bijɑ/	IMP.come.2S	[ɡija], [pih]
/na/	no	[na], [nanana]
/in/	this	[iʔ]
/ʧie/	what-be.3S	[ʦieʔ]
/ku/	where	[ɡu], [ɡu:], [ɡu:ʔ], [ku], [ku:], [kuh]

Nikoo's consonant inventory

At 1;3.25 Nikoo's consonant inventory had eight consonants two of which were produced at least twice in either onset or coda positions (Table 28). She also produced the voiceless alveolar affricate /ʦ/ as a replacement for the target alveolar affricate /ʧ/.

Table 28 Nikoo's CI at 1;3.25

	labio-labial/ labio-dental	apico-alevolar/ apico-dental	dorso-post-alveolar	dorso-palatal	dorso-prevelar	dorso-post-velar	dorso-uvular	glottal
stop	p	t **d**			k **g**			ʔ
affricate								
fricative			[ʦ]					
nasal		**n**						
liquid								
glide				**j**				

The consonants produced at least twice, either in onset or coda positions, are marked in **bold face**.

The consonant produced in imitated words and as a replacement for a target word is inside brackets [ ].

At 1;8.16 Nikoo's consonant inventory size grew massively (Table 29). She produced 19 consonants, 12 of which appeared at least twice in either of the positions.

Table 29 Nikoo's CI at 1;8.16

	labio-labial/ labio-dental	apico-alevolar/ apico-dental	dorso-post-alveolar	dorso-palatal	dorso-prevelar	dorso-post-velar	dorso-uvular	glottal
stop	**b**	**t** **d**			k **g**		ɢ	**ʔ**
affricate			**ʧ** **ʤ**					
fricative		**s**	**ʃ** ʒ			x		h
nasal	m	**n**						
liquid		l **r**						
glide				**j**				

The consonants produced at least twice, either in onset or coda positions, are marked in **bold face**.

Table 29 shows Nikoo's consonant inventory. She had 19 consonants in her inventory. She produced 17 consonants in at least two word tokens in either of the positions.

Table 30 Nikoo's CI at 2;1.19

	labio-labial/ labio-dental	apico-alevolar/ apico-dental	dorso-post-alveolar	dorso-palatal	dorso-prevelar	dorso-post-velar	dorso-uvular	glottal
stop	**p** **b**	**t** **d**			**k** **g**		ɢ	**ʔ**
affricate			**ʤ**					
fricative	**v**	**s** z	**ʃ**			**x**		**h**
nasal	**m**	**n**						
liquid		**r**						
glide				**j**				

The consonants produced at least twice, either in onset or coda positions, are marked in **bold face**.

Nikoo's syllable structures

Looking at Figure 26, we can see that Nikoo's top syllable structure that accounts for at least 50% of her production at the age of 1;3.25 is the monosyllabic CV. During the next two months we see emergent use of disyllabic VCV and CVCV. At 1;6.1 we finally see the production of codas in her top CV structures. Until 1;7.21 there is not any two-syllable words with a final coda or any three-syllable long words among her top structures.

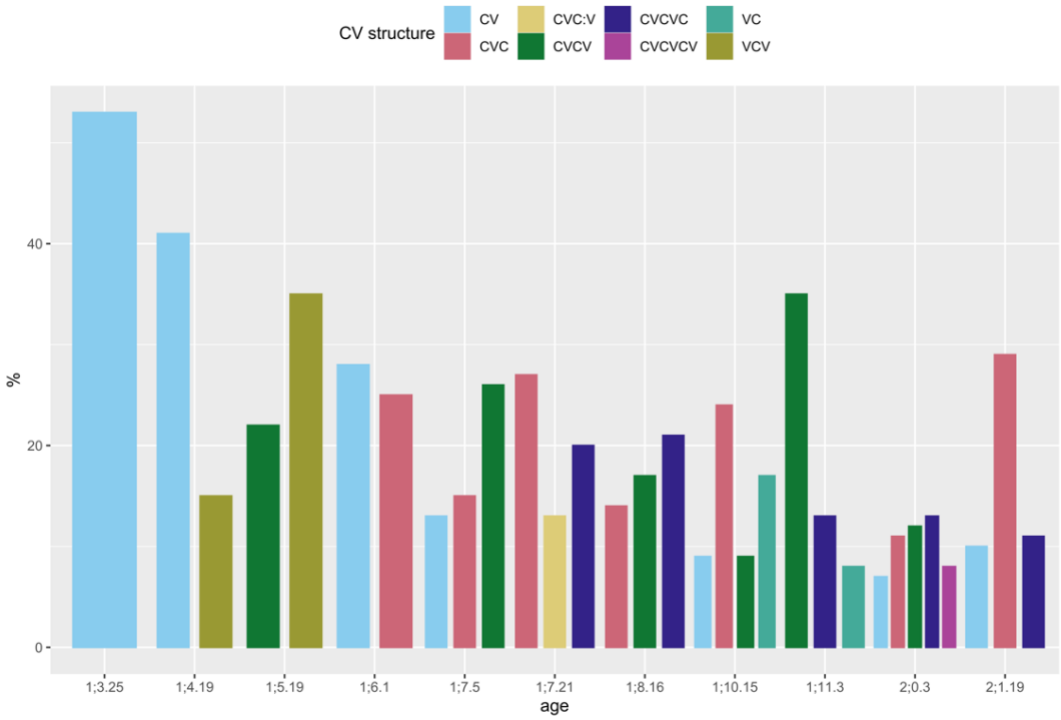

Figure 26 Nikoo's top syllable structures

Nikoo's syntactic development

Nikoo's first combinations can be found in Table 31. Her first recorded combination consists of an unintelligible word with /ku/ 'where'. The second combination is a fixed phrase made of two of her toys' names joined with /-o/ 'and'. Her mother and sister always referred to these two toys' names as one. So, despite the fact that this phrase consists of two separate words, it cannot be considered as a phrase combined by Nikoo. Her other combinations mostly consist of /ino/ 'this-OM' plus a verb.

Table 31 Nikoo's first combinations

Child form	Adult target	Gloss	Combination type	Age
[ʃi du]	xx ku	xx where	N + where	1;4.19
[dadojo dadojo]	aduso adusak	toy name-and toy name	N + o + N	1;5.19
[ino mixɑ]	in-o mi-xɑ-m	this-OM DUR-want-1S	object-om + V	1;5.19
[ino bete]	in-o be-de	this-OM IMP-give. 2S	object-om + impV	1;5.19
[o begi]	in-o be-gir	this-OM IMP-take.2S	object-om + impV	1;5.19
[ɑhi: biʔo]	mahi bi-ja	fish IMP.come.2S	N + impV	1;5.19

Figure 27 shows Nikoo's MLU scores. Second lowest MLUm, where a massive drop was seen at the age of 1;6.1, is the same age where consonant codas first emerge in her top CV structures. Highest MLUw, MLUm, and pMLU recorded on 1;11.3. This is interesting because all of her lexical measures drop (WType, WToken, Lem) on 1;11.3 (Figure 25). This is also the same age where her CVar score (Figure 28) and her PDSS (means) peak (Figure 29).

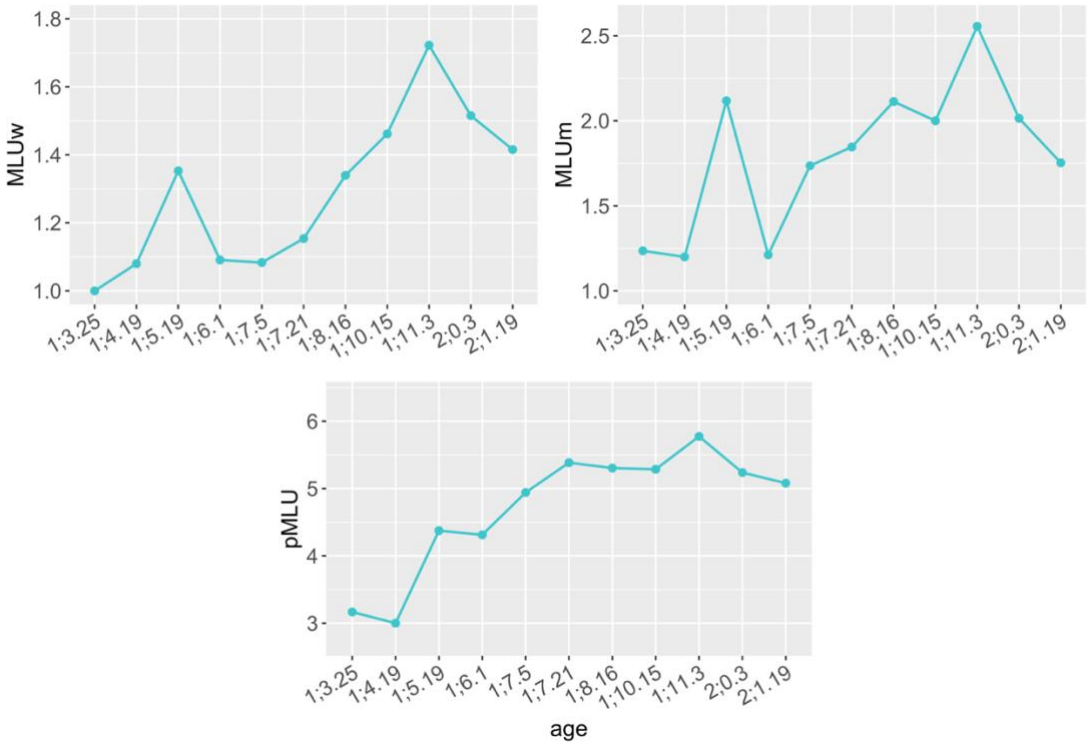

Figure 27 Nikoo's MLU scores

Nikoo's phonological development

Looking at Figure 28, it can be seen that Nikoo's PWC score peaks at session three at the age of 1;5.9. We also see a rise in her WComb. However, this session recorded a drop in all her lexical measures (WType, WToken, Lem) (compare with Figure 25). Her lowest PWC was recorded at 1;10.15. This coincides with the highest number of errors of omission recorded.

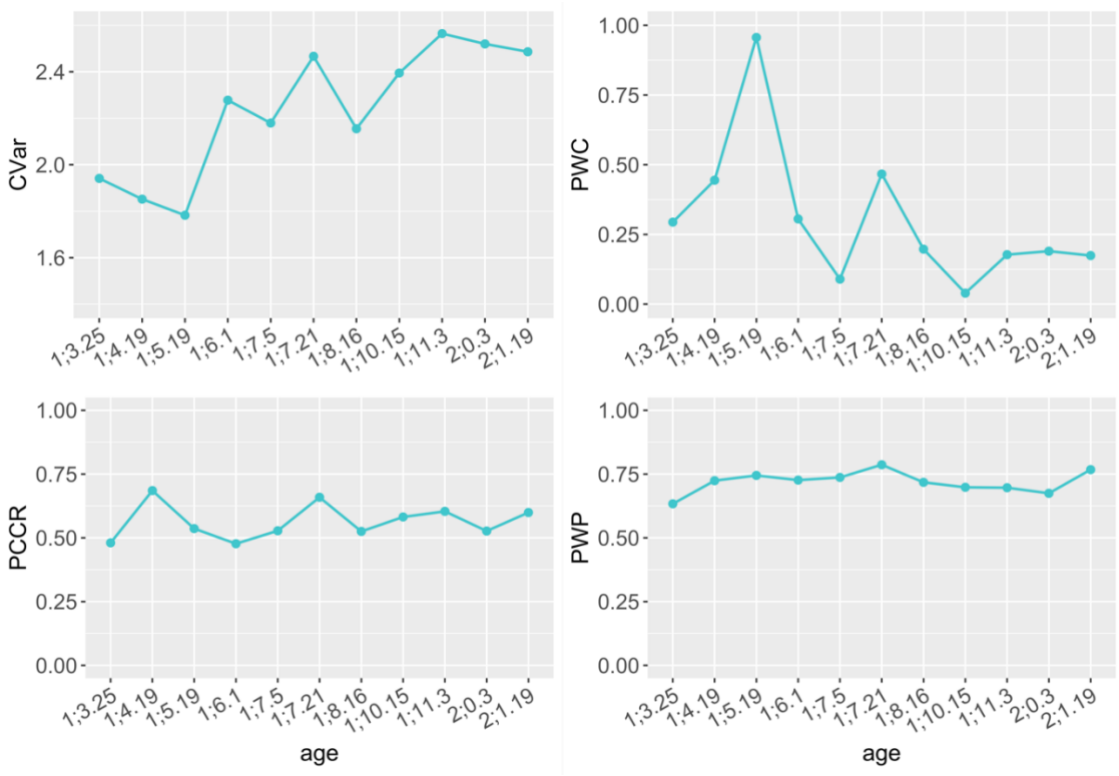

Figure 28 Nikoo's phonological development

Nikoo's morphological development

Nikoo's PDSS score

Nikoo's PDSS information can be seen in Table 32 and Figure 29 in detail. Her mean score more than doubled during the course of data collection. Her use of verbs from the first session is reflected in her PDSS score.

Table 32 Nikoo's PDSS score

Session number	Session length	age	PDSS_mean	PDSS_sd	PDSS_min	PDSS_max	PDSS_total
1	38	1;3.25	1.31	0.48	1	2	17
2	58	1;4.19	2	1.07	1	3	16
3	47	1;5.19	2.76	1.99	1	7	47
4	41	1;6.1	2.83	1.33	2	5	17
5	41	1;7.5	2.16	1.07	1	7	110
6	43	1;7.21	2.5	0.93	2	4	20
7	40	1;8.16	2.33	1.31	1	6	114
8	36	1;10.15	2.56	1.94	1	10	105
9	30	1;11.3	3.2	1.87	1	7	80
10	56	2;0.3	3	1.31	1	6	105
11	41	2;1.19	3	1.69	1	7	93

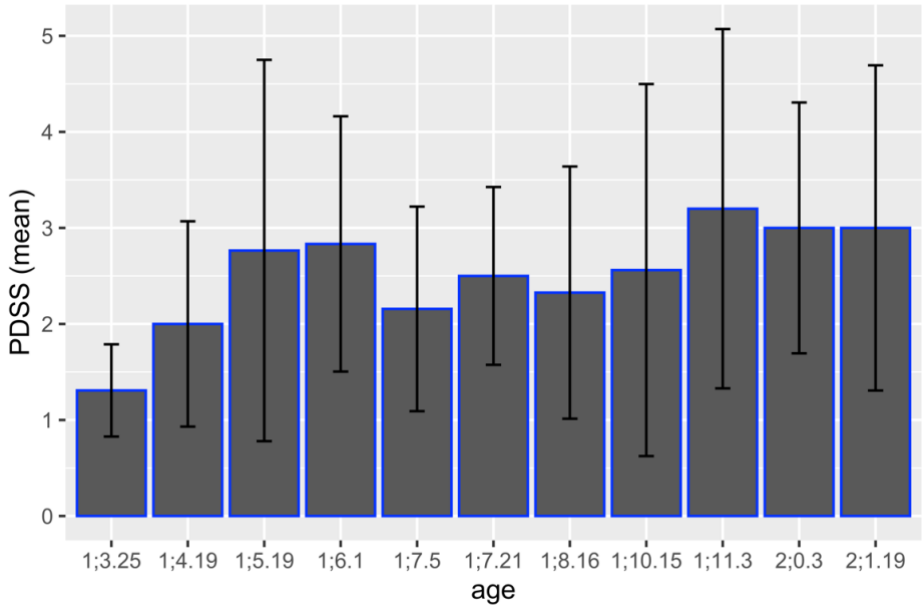

Figure 29 Nikoo's PDSS score

Among the sub-categories, Gmorph and Vmorph saw the biggest increase (Figure 30).

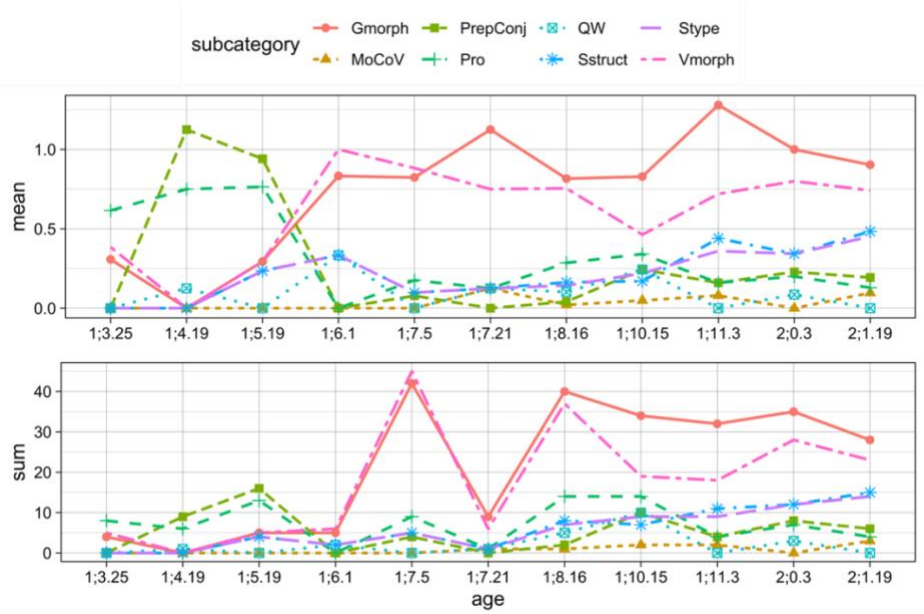

Figure 30 Nikoo's sub-categorical PDSS scores

Nikoo's morphemes

Nikoo's correct use of morphemes is presented in Figure 31 on the next page.

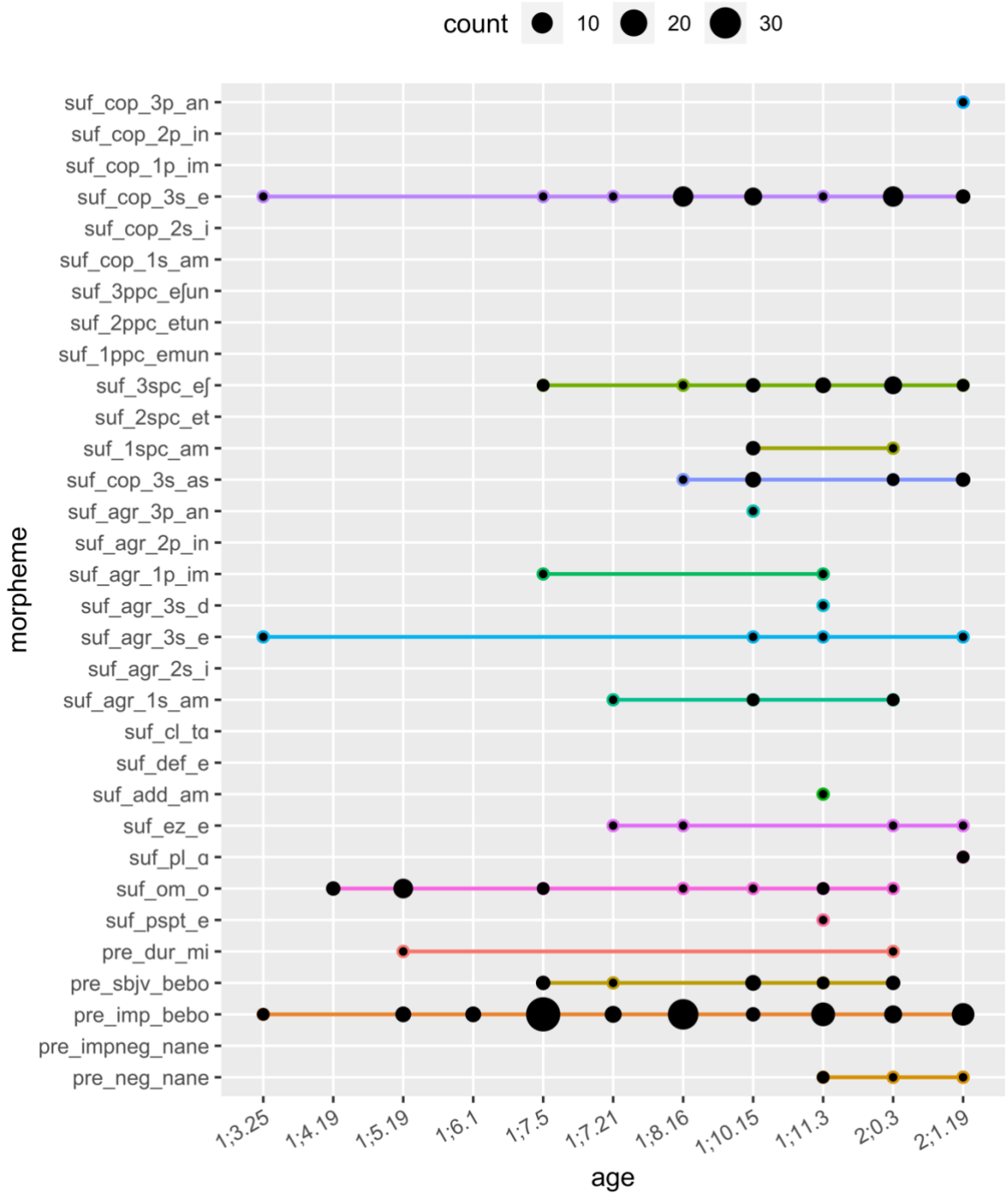

Figure 31 Nikoo's correctly used morphemes

At the age of 1;3.25, Nikoo's speech already had three morpheme types. However, none were productive and Nikoo probably learnt these words as a whole. These were [pih] and [gija] for /bi-jɑ/ 'IMP-come.2S', [ie:] for /mi-r-e/ 'DUR-go-3S', and [ʦieʔ] for / ʧi-je/ 'what-be.3S'.

At 1;4.19 Nikoo first used OM /-o/ with /in/ 'this' three times. She did not attempt producing

it without the morpheme during this session. She did, however, produce [in] 'this' at 1;3.25.

At 1;5.19 DUR /mi-/ first emerged in the following utterance:

(6.1)	CF:	in-o	mi-xɑ
	AT:	in-o	mi-xɑ-m
		this-OM	DUR-want-1S
		'I want this'	

Nikoo produced [in] 'this' and [in-o] 'this-OM' contrastively at this age. She started using IMP /be-,bo-/ with two new lemmas:

(6.2)	CF:	be-gi	vs.	be-te	vs.	bi-je
	AT:	be-gir	vs.	be-de	vs.	bi-jɑ
		IMP-take.2S		IMP-give.2S		IMP-come.2S

At 1;6.1 Nikoo only produced IMP /be-,bo-/ in two verb types. No new morphemes recorded in her speech.

At 1;7.5 she used IMP /be-,bo-/ with more verbs. The first three instances of SBJV /be-,bo-/ were recorded in [be-ʃ:i] for /be-ʃin-am/ 'SBJV-sit-1S', [i-ʃin-im] for /be-ʃin-im/ 'SBJV-sit-1P', and [bu-xu] for /bo-xor-im/ 'SBJV-eat-1P'. Verbal agreement 1P emerged at this age as seen in [i-ʃin-im] for /be-ʃin-im/ 'SBJV-sit-1P'. One more case of copula 3S /-e/ was recorded in [kije] 'who-be.3S'. She produced two tokens of 3SPC /-eʃ/ which functions as the object of the verb: [be-de-ʃ] 'IMP-give.2s-3SPC (give it to me)'.

At the age of 1;7.21 there were two new morphemes recorded in Nikoo's speech, EZ particle and verbal agreement 1S /-am/. The use of EZ particle is presented in the following utterance:

(6.3)	CF:	ʧɑʧ:-e	ban
	AT:	ʧɑji-e	man
		tea-EZ	I
		'my tea'	

Verbal agreement 1S /-am/ was used contrastively as follows:

(6.4)	CF:	ba:dɑr-am	vs.	ba:dɑr
	AT:	bardɑr-am	vs.	bardɑr
		(SBJV)pick up-1S		(IMP)pick up.2S

As in the previous data collection sessions, copula 3S /-e/ only appeared with one word type, as can be seen in the following utterance:

(6.5)	CF:	birih	in-e
	AT:	sibil	in-e
		moustache	this-be.3S
		'This is the moustache'	

At the age of 1;8.16 there was only one instance of a new morpheme: copula 3S /-as/. Nikoo started using copula 3S /-e/ productively at this age with different lemmas, for example: [ɢaʃanʃ-e] for /ɢaʃang-e/ 'beautiful-be.3S' and [du-e] for /ʤuʤu-e/ 'birdie-be.3S'. She used some of these lemmas without the morpheme as well.

At 1;10.15 two new morphemes emerged for the first time in Nikoo's speech. One was 1SPC /-am/ that attached to two different lemmas: [dad-ɑ:m] for /dast-am/ 'hand-POSS.3SPC' and [dub-am] for /tup-am/ 'ball-POSS.3SPC'. The second morpheme that emerged was verbal agreement 3P /-an/ in [ba tan-ah] for /bɑz mi-kon-an/ 'open DUR-do.3P'.

Copula 3S /-as/ continued to be added to more lemmas such as [oʤɑ-s] for /koʤɑ-s/ 'where-be.3S' and [goba-h] for /gorba-s/ 'cat-be.3S'.

At the age of 1;11.3 Nikoo had four new morphemes. One was NEG /na-,ne-/ as in [na-dɑj-i] for /na-dɑr-im/ 'NEG-have-1P'. The next one was PSPT 3S /-e/ in [daf-hi] for /raft-e/ 'went-PSPT.3S'. The third morpheme was verbal agreement 3S /-d/ in [bi-jɑ-d] 'SBJV-come-3S'. and the fourth case was the ADD morpheme /-am/ in the following utterance:

(6.6)	CF:	nɑreg-a	be-de

AT: nɑrengi-jam be-de
tangerine-ADD IMP-give-2S
'Also give (me) tangerine!'

At 2;0.3 new no morphemes emerged in Nikoo's speech. But Nikoo used /-eʃ/ with two different lemmas as in [doʃ] for /tu-ʃ/ 'inside-3SPC' and the following utterance:

(6.7) CF: da-ʃɪ bi-de
AT: dar-eʃ-o be-de
cap-POSS.3SPC IMP-give.2S
'Give (me) its lid'

At 2;1.19 Nikoo produced two new morphemes: PL /-(h)ɑ/ and verbal agreement 3P /-an/. PL /-(h)ɑ/ appeared on two different lemmas: [dub:-ɑ] for /tup:-ɑ/ 'ball-PL' and [ɢuɢu-hɑ] for /ʤuʤu-hɑ/ 'birdie-PL'. Verbal agreement 3P /-an/ appeared in the following utterance:

(6.8) CF: ʃu bɑj-ani
AT: tu ɑb-an
in water-be.3S
'They are in the water'

Nikoo's inflectional errors

Nikoo's speech only contained one omission error at the age of 1;3.25 (Figure 32).

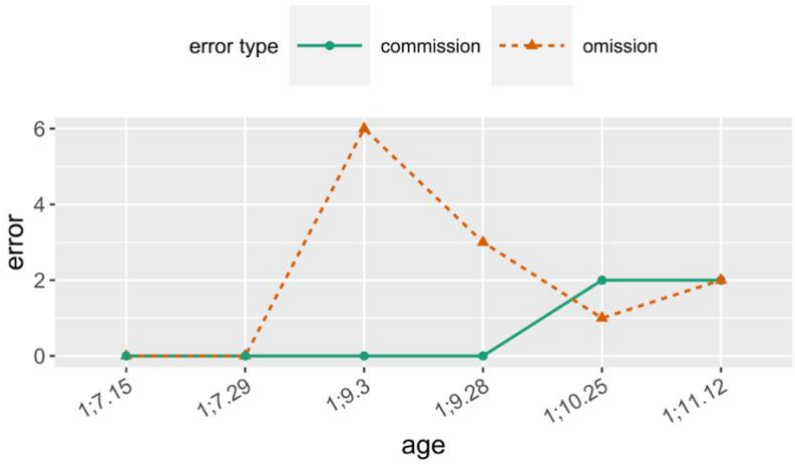

Figure 32 Nikoo's commission vs omission errors

This error was the omitted DUR /mi-/ in [i-e:] for /mi-r-e/ 'DUR-go-3S'. No errors were recorded on 1;4.19. At the age of 1;5.19 two omission errors were recorded. One was OM /-o/ in [in] for [in-o] 'this-OM'. The other error was the case of a missing verbal agreement 1S /-am/ which occurred in the following utterance:

(6.9) CF: in-o mi-xɑ
AT: in-o mi-xɑ-m
this-OM DUR-want-1S
'I want this (one)'

No errors were recorded at 1;6.1. At the age of 1;7.5 Nikoo omitted IMP /be-,bo-/ with three verb tokens (one lemma) as in the following: [de] for /be-de/ 'IMP-give.2S' and in the word combination below, which she produced as a single word:

(6.10) CF: um:-deh
AT: in-o be-de
this-OM IMP-give.2S
'Give me this (one)'

One case of omitted OM /-om/ was recorded, as seen in the above production. She also omitted verbal agreement 1S /-am/ in [be-ʃ:i] for /be-ʃin-am/ 'SBJV-sit-1S' and verbal agreement 1P /-im/ in [bu-xu] for /bo-xor-im/ 'SBJV-eat-1P'.

No errors were recorded at 1;7.21. Nikoo's errors increased at 1;8.16. She omitted DUR /mi-/ and verbal agreement 1S /-am/ in [xɑ:] for /mi-xɑ-m/ 'DUR-want-1S'. One commission error of 3SPC /-eʃ/ in [koʤɑ-s-uʃ] for /koʤɑ-s/ 'where-be.3S' appeared in her speech. Nikoo also used copula 1P /-im/ instead of copula 3S in [ɢaʃ-im] 'beautiful-be.1P' for /ɢaʃang-e/ 'beautiful-be.3S'.

At 1;10.15, omission of DUR /mi-/ continued: [ba tanah] for /bɑz mi-kon-an/ 'open DUR-do-3P'. EZ particle omission was recorded three time in the two following utterances: [nak deda] for /ejnak-e nirvɑnɑ/ 'glasses-EZ proper name' and [ɑʃɑ mah] for /ɑʃraf-e man-e/ 'proper name-EZ I-be.3S'. In the latter utterance, copula 3S /-e/ was omitted.

Nikoo used verbal agreement 2S /-i/ instead of 1P /-im/ in [bɑ:di kon-i] 'play (SBJV)do-2S for /bɑzi kon-im/ 'open (SBJV)do-1P'. Copula 3S /-as/ was missing in [kijɑ] for /kojɑ-s/ 'where-be.3S'. In [inɑ:ʃ-e:] 'here-3SPC-be.3S' an extra /-e/ was recorded (for /inɑhɑʃ/ 'here-3SPC').

At 1;11.3 Nikoo made more OM /-o/ errors, for instance:

(6.11)	CF:	iʔ	be-zɑ	inʤɑ
	AT:	in-o	be-zɑr	inʤɑ
		this-OM	IMP-put.2S	here
		'Put this here!'		

Nikoo's use of verbal agreement 2S /-i/ instead of 1P /-im/ continued at 1;11.3:

(6.12)	CF:	ɑb:ɑʃ	na-dɑj-i
		water game	NEG-have-2S
	AT:	ab:ɑzi	na-dɑr-im
		water game	NEG-have-1P
		'We're not gonna play water games'	

At 2;0.3 Nikoo verbal agreement 3S /-d/ was omitted in two cases for the verb *umadan* 'to come': [bi-je] for /bi-jɑ-d/ 'SBJV-come-3S' and [me-ʔa] for /mi-jɑ-d/ 'DUR-come-3S'. Erroneous use of copula 3S /-as/ continued, as in the example below:

(6.13) CF: un gij-as
AT: in ʧi-e
this what-be.3S
'What is this?'

(6.14) CF: ɑb:ʃ tu
AT: ɑb (be-riz) tu-ʃ
water (IMP-pour.2S) inside-3SPC
'Pour water inside it'

At 2;1.19 one more omission error of copula 3S /-e/ was recorded, as well as commission errors of 3SPC /-eʃ/. An example of this commission error follows:

(6.15) CF: na-dɑd-eʃ
NEG-have-3SPC
AT: na-dɑr-e
'NEG-have-3S'
'He/she/it doesn't have (it)'

Nikoo's errorneous use of morphemes can be seen in Figure 33.

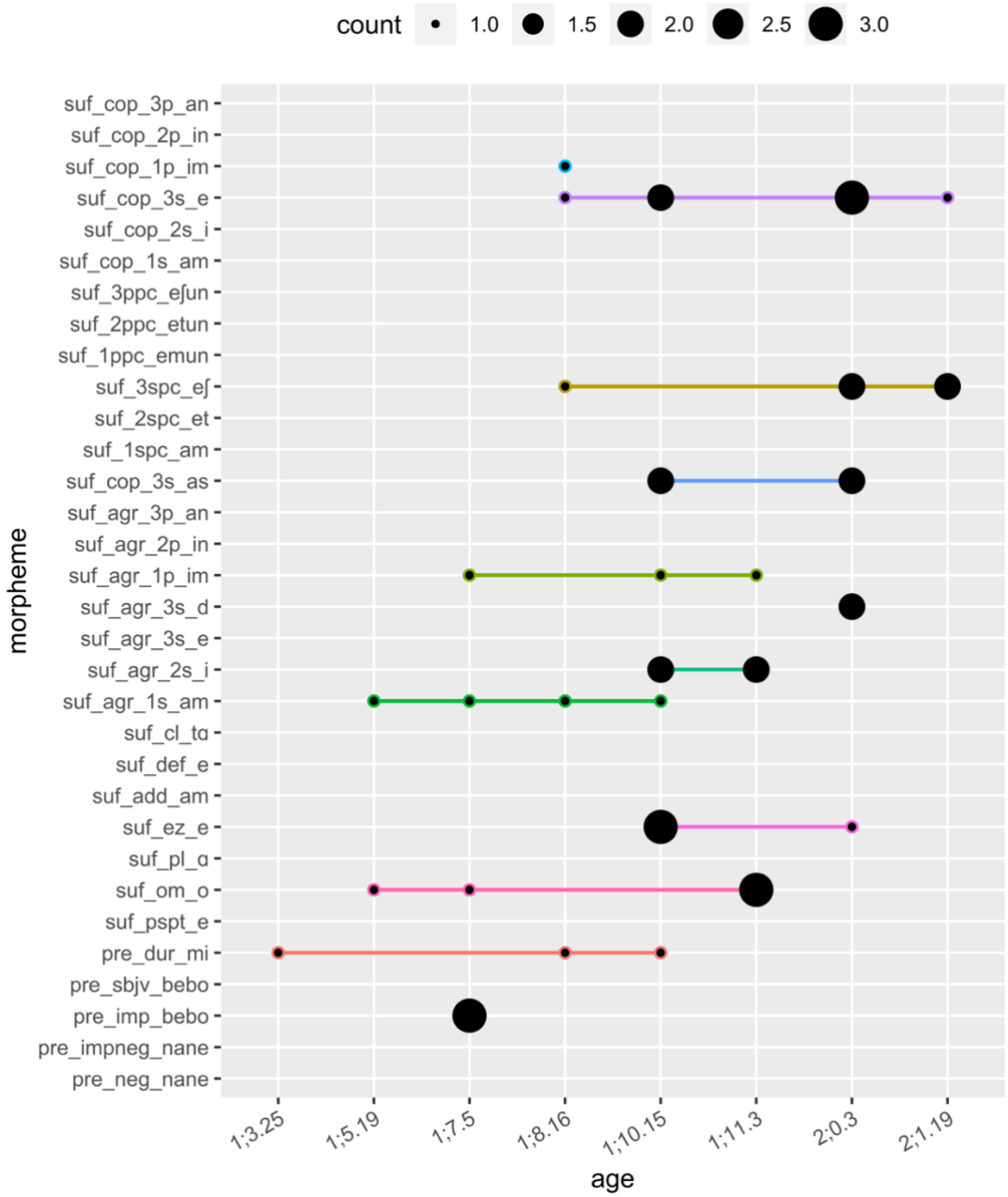

Figure 33 Nikoo's erroneous use of morphemes

First appearance vs. contrastive use

Finally, the first appearance of Nikoo's inflectional morphemes and the first contrastive use are presented in Figure 34. It is interesting to see that six of her inflectional morphemes were first appeared and used at the same age. But there are also inflectional morphemes that she used contrastively with a short delay, for example OM /-o/ (1;4.19 for first appearance versus 1;5.19 for contrastive use), and with a longer delay, as in the case of copula 3S /-e/ (1;3.25 for first appearance versus 1;8.16 for contrastive use).

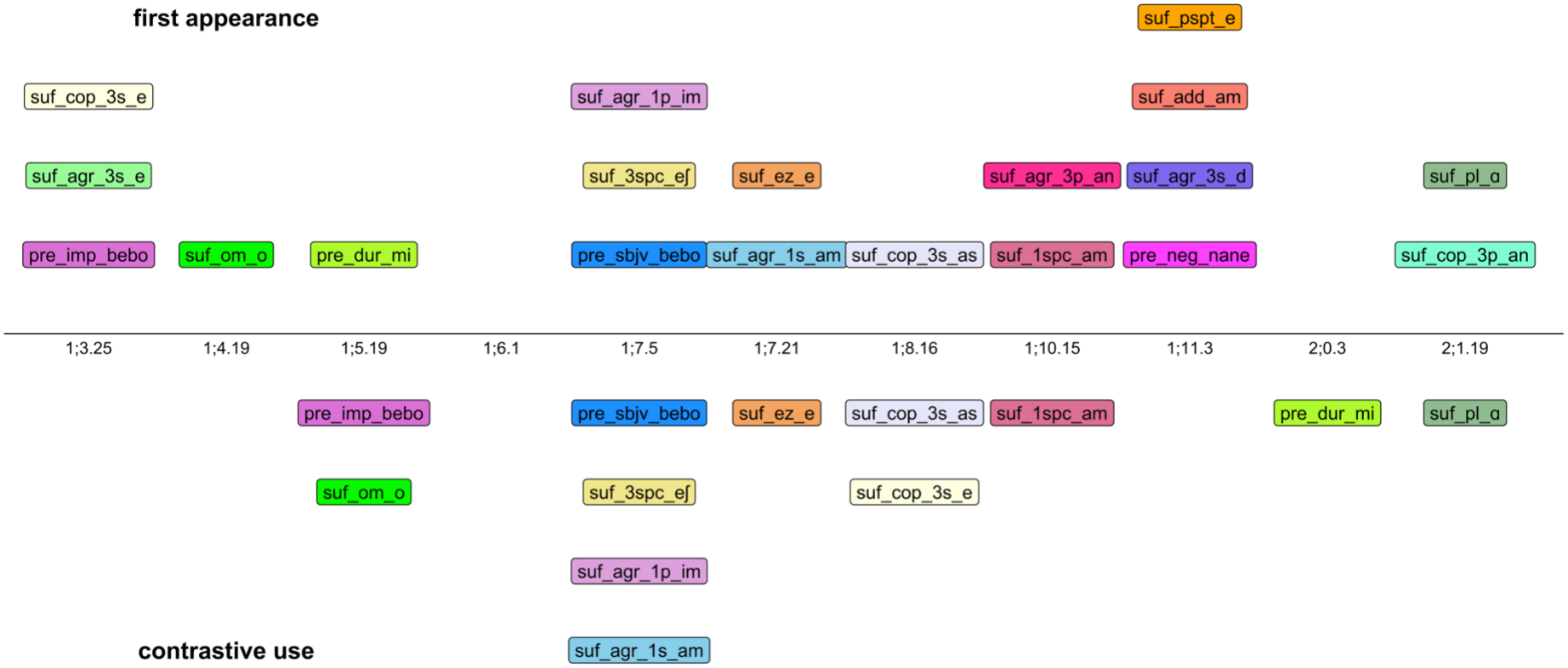

Figure 34 Nikoo's morphemes: first appearance vs. contrastive use

Summary

This chapter presents an examination of the linguistic development of Nikoo. Data was collected through recording sessions with Nikoo from the age of 1;3.25 to 2;1.19. The study found that Nikoo began with no word combinations in the first session and progressed to producing 32 word combinations in the last session. Notably, Nikoo began producing inflectional morphemes before producing any word combinations. The first productive use of inflectional morphemes by Nikoo was observed at 1;4.19, at which point she had produced 6 word combinations. The chapter provided an in-depth analysis of Nikoo's development, highlighting the unique aspects of her language acquisition process.

7. Hediye

Introduction

Hediye is the first child of a family living in Tehran. She has a younger sister. Her mother did not work at the time of the data collection and was her primary caregiver. Her receptive and productive scores were 270 and four on the pre-session at the age of 1;8.8. As mentioned in chapter 3, Hediye's father attempts to expose her to German after returning home from work, accounting for approximately half the time he spends with her. Despite her father's attempt to expose her to German, Hediye did not use any German words in her speech during the period of data collection except for the words [omɑ] 'grandmother' and [opɑ] 'grandfather' for her German grandparents.

The data collection started at the age of 1;8.16 and continued to the age of 2;1.10. A total number of nine sessions were recorded, however one session at the age of 1;8.16 had to be cancelled mid-session due to the child being unwell.

Hediye's lexical development

Table 33 and Figure 35 displays Hediye's lexical development. Looking at this table, it is apparent that she had passed the 5-word point when the data collection started. She shows a steady rise in all the measures. At 2;1, there is a sharp increase especially in word tokens, word types and word combinations. She produced the word combination [in ʧi-je] 'what this-be.3S (= What is this?)' and [in xub-e] 'this good-be.3S (= This is good)' 198 and 73 times, respectively. As we will see later, this reflects in her higher PDSS score for this session.

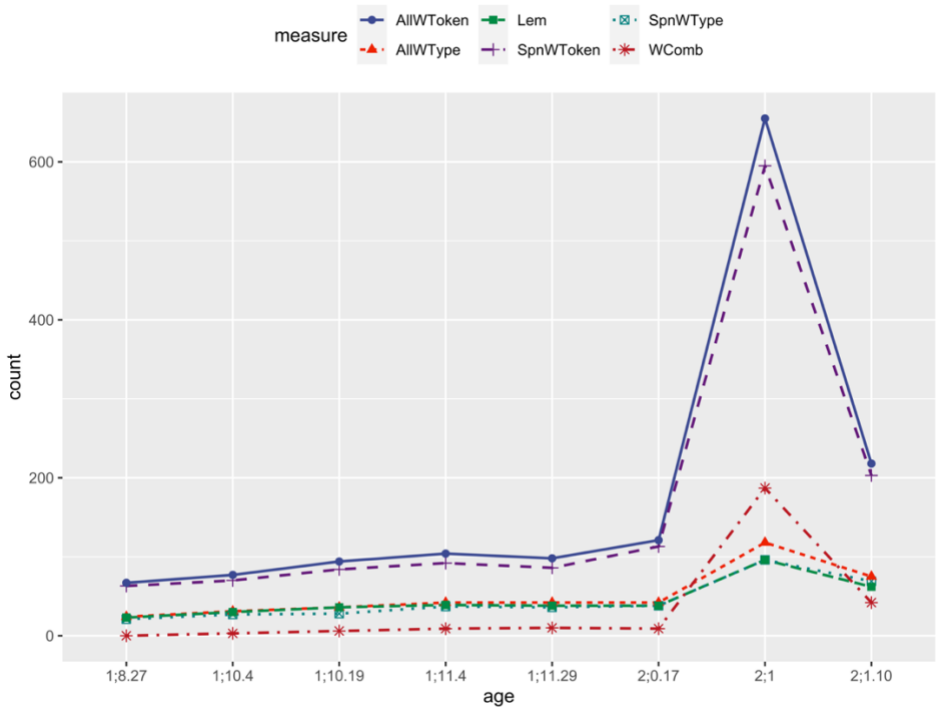

Figure 35 Hediye's lexical development

Table 33 Hediye's lexical development

session number	1	2	3	4	5	6	7	8
age	1;8.27	1;10.4	1;10.19	1;11.4	1;11.29	2;0.17	2;1	2;1.10
session length (min)	34	40	42	45	40	39	45	50
SpnWToken	63	70	84	92	86	113	595	203
SpnWType	21	27	28	37	36	38	96	69
AllWToken	67	77	94	104	98	121	655	218
AllWType	24	31	36	42	42	42	118	75
lemmas	23	30	36	39	38	38	96	62
word combinations	0	3	6	9	10	9	187	42

Hediye's first recorded words

Table 34 presents Hediye's words recorded on the first session. An interesting fact about her first identified words is the combination words produced as single words. One structure is a person (or an animal) plus the imperative form of the verb 'coming', i.e., [ɑjbijɑ] for /ɑje bijɑ/ 'proper name IMP.come.2S' and [abija] for /hɑpu bijɑ/ 'doggie IMP.come.2S. The other is made from omitting the second syllable of the first word and the first syllable of the second

word, i.e., [ʧ:ʤe] for /ʧi ʃod-e/ 'what become-PSPT.3S'. As the data presented in this table shows, she had several variants for a single adult target.

Table 34 Hediye's first recorded words

Target Word	Gloss	Child forms
/alo/	phone greeting	[aboɑ]
/asb/	horse	[a:s]
/ɑj/	ouch	[ɑji]
/ɑje bijɑ/	proper name IMP.come.2S	[ɑjbijɑ]
/ɑx/	ouch	[aox], [aoʔ], [ɑ:x], [ɑx], [o:]
/bijɑ/	IMP.come.2S	[bi], [bi:jɑ], [bi:jɑ:], [bija], [bijɑ], [bijɑ:], [bijɑh]
/dɑl:i/	peek-a-boo	[daʔ]
/do/	two	[ido] IM
/ʤuʤu/	birdie	[dudah], [dudu], [duduɑ]
/hɑdi/	proper name	[ad:i]
/hɑpu/	doggie	[ɑbox], [hɑpu:] IM
/hɑpu bijɑ/	doggie IMP.come.2S	[abija]
/in/	this	[iʔ]
/mɑmɑnʤun/	grandmother	[mamɑ] IM
/mersi/	thanks	[a:si:], [e:si:], [es]
/mio/	kitty	[mio]
/mixɑj/	DUR-want-2S	[xɑ], [xɑ:], [xɑv]
/nɑzi/	petting	[ʔa:h]
/oh/	oh	[ɑo]
/omɑ/	grandmother (German)	[omɑ:] IM
/ʧi ʃode/	what become-PSPT.3S	[ʧ:ʤe]
/ʧie/	what-be.3S	[ʧie]
/ʔe/	oh	[ʔɑ:], [ʔɑo], [ʔɑ], [eh], [ʔe], [ʔo]

IM = imitated and modelled words

Hediye's consonant inventory

At the beginning of the data collection, at the age of 1;8.27, Hediye produced 11 consonants (Table 35). The stops /d, ʔ/, the nasal /m/ and the glide /j/ appeared in two word tokens in either of the positions.

Table 35 Hediye's CI at 1;8.27

	labio-labial/ labio-dental	apico-alevolar/ apico-dental	dorso-post-alveolar	dorso-palatal	dorso-prevelar	dorso-post-velar	dorso-uvular	glottal
stop	p **b**	**d**						**ʔ**
affricate			ʧ					
fricative	v	s				x		h
nasal	**m**							
liquid								
glide				**j**				

The consonants produced at least twice, either in onset or coda positions, are marked in **bold face**.

Table 36 shows that Hediye's consonant inventory size grew at 1;11.4. Hediye had 14 out of her 19 consonants produced twice in either of the positions.

Table 36 Nikoo's CI at 1;11.4

	labio-labial/ labio-dental	apico-alevolar/ apico-dental	dorso-post-alveolar	dorso-palatal	dorso-prevelar	dorso-post-velar	dorso-uvular	glottal
stop	**p** **b**	**d**			**k** **g**		**ɢ**	**ʔ**
affricate			ʧ					
fricative	f	s z	ʃ			**x**		**h**
nasal	**m**	**n**						
liquid		l **r**						
glide				**j**				

The consonants produced at least twice, either in onset or coda positions, are marked in **bold face**.

At 2;1 Hediye's consonant inventory was almost adult-like (Table 37). She produced 22 out of 23 consonants. She produced 21 consonants in either onset or coda positions. The only consonant not present in her CI was the fricative /ʒ/.

Table 37 Hediye's CI at 2;1

	labio-labial/ labio-dental	apico-alevolar/ apico-dental	dorso-post-alveolar	dorso-palatal	dorso-prevelar	dorso-post-velar	dorso-uvular	glottal
stop	**p b**	**t d**			**k g**		**ɢ**	**ʔ**
affricate			**ʧ ʤ**					
fricative	**f v**	**s** z	**ʃ**			**x**		**h**
nasal	**m**	**n**						
liquid		l **r**						
glide				**j**				

The consonants produced at least twice, either in onset or coda positions, are marked in **bold face**.

Hediye's syllable structures

Hediye's top syllable structures are mostly monosyllabic at the beginning of the data collection along with the disyllabic CVCV. First codas emerged on 1;10.19 in CVC:VC and CVCVC. At 2;0.17 we see a big increase in the use of VC structure. This is the session which was discussed earlier. The repetitive use of the word *in* 'this' in the phrases [in ʧi-je] 'what this-be.3S (= What is this?)' and [in xub-e] 'this good-be.3S (= This is good)' contributed to CV being her top structure in that session (Figure 36).

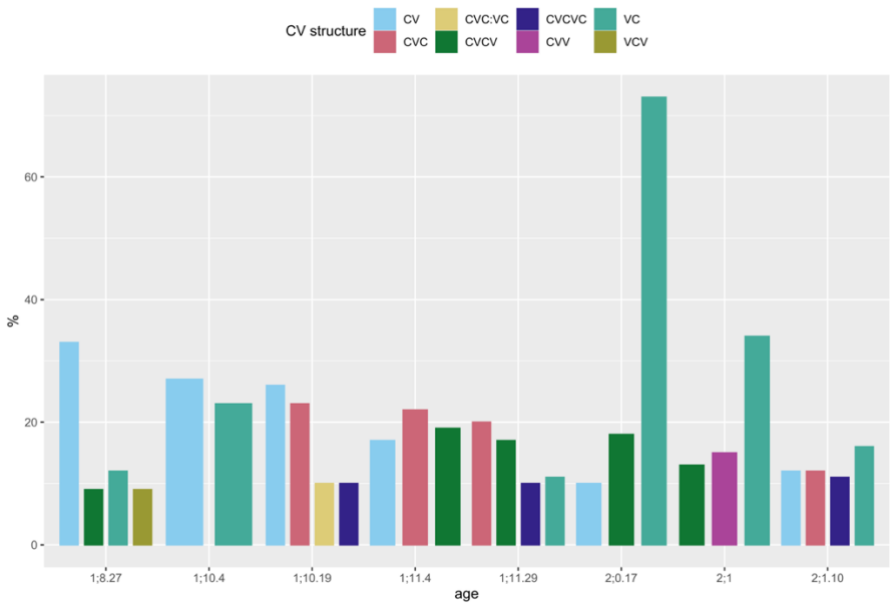

Figure 36 Hediye's syllable structures

Hediye's syntactic development

Hediye's first combinations at the age of 1;10.4 vary in type. She has N + V, N + N, and N + impV (see Table 38).

Table 38 Hediye's first combinations

Child form	Adult target (by word)	Gloss	Combination type	Age
[hɑp mixɑ]	/hɑpu mixɑm/	doggie DUR-want-1S	noun + verb	1;10.4
[bɑz nis]	/bɑzi nis(t)/	game NEG.be.3S	noun + verb	1;10.4
[e ɑji]	/ʔe dɑl:i/	oh! peek-a-boo	NA	1;10.4
[um bede]	/un-o be-de/	that-OM IMP-give.2S	object-om + impV	1;10.19
[ʃi be:]	/ʃir be-de/	milk IMP-give.2S	INDobject + impV	1;10.19
[go:r bɑs]	/gorg bɑz/	wolf open	noun + noun	1;10.19
[bɑ:s ka:]	/bɑz kard/	open did.3S	noun + verb	1;10.19
[abu dada]	/hɑpu dadar/	doggie outside	noun + noun	1;10.19

Hediye's MLU scores all show a rising trend (Figure 37) with all being the lowest at 1;8.27 and peaking at 2;1. All her lexical measures as well as WComb peak at 2;1 (see Figure 35).

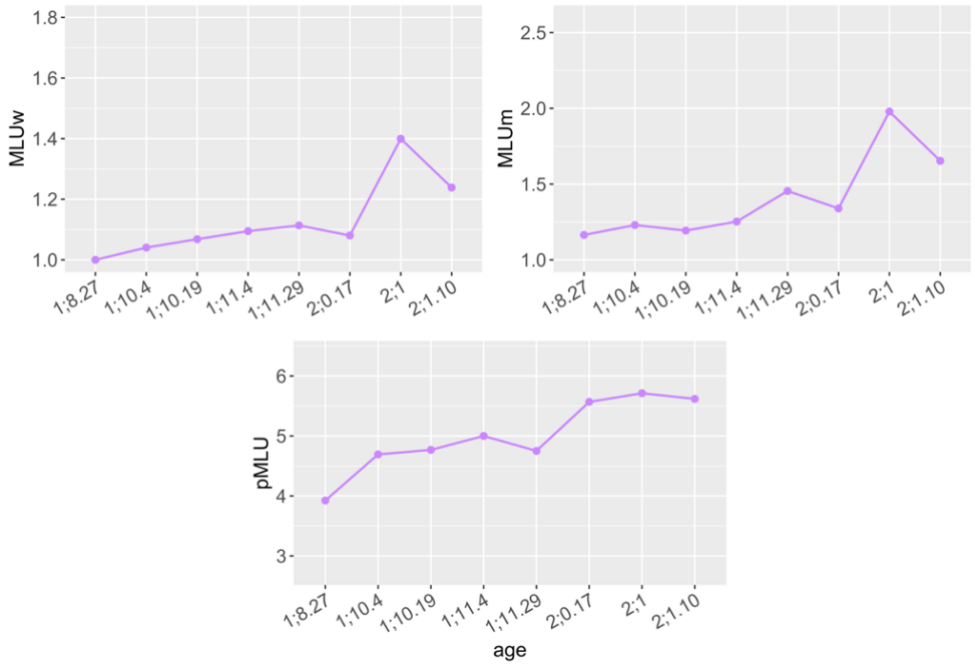

Figure 37 Hediye's MLU scores

Hediye's phonological development

Hediye's CVar shows a rising trend up to 1;11.29 followed by a falling trend. Her PCC-R and PWP scores change the least throughout the data collection period. Her PWC score more than doubles on 2;1 followed by a sharp decrease. These are shown in Figure 39.

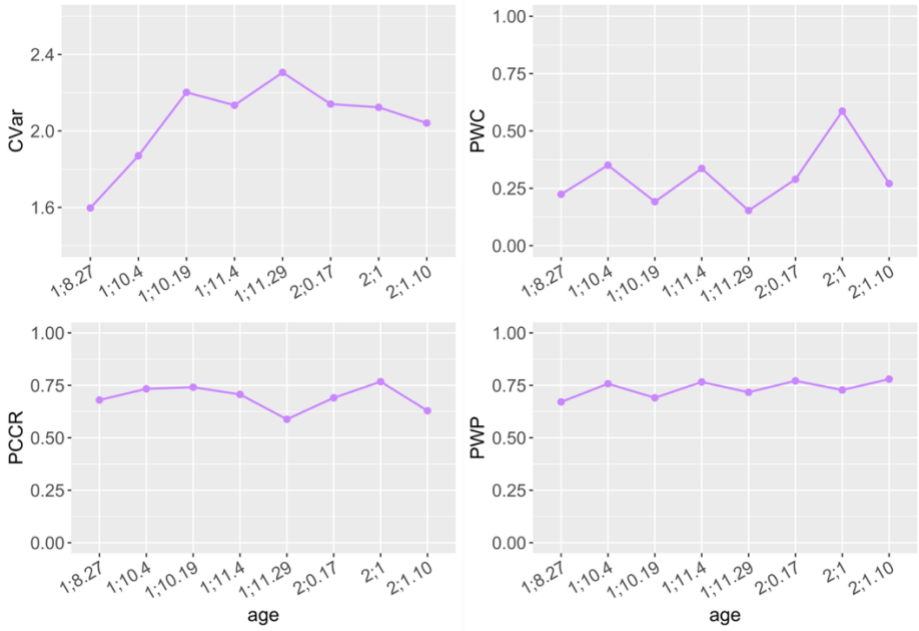

Figure 38 Hediye's phonological development

Hediye's morphological development

Hediye's PDSS score

Hediye's PDSS score, the standard deviation, the minimum, and the maximum scores are presented in Table 39 and Figure 39. As mentioned earlier, her repetitive use of two phrases resulted in a very high PDSS score at 2;1 (PSSS total = 1230, PDSS mean = 3.82). Putting this session aside, her PDSS score increase steadily with a few small decreases in the course of the data collection.

Table 39 Hediye's PDSS score

Session number	Session length	age	PDSS_mean	PDSS_sd	PDSS_min	PDSS_max	PDSS_total
1	34	1;8.27	1.81	0.75	1	4	29
2	40	1;10.4	1.67	0.87	1	4	40
3	42	1;10.19	2.5	1.09	1	5	35
4	45	1;11.4	2.43	1.69	1	6	51
5	40	1;11.29	2.23	1.33	1	5	87
6	39	2;0.17	1.95	1.46	1	8	82
7	45	2;1	3.82	2.13	1	7	1230
8	50	2;1.10	2.58	1.71	1	7	235

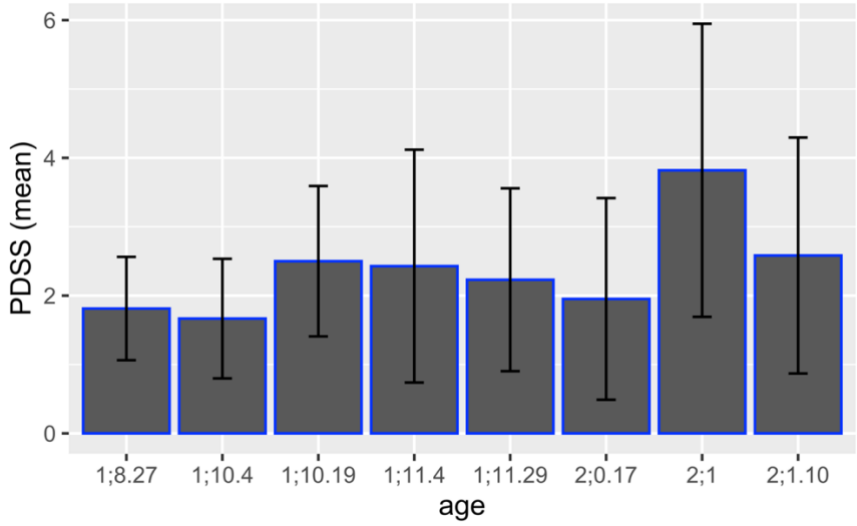

Figure 39 Hediye's PDSS score

Hediye's sub-categorical mean and total PDSS scores are shown in Figure 40. Comparing the mean PDSS and the total PDSS scores for all the sub-categories yet again shows the importance of providing both. This figure correctly depicts the impact of repeating two phrases on PDSS total.

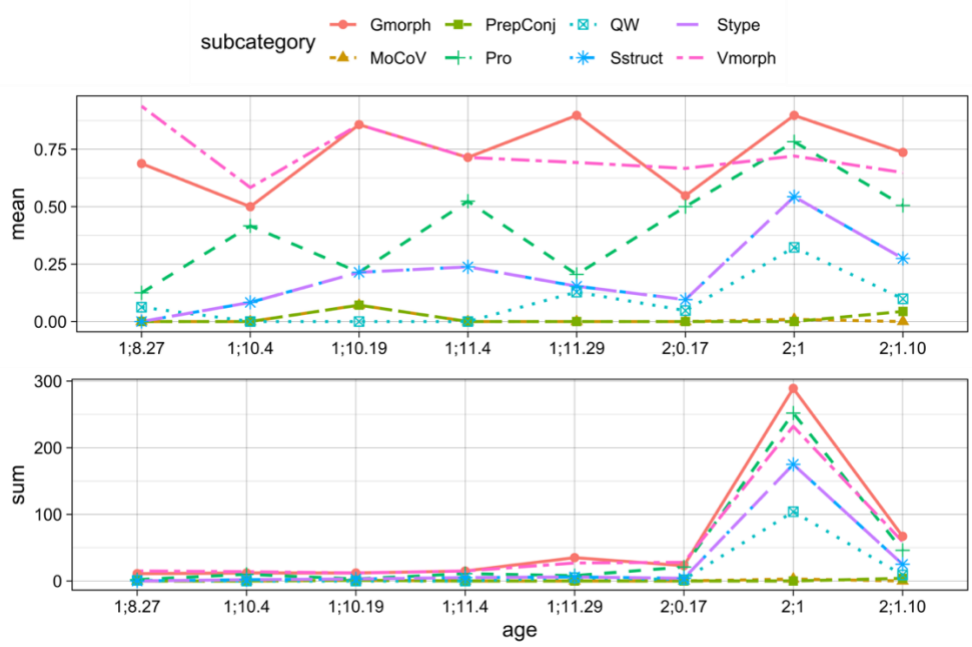

Figure 40 Hediye's sub-categorical PDSS score

Hediye's Morphemes

Hediye's correctly used inflectional morphemes are presented in Figure 41. A detailed analysis of her use of inflectional morpheme follows.

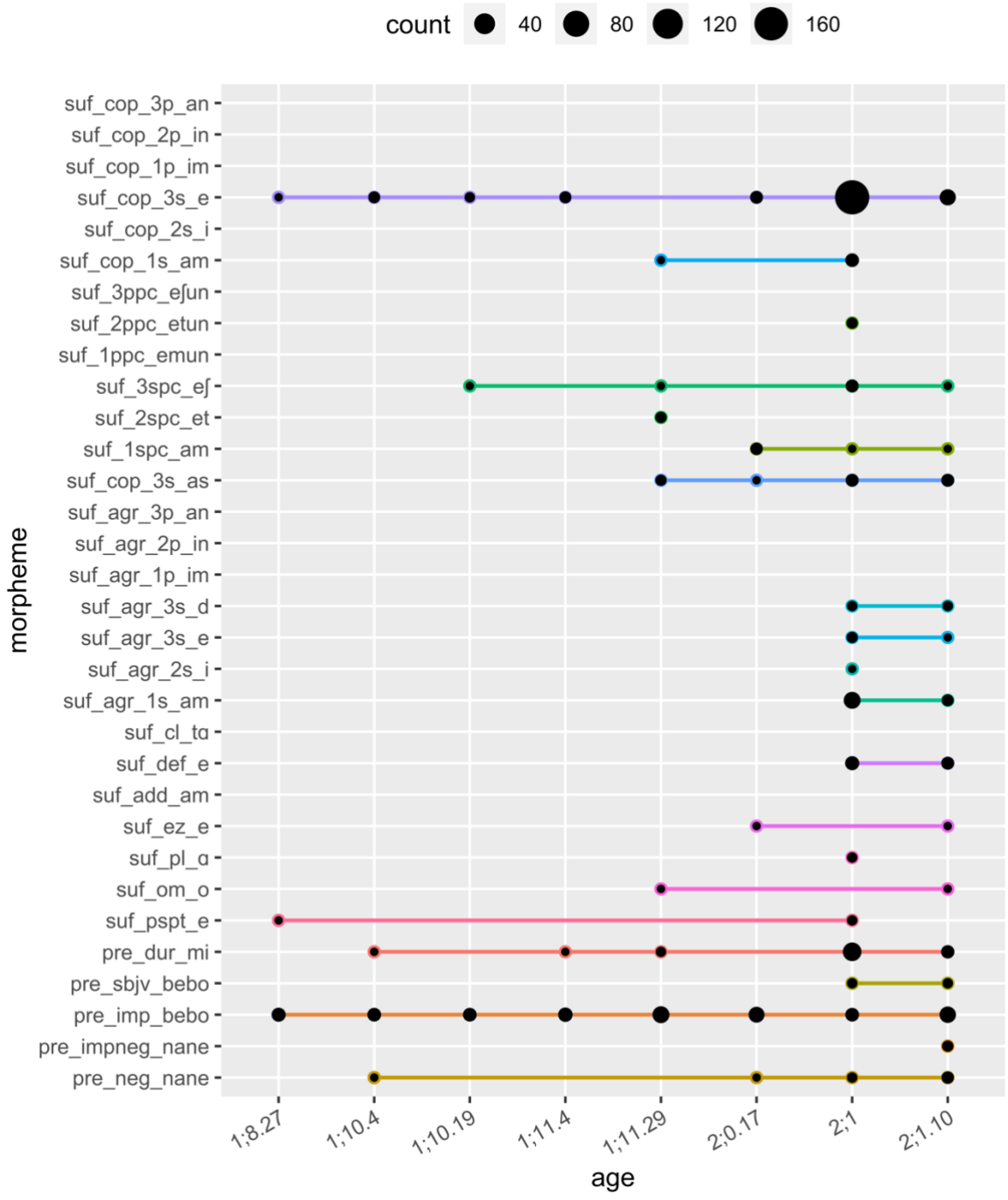

Figure 41 Hediye's correctly used morphemes

Hediye only had three types of morphemes at the age of 1;8.27. The first one was IMP /be-, bo-/. She only used this suffix with the verb *umadan* 'to come' in 2S form, which can be

considered the base form since it has no verbal agreement: [bijɑ] for /bi-jɑ/ 'IMP-come.2S'. It can be argued that she learnt this verb form as a whole, which means this cannot be used as an indication of her productive use of this morpheme. She used this verb to form word combinations which she pronounced as a single word in two different cases. An example of this is presented in (7.1). This phenomenon was not seen in the other children and soon stopped happening.

(7.1) CF: ɑjbijɑ
AT: ɑje bijɑ
proper name IMP-come.2S
'Aje! Come (here)!'

Hediye's other two used morphemes were PSPT suffix /-e/ and the copula suffix 3S /-e/, both of which were only used once. This again can be an indication of Hediye learning these words as a whole. In the first case, she combined two words which she pronounced as one:

(7.2) CF: ʧ:ʤe
AT: ʧi ʃode
what became-PSPT(3S)
'What's up?'

Hediye used more morphemes at the age of 1;10.4. The first use of IMP /be-, be-/ with more than one verb was recorded at this age. She also produced her first case of DUR /mi-/ at this stage. This prefix was missing in obligatory contexts at 1;8.27, as will be explained in the next section. The first productive use of copula 3S /-e/ was recorded at 1;10.4 as she used one a lemma with and without 3S /-e/ in different contexts:

(7.3) CF: ap:u-e vs. hɑp mi-xɑ
AT: hɑpu-e vs. hɑpu mi-xɑ-m
doggie-be.3S doggie DUR-want-1S
'It's a doggie' 'I want a doggie'

The first use of NEG /na-,ne-/ emerged at this age. Hediye only produced it once. Hediye did

not use this morpheme again until much later at 2;0.17.

At 1;10.19 a new morpheme first appeared in Hediye's speech. She produced 3SPC /-eʃ/ only once: [mɑmɑ-ʃ] for /mɑmɑn-eʃ/ 'mum-POSS.3SPC'. Contrastive use of IMP /be-,be-/ and copula 3S /-e/ continued.

No new morphemes emerged at 1;11.4. Hediye used DUR /mi-/ with a new verb at this age:

(7.4)	CF:	in	ɑb	mi-xoj-e (1;11.4)	vs.	hɑp	mi-xɑ (1;10.4)
	AT:	in	ɑb	mi-xor-e		hɑpu	mi-xɑ-m
		this	water	DUR-eat-3S		doggie	DUR-want-1S
		'This is drinking water'				'I want a doggie'	

Four new morphemes appeared a 1;11.29. Verbal agreement 2S /-et/ emerged three times attached to only one noun: [das:-a] for /dast-et/ 'hand-2S'. The first uses of copula 1S /-am/ and 3S /-as/ were also recorded at this age each: [man-am] "I-be.1S" and [dɑ:s] for /koʤɑs/ 'where-be.3S'. The first use of OM /-o/ was recorded in an imitated word at this age.

At the age of 2;0.17, two more morphemes emerged first in Hediye's speech. 1SPC /-am/, and EZ particle. 1SPC was attached to two different stems: [xod-am] 'self-POSS.1SPC' and [dad-a:m] for /dard-am/ 'pain-POSS.1SPC'. EZ particle was used in the utterance:

(7.5)	CF:	mal-e	xod-am-e
		possession-EZ	self-POSS.1SPC-be.3S
		'(This) is mine'	

Additionally, NEG /na-,ne-/ re-emerged. However, it was used with the same verb and was only recorded once: [ni:] for /nist/ 'NEG.be.3S'.

The greatest number of new appearances was recorded at 2;1 for both verbal and non-verbal morphemes. Hediye used three new verbal agreements at 2;1. She produced verbal agreement 3S /-d/ with two different verbs: [bi-jɑ-d] 'SBJV-come-3S' vs. [mi-xɑ-d] 'DUR-want-3S'. Verbal agreement 3S /-e/ also appeared productively: [da:-me] for /dar-mi-zan-e/ 'door-DUR-hit.3S: knocked' vs. [mi-xan-:ah] for /mi-xand-e/ 'DUR-laugh-3S'. Verbal agreement 1S /-am/ used

with four different verbs: [be-xɑb-am] 'SBJV-sleep-1S' vs. [mi-jɑ-m] 'DUR-come-1S'. First use of NEG /na-, ne-/ with a different verb was recorded:

(7.6)	CF:	in-da:-m (2;1)	vs.	ni:s (1;10.4)	and	ni: (2;0.17)
	AT:	ne-mi-dun-am		nis(t)		
		NEG-DUR-know-1S		NEG-be.3S		
		'(I) don't know'		'(it) is not'		

SBJV /be-,bo-/ appeared with two different stems: [bijɑd] 'SBJV.come.3S' vs. [be-xɑb-am] 'SBJV-sleep-1S'. She used PSPT /-e/ with two different stems: [xord-e] 'ate-PSPT.3S' [jixte] for /rixt-e/ 'poured-PSPT.3S'. DUR /mi-/ was attached to seven different verb stems.
In addition, plural /-(h)ɑ/ first appeared with two different nouns at this age: [mu-hɑ-ʃ] 'hair-PL-POSS.3SPC' vs [dɑʃ-a-ʃ-e] for /kafʃ-ɑ-ʃ-e/ 'shoe-PL-POSS.3SPC-be.3S'. Hediye produced the same noun once with definite /-e/ and once without: [go:] for /gorg/ 'wolf' vs. [ɑ gorg-e:] for /ɑɢɑ gorg-e/ 'mr. wolf-DEF'. Finally, 2PPC /-etun/ appeared with one noun only in three different occasions.

The only new morpheme appearing at 2;1.10 was IMPNEG /na-,ne/ with the verb *raftan* 'to go': [na-ro] 'IMPNEG-go.2S'. The first correct use of OM /-o/ was also recorded:

(7.7)	CF:	in-o	dɑʔ-e
	AT:	in-o	dɑr-e
		this-OMhave-3S	
		'(He/she/it) has it'	

Hediye's inflectional errors

This section will focus on Hediye's erroneous use of inflectional morphemes.

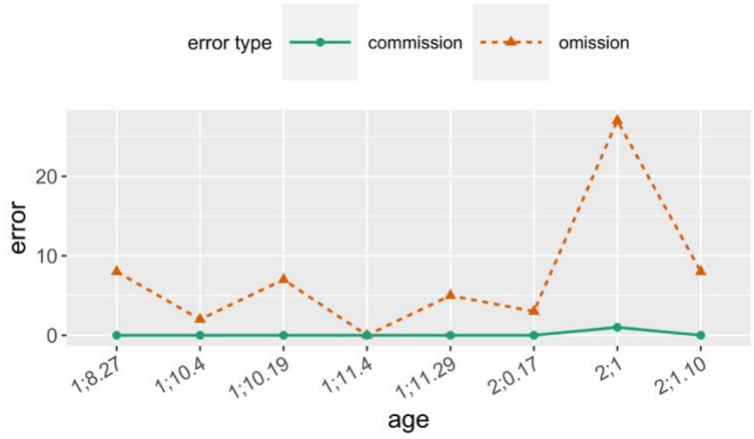

Figure 42 Hediye's omission vs commission errors

Looking at Figure 42, it is apparent that the majority of Hediye's errors are omission errors. In fact, she made only one commission error. This error was made at the age of 2;1. This commission error was the wrong use of OM /-o/ in the following utterance:

(7.8)	CF:	in-o	xub-e
		this-OM	good-be.3S
	AT:	in	xub-e
		this	good-be.3S
		'This is good'	

This is an interesting error because Hediye, as explained in the previous section, only produced two cases of OM /-o/. The first one at the age of 1;11.29 in an imitated form which cannot be considered as a sign of its acquisition. The second one was produced in the following utterance at the age of 2;1.10:

(7.9)	CF:	in-o	dɑʔ-e
	AT:	in-o	dɑr-e
		this-OM	have-3S
		'(He/she/it) has it'	

This successful use of OM /-o/ is followed by a number of omission errors starting at the age of 1;10.19 which continued until the last session of data collection (2;1.10). Most of Hediye's omission errors occur on verbs. Examples of this are [xɑ] for /mi-xɑ-m/ 'DUR-want-2S' or

/mi-xɑ-j/ 'DUR-want-2S'. These child forms miss both the prefix and the suffix. The suffix omission is seen for both 1S and 2S verbal agreements. Hediye's erroneous use of inflectional morphemes can be seen in Figure 43.

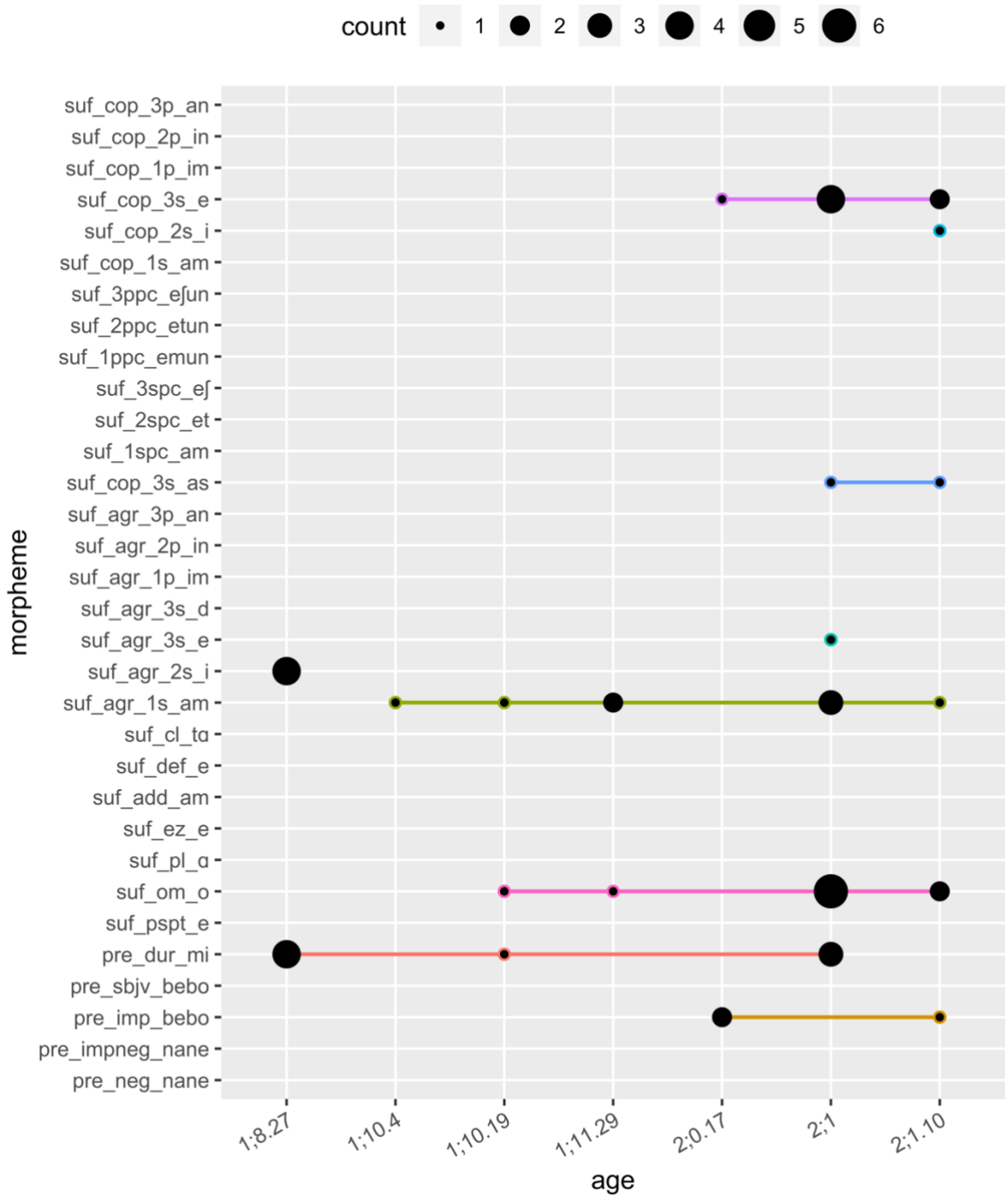

Figure 43 Hediye's erroneous use of morphemes

First appearance vs. contrastive use

To conclude this chapter on Hediye's data, we look at her inflectional morphemes, when they

first appeared versus when she first used those morphemes contrastively (Figure 44). This figure reveals that Hediye's first use of inflectional morphemes is followed by a delay in their contrastive use. For example, PSPT 3S first appeared at 1;8.27 but Hediye first used it contrastively at 2;1. Other cases like copula 3S /-e/, and IMP /be-,bo-/ are used contrastively with a shorter delay. What is striking about Hediye's morphological development, is the big difference between her 7th session (at 2;1) and the rest of the sessions in terms of the number inflectional morphemes recorded. This session recorded the highest number of both new appearances and contrastive use.

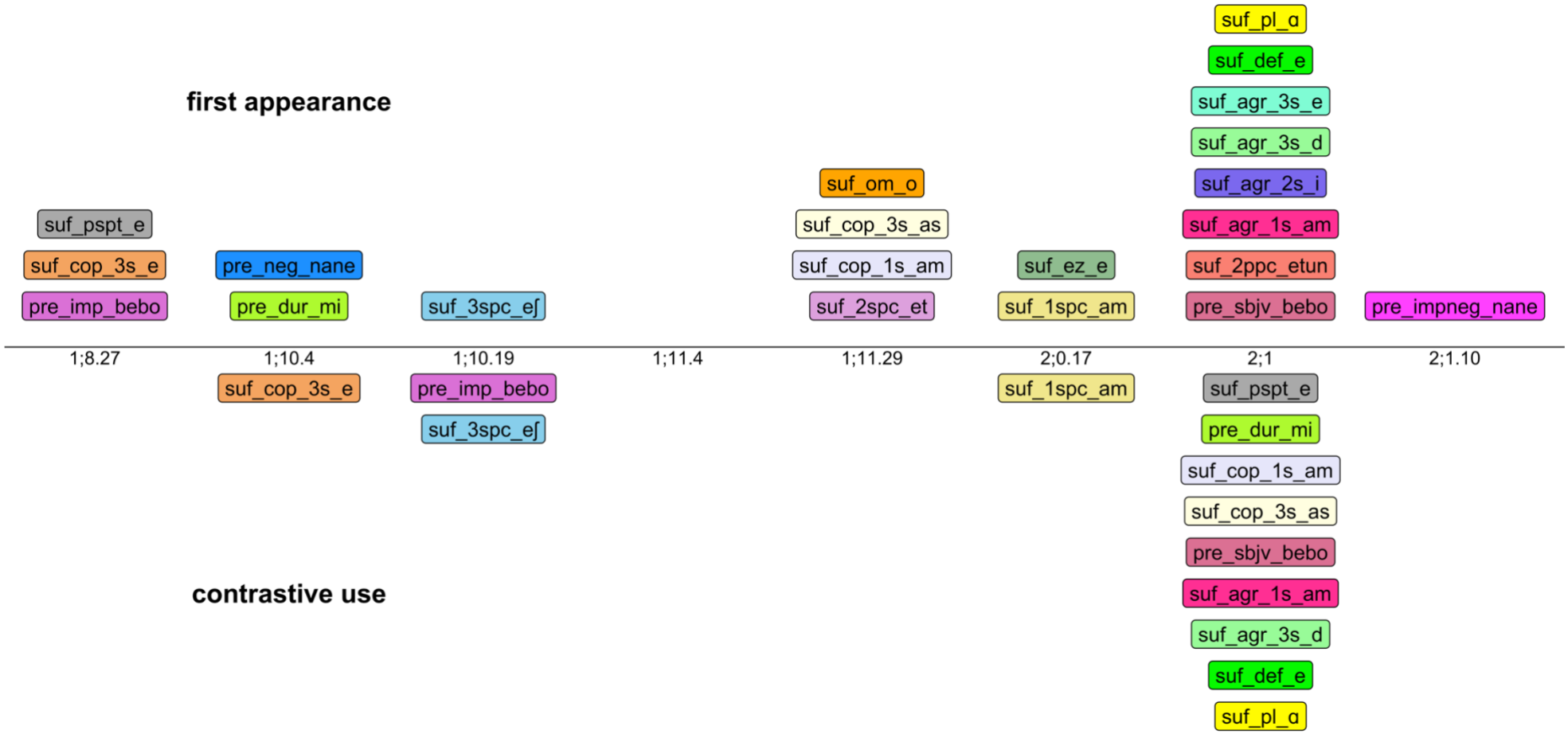

Figure 44 Hediye's morphemes: first appearance vs. contrastive use

Summary

This chapter presented an examination of the linguistic development of a child, referred to as "Hediye". The chapter delves into the unique aspects of Hediye's language development, focusing on her progression from producing single words to producing proper word combinations. The study also examines the timing of Hediye's use of inflectional morphemes, and the accuracy of her use of these morphemes. Data was collected through recording sessions with Hediye from the age of 1;8.27 to 2;1.10. The study found that Hediye began with zero proper word combinations in the first session and progressed to producing 42 proper word combinations in the last session. The study also found that Hediye produced word combinations as single words, a phenomenon that has been observed in previous research on language acquisition. However, it was not observed in the other children in this study and it soon stopped.

8. Data analysis

Statistical analysis

In this thesis, all statistical analysis tests are generated in R (R Code Team, 2019). The Shapiro-Wilk test of normality was performed using the `shapiro.test()` function. The Spearman's correlation coefficient (Spearman, 1987) is calculated using the `ggscatter()` function from the ggplot2 package, `stat_cor()` in the ggpubr package, and `Hmisc::rcorr()` function from the Hmisc package. A linear mixed-effects regression model was fitted using the `lmer()` in the lme4 package. An Anova test was done using `anova()` function to measure the likelihood ratio of the linear mixed-effects model.

To make sure the session length did not affect *count* measures, (i.e., SpnWToken, SpnWType, AllWType, AllWToken, Lem, and WComb) six new measures were introduced. Each *count* measure was divided by the session length (in minutes) to obtain *count per minute* values, resulting in the following: SpnWToken/m, SpnWType/m, AllWType/m, AllWToken/m, Lem/m, and WComb/m.

Normality

Normality testing is a statistical procedure that is used to assess whether a sample of data is drawn from normally distributed population. The normal distribution of data was assessed using the Shapiro-Wilk's test at $p > 0.05$. The Shapiro-Wilk's test confirmed the normal distribution for the Lem ($p = .09$), Lem/m ($p = .3$), pMLU ($p = .9$), PCC-R ($p = .018$), CVar ($p = .35$), and PDSS mean ($p = .06$). The result of this test showed that the rest of the variables are not normally distributed.

Linear mixed-effect regression model

In recent years, mixed models have gained widespread acceptance in the field of linguistics because mixed models are well-suited for handling the types of groupings that are often encountered in linguistic data (Speelman et al., 2018). When data includes grouped observations and the possibility of measurements within the same group being related, group-specific random effects can be included in a regression model to account for these associations (Speelman et al., 2018). Mixed models are currently considered to be one of the

most versatile options for analysing data of this nature (Verbeke et al., 2018). The decision to utilize this specific model was informed by the longitudinal nature of the study, in which subjects were repeatedly assessed at various intervals. This design necessitated the consideration of within-subject measurements, rendering the mixed-effects model an appropriate choice. The utilization of this model facilitates the attainment of more accurate and precise estimates of the fixed effects. Additionally, it enables the examination of both individual-level and group-level effects on the morphological development of the children. Furthermore, the application of a linear mixed-effects model allows for the incorporation of both fixed and random effects within the model, thereby providing a more comprehensive analysis of the data.

The following full model was employed:

$$\text{MorphDevMeasure} \sim \text{LexDevMeasure} + \text{PhonDevMeasure} + \text{SyntDevMeasure} + \text{Age} + (\text{LexDevMeasure} + \text{PhonDevMeasure} + \text{SyntDevMeasure} + \text{Age} | \text{Child})$$

In this study, a mixed-effect linear model was utilized to investigate the association between morphological development and various factors such as phonological development, lexical development, syntactic development, and age. The model employed in this analysis incorporated both fixed and random effects. This approach allows for a more comprehensive examination of the data by accounting for both individual and group-level effects. The fixed effects were the independent variables of phonological development, lexical development, syntactic development, and age, which were all included in the model to explain the variation in morphological development. An ideal approach would have been to use the above-mentioned full maximal model. However, the inclusion of random slopes in this model resulted in convergence issues. As a result, a simplified version of the model was used, which only included a random intercept term.

$$\text{MorphDevMeasure} \sim \text{LexDevMeasure} + \text{PhonDevMeasure} + \text{SyntDevMeasure} + \text{Age} + (1 | \text{Child})$$

Moreover, it would have been valuable to examine the interactions between the language development measures and age. However, it was not possible to include these interactions in the model due to the risk of overfitting and the limited sample size. The age ranges varied by

child in the study, which could have influenced the relationship between the language development measures and age. Therefore, the exclusion of these interactions from the model represents a limitation of the present study.

The random effect was the child, which was included in the model as a random intercept to account for the between-subject variability in the data. This random intercept was included to control for the fact that some children may have different morphological development levels, even when other factors are held constant. This mixed-effect linear model allowed us to examine the unique contributions of each predictor on morphological development while accounting for the inherent variability among children.

It is worth noting that the estimates of the fixed effects are based on the assumption that the predictor variables are independent of each other. If the predictor variables are correlated, the estimates of the fixed effects may be biased. To mitigate this potential issue, we chose not to combine the measures, but rather selected the most appropriate measure in each category. It is therefore important to carefully consider the relationships between the predictor variables when interpreting the results of the model.

In the present study, as stated previously in Chapter 3, a variety of lexical measures were introduced. However, lemma was selected for statistical analysis as it is believed to be the most appropriate measure to reflect the vocabulary size of the child. This decision was based on the assumption that lemma provides a more accurate representation of a child's vocabulary size over time (hence their lexical development) compared to other measures.

In the selection of an appropriate phonological measure, it is crucial to consider various factors and evaluate the strengths and limitations of each option in relation to the specific research question and data being analysed. When determining the optimal phonological measure, it is important to consider the potential overlap between phonological development and morphological development. Specifically, as a child produces more inflectional morphemes, their phonological scores also tend to increase using the measures introduced in this thesis. This may be due to the fact that some inflectional morphemes contain consonants, which can contribute to higher phonological scores. In order to address the overlap between phonological and morphological development, it is necessary to select a measure with minimal overlap. Among the phonological measures discussed in Chapter 3, the Cvar measure exhibits

the least overlap. It is important to reiterate the methodology employed for the calculation of this score. As described in that chapter, the Cvar score is based on a scale of three. A word that contains only vowels, glides, and no true consonants receives a score of one. A word that contains only one true consonant receives a score of two. And a word with more than one true consonant receives a score of three. When considering the inflectional morphemes of Persian, it is important to note that some of these morphemes consist only of a single vowel, which does not affect the Cvar score of the word. Additionally, most of the words produced by children that receive inflectional morphemes already have more than one true consonant, meaning that the presence of a true consonant in the inflectional morpheme does not alter the Cvar score. There are only a few exceptions to this in the data collected, such as the words [pa-m] 'foot-POSS.1S' and [ni-s] 'NEG-be.3S'. Given these considerations, it appears that the Cvar measure is the most appropriate for this analysis.

Returning to our statistical model, to apply the mixed-effect model to the actual measures, we employed the following full model with the following measures:

$$\text{MLUm} \sim \text{Lemm} + \text{Cvar} + \text{MLUw} + \text{Age} + (1 \mid \text{Child})$$

The full model was then compared to the null model using a likelihood ratio test which tests the hypothesis that the fixed effects in the full model are not siginicantly different from zero. The likelihood ratio test was done to measure how well the models fit the data using Anova test. The likelihood ratio test of the models revealed a siginificant different between the two models ($\chi^2(4) = 83.09$, $p < 0.001$). The full model was significantly better at predicting morphological development than the null model. All p values are two-tailed and based on a significance level of 0.001.

The results of the linear mixed-effects model suggest that there is a significant association between the outcome variable (MLUm) and one predictor variable. In particular, the model estimates that the predictor variables Lemm, Cvar, and age are not significantly associated with the outcome, while the predictor variable MLUw ($\beta = 2.048$, $p < 0.001$) is significantly associated with the outcome (see Table 40, Figure 45, Figure 46, Figure 47, and Figure 48).

Table 40 Coefficients and significant tests for the model

Predictor	Coefficient	Std. Error	t-value	P-value
(Intercept)	-0.703991	0.315	-2.232	0.0323*
Lemm	-0.011992	0.080	-0.150	0.8815
Cvar	0.117720	0.115	1.021	0.3145
MLUw	2.048405	0.169	12.116	<0.001***
age	-0.000684	0.001	-1.340	0.189

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

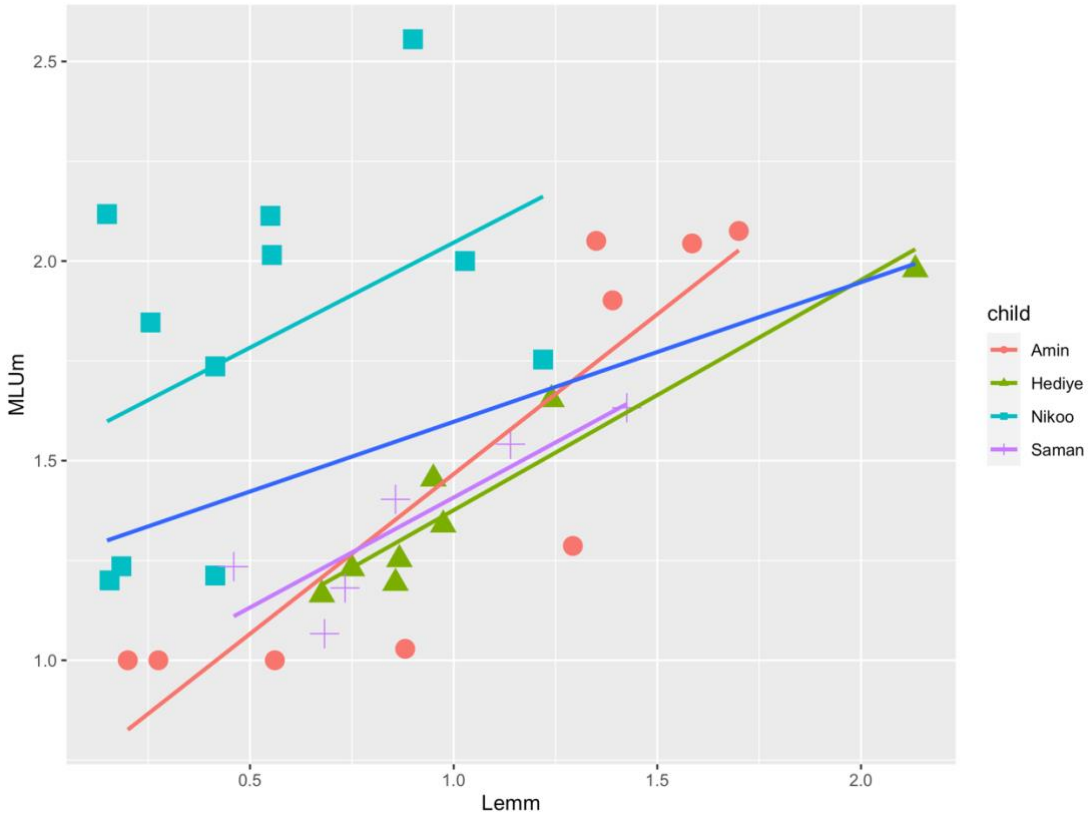

Figure 45 MLUm x Lemm

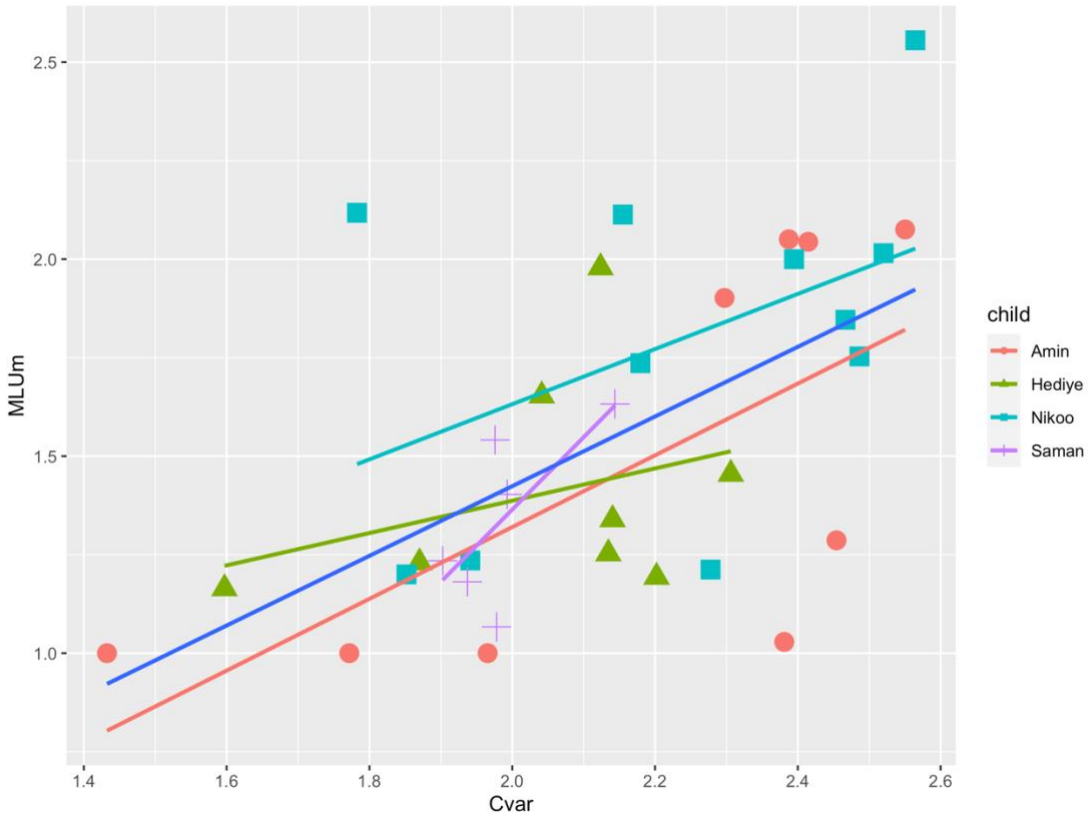

Figure 46 MLUm x Cvar

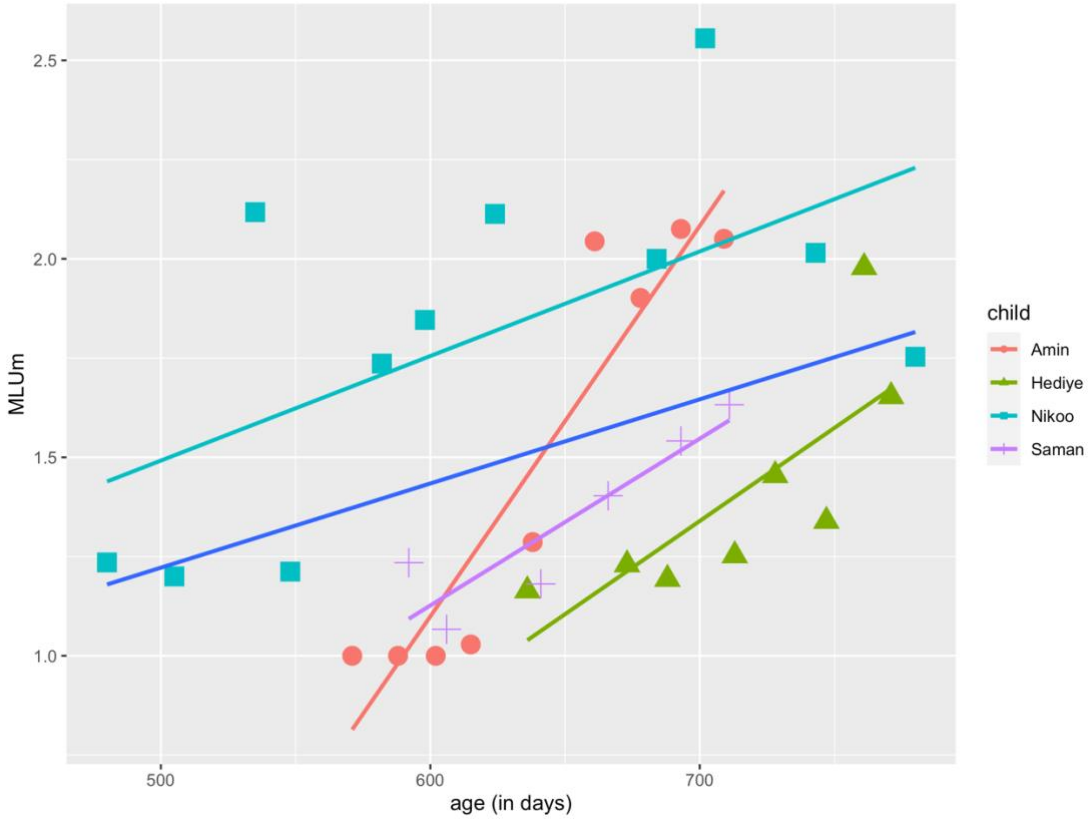

Figure 47 MLUm x Age

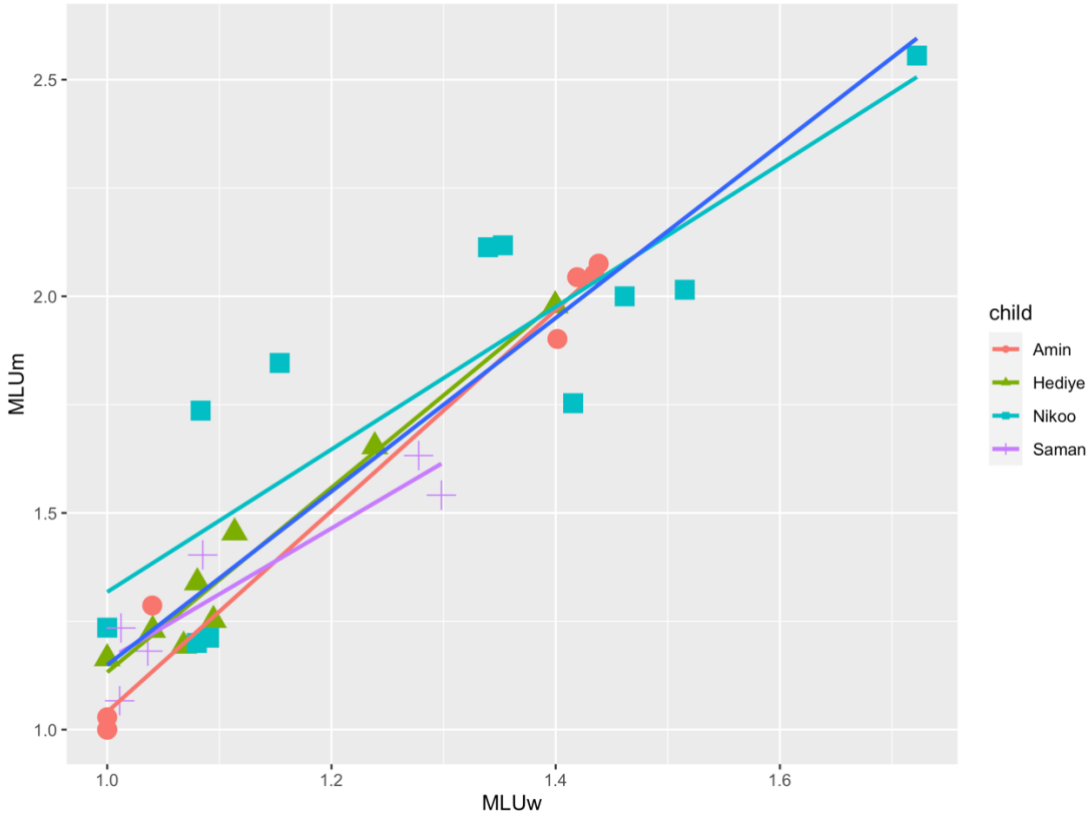

Figure 48 MLUm x MLUw

The estimates of the fixed effects indicate the magnitude and direction of the association between the predictor variables and the outcome. For example, the estimate for the predictor variable MLUw is 2.048, which means that for every unit increase in MLUw, the outcome (MLUm) is expected to increase 2.048 units on average, holding all other variables constant.

The results of the model also indicate the presence of random effects, which suggests that there is variability in the outcome that is not explained by the predictor variables. The variance of the random effects indicates the amount of this unexplained variability, and the standard deviation provides a measure of the dispersion of the random effects.

To sum, the results obtained from the linear mixed-effects model indicate that Lemm, Cvar, and age do not exert a statistically significant influence on the outcome variable (MLUm). Conversely, the predictor variable MLUw demonstrates a statistically significant relationship with the outcome. These findings suggest that MLUw is a more robust predictor of the outcome variable in comparison to Lemm, Cvar, and age. However, it is important to note that the results obtained should be interpreted with caution, as the current study may have

been limited by certain factors such as sample size, measurement error, and confounding variables that may have affected the outcome. This topic will be further discussed in the subsequent chapter, providing a more in-depth examination of the implications and potential avenues for future research.

These findings may have important implications for understanding the factors that influence the outcome and for predicting the outcome in future samples. For example, the estimate for the predictor variable MLUw suggests that an increase in MLUw is associated with an increase in the outcome, while the estimate for the predictor variable Lemm suggests that there is no significant association between Lemm and the outcome.

The presence of random effects in the model indicates that there is unexplained variability in the outcome that is not captured by the predictor variables. This may be due to factors such as measurement error, unmeasured confounding variables, or other sources of variability that are not included in the model. The variance of the random effects provides a measure of the amount of this unexplained variability, and the standard deviation provides a measure of the dispersion of the random effects. Understanding the sources of the random effects may be important for identifying potential sources of bias or for developing more accurate models in the future.

Children's error rate

The following section delves into a comparative analysis of the error patterns exhibited by individual children in the study. The analysis is presented in Figure 49, which is a visual representation of the trends in error rates over time. As Figure 49 shows, there is some resemblance between the overall error patterns of Hediye and Nikoo. Almost each increase in the number of errors is followed by a decrease. What stands out in Hediye's error pattern is it peaks at 2;1. This high error rate could be due to her high production on that session. Amin's error pattern shows a steadier increase followed by a sudden fall, which could be a direct result of less production on the last session. Saman's error pattern peaks halfway through the data collection, followed by a rapid decrease.

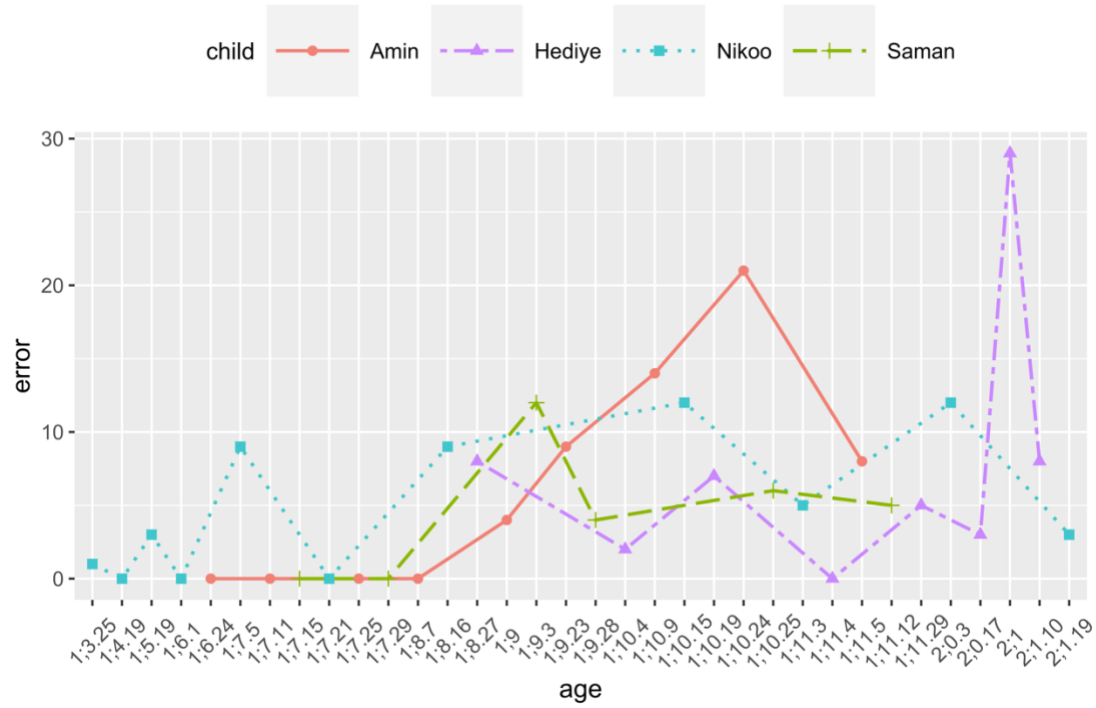

Figure 49 Group error rate

Comparative analysis

This section provides a comparative analysis of the children's scores in order to gain a more comprehensive understanding of the early stages of morphological development in Persian. By examining the similarities and differences between the children, a broader picture of the acquisition process can be formed. Through this analysis, key trends and patterns in the children's language development can be identified and further explored. Additionally, this examination allows for a deeper understanding of the unique characteristics and individual variations within the population of Persian-speaking children studied. The aim of this section is to provide a detailed and nuanced account of the children's morphological development, shedding light on the complex and multifaceted nature of language acquisition.

One unique example of an individual difference is Hediye's word combinations that were produced as one word on the first day at the age of 1;8.27. This phenomenon was not seen in the other children and soon stopped happening. This highlights the importance of considering individual differences when making generalizations about language acquisition.

However, there were also similarities between the children's language development. For example, all of the children produced inflectional morphemes on single words on the final stages of the single-word period prior to starting to produce word combinations. This suggests that inflectional morphemes as seen in other agglutinating languages like Turkish appear early on. Additionally, the existence of words such as /bi-jɑ/ 'IMP-come-2S' and / ʧi-e/ 'what-be.3S' among the early inflected word forms was a common occurrence. The reason for this similarity might be in the semantics of these words or just the fact that they are higher in frequency in the input.

Another similarity was the omission of the final morpheme in words like /mi-x-ɑm/ 'DUR-want-1S' and the use of bare stem without any other inflectional morphemes present, which is an interesting finding that can be looked into in future research. Additionally, the study found that Amin showed evidence of the use of the first person singular pronoun /man/ among his first word combinations despite the fact that Persian is a pro-drop language and the use of pronoun is not necessary. This may suggest the influence of frequency in the input, something that can be looked into in future research.

The following section will examine a phenomenon observed in the productions of these children.

Children's use of bare stems

Another similarity that was observed across all of the children studied in this research was the utilization of bare stem forms without any inflectional prefixes or suffixes, which can be classified as an error of omission. This phenomenon may have implications for understanding the mechanisms underlying language acquisition in Persian-speaking children. Considering the complexity of the inflectional system of Persian verbs, and the fact that bare stems do not surface unmarked (except for 3S simple past), production of such forms are not predicted. This is because children are not exposed to these base forms in Persian (Marvasti, 2014). However, such base stems have been recorded in this study (Table 41). The base stems in the adult target are marked in **boldface**. It is worth noting, however, that the data contained only a total of 16 instances where bare stem forms were used which is not sufficient for conducting any robust statistical analysis.

Table 41 Children's verbal bare stems

Child form	Adult target	Gloss	Child	Age
bih	be-**bin**-am	SBJV-see-1S	Saman	1;9.28
xɑ	mi-**xɑ**-j	DUR-want-2S	Hediye	1;8.27
bih	be-**bin**	IMP-look.2S	Hediye	1;11.29
bih	be-**bin**	IMP-look.2S	Hediye	2;0.17
dɑj	**dɑr**-am	have-1S	Amin	1;9.23
bin	be-**bin**	IMP-look.2S	Hediye	2;0.17
xɑ	mi-**xɑ**-m	DUR want-1S	Nikoo	1;8.16
xɑ	mi-**xɑ**-m	DUR-want-1S	Hediye	2;1
gi:r	be-**gir**	IMP-take.2S	Amin	1;9.23
de	be-**de**	IMP-give.2S	Nikoo	1;7.5
deh	be-**de**	IMP-give.2S	Hediye	2;1.10

The base stems in the adult target are marked in **boldface**.

Upon closer analysis of the children's productions, it was found that only three types of errors of bare stem forms were present. The first type was the omission of inflectional markers in the imperative singular second person. The second type involved the omission of inflectional markers in the first and second singular person of the durative verb *xɑstan* 'to want'. Lastly, the third type of error pertains to the absence of both suffixes and prefixes on subjunctive first person singular.

The study conducted by Savičiūtė and Ambridge, (2018) suggests that when a child is in the process of acquiring a language like Lithuanian, which does not have 'bare' nouns or verbs, and has not yet learned the proper word form or morpheme, they may resort to using a form that they have already learned. This can result in errors. This phenomenon may also be observed in the current study, as the observed errors may be indicative of the child's limited knowledge of the target language's morphological rules.

Summary

The aim of this chapter was to provide a better understanding of the language acquisition process in Persian-speaking children. To accomplish this, a statistical analysis of data collected from four children over a period of several months was conducted. Through this analysis, a number of commonalities and differences in the children's language development were

identified.

One of the key similarities observed was the use of bare stem without any inflectional prefixes or suffixes, which was present in all of the children studied. This finding may suggest that this is a common feature of inflectional morphology acquisition in Persian-speaking children in the early stages and warrants further investigation in future research. Additionally, we noted the presence of certain inflected words among the early inflected word forms, which could be due to the semantics of these words or their frequency in the input.

Furthermore, we noted that all of the children produced inflectional morphemes on single words during the final stages of the single-word period prior to beginning to produce word combinations as predicted.

While these similarities were present, we also identified differences between the children studied. For example, Hediye's word combinations that were produced as one word on the first day at the age of 2;1. This phenomenon was not seen in the other children and soon stopped happening.

In conclusion, this chapter has provided a deeper understanding of the similarities and differences in the language acquisition process of Persian-speaking children. The findings highlight the importance of considering individual differences in language.

9. Discussion

Introduction

This study's first goal was to investigate the development of inflectional morphology in Persian in the earliest stages, qualitatively. The second goal was to investigate the relation of morphological development with phonological advance, lexical development, and syntactic development of the child, quantitatively. To understand the mechanisms of morphological acquisition in Persian certain measures and analyses were introduced and performed. The following questions were attempted to be answered through the course of this research:

1. Can the lexicon size of children influence their morphological acquisition?
2. Can the phonological abilities of children impact their morphological acquisition?
3. Can the child's syntactical development be used as a predictor for their morphological acquisition?

Based on the research questions, the following hypotheses were proposed:

1. There is a correlation between lexical development and morphological development.
2. There is a correlation between phonological development and morphological development.
3. There is a correlation between syntactic development and morphological development.

The study aimed to examine the above-stated hypotheses by gathering naturalistic data from children who had not yet exhibited use of inflectional morphemes or word combinations, as reported by their parents. The data was transcribed and coded, and the linguistic development of each child was thoroughly evaluated and described using various measures. A statistical analysis was then performed to analyse the correlation between the children's lexical, phonological, syntactic, and morphological development. To investigate the connections among these areas, a range of different measures were employed. The research questions were subsequently examined through the application of statistical analysis.

The first chapter provided a background for the study, the second chapter discussed an overview of adult Persian morphology, the third chapter outlined the methodology used in the study. The fourth to seventh chapters presented the analysis of the children's data. The eighth

chapter presented an overall analysis of the results. The ninth chapter, the current chapter, concludes the thesis by offering a discussion, summary, and limitations of the study, as well as suggesting potential avenues for future research.

For the purpose of this investigation four Persian-speaking children's linguistic development were examined. An extensive analysis of their phonological development, lexical development, syntactic development, and morphological development was performed. To get an insight into their phonological development, their consonant inventory at certain points were measured. To be able to follow the children's phonological development whole-word measures like pMLU, PWP, and PWC, and a segmental measure like PCC-R were used. Some of these measures were used as relational measures. On the one hand, pMLU, PWP, PWC, and PCC-R were used because of their relational nature. That is to say, these measures calculate the phonological accuracy of a word with reference to the adult target. On the other hand, CVar, syllable structures, CI size were measured independently (i.e., without reference to the adult target) to better understand the child's articulatory resources and planning.

The number of word types, word tokens, and lemmas were counted on a session-by-session basis to keep track of the children's lexical development. With the beginning of the production of word combinations, the number of those along with MLUw were calculated. Next, their inflectional morphemes were identified. The point of emergence of these morphemes were recorded. Their first contrastive use of inflection was determined. Their productive use of these morphemes was discussed as well. Any errors in using the inflectional morphemes were recorded and analysed. The children's overall morphological development was investigated by the use of two specific measures: PDSS and MLUm.

Given that the morphological structure of Persian is primarily agglutinating, we anticipated that the first instances of inflected forms would occur shortly before or around the 25-word stage. Additionally, it was predicted that during the one-word stage, there would be a presence of non-productive, rote-learned single words with inflectional morphemes followed by a productive period of using inflectional morphemes. It was also expected that the emergence of inflected single words would be observed shortly before the onset of the two-word stage in this study.

It was further proposed that the acquisition of inflectional morphemes in Persian by language

learners would exhibit a U-shaped developmental pattern. Specifically, it was expected that initial progress in acquiring inflectional morphemes would be rapid, followed by a period of relative stagnation, and ultimately culminating in a second period of accelerated acquisition. It was expected that the acquisition of inflectional morphemes would follow a predictable sequence, where simpler forms would be acquired before more complex forms. In particular, children were expected to first demonstrate proficiency in using inflectional forms that are more frequent and have higher number of allomorphs, before gradually progressing to the productive use of less frequent and complex inflectional forms. It was hence predicted that, in the process of language acquisition, individuals would acquire simpler inflectional morphemes before those with strong allomorphs.

Error analysis

It was shown that the children had a low error rate. Children's overall low error rates have been recorded in the literature for different languages as well. However, detailed analysis reveals that the low error rate is not a sign of the child's mastery of inflectional morphology. This low error rate can be attributed to a number of causes. Firstly, early inflected words in the child's production might be rote-learnt as a whole. Frozen phrases frequent in the input might be reproduced by the child later and taken as evidence of correct use of inflectional morphemes. Another reason might be due to the typology of Persian. As mentioned in Chapter 2, Persian is a pro-drop language. This means that the subject of the verb can be dropped. In the absence of the subject, determining the obligatory context of the subject-verb agreement becomes very difficult or even impossible. This results in depending on non-linguistic cues only for determining the obligatory context. In the absence of solid linguistic cues, the inflectional morpheme in question might be deemed correct by mistake.

Previous studies have shown that children are more likely to make errors of omission than commission (Szreder et al., 2021). However, Marvasti (2014)'s study demonstrated that the majority of the errors made by the Persian-speaking children in her study were commission. She concluded that her results support the claim that in morphologically rich languages "morphological development should be conceived of as the acquisition of the ability to REPLACE grammatical morphemes according to the rules of the language rather than the ability to ADD them to the basic forms when required" (Smoczynska, 1985: pp 596-598). Marvasti (2014)'s study also showed that Persian speaking children do not produce bare stem verbs. Her results were reported to be in line with other studies on morphologically rich

languages like Spanish and Italian. This is while it was shown in this study that such bare stems are actually produced by children. At first glance, the results obtained here and those of Marvasti seem contradictory regarding two topics: 1) the rate of omission to commission errors and 2) the appearance of bare uninflected stems in the child production. However, the children in Marvasti's study were at a later stage in their morphological development compared to the children reported here. In fact, comparing the two studies provides a better picture of the child's morphological development, in which the child does use verbal bare stems leaving out certain inflectional morphemes (i.e., making more omission errors) in the earliest stages.

The errors observed in this study demonstrate that children exhibit overgeneralization of inflectional forms, as predicted. What was not anticipated in this study was the utilization of illegal forms (i.e., bare stems) without any inflectional forms as stated above. This finding deviates from the predictions and warrants further investigation. It is possible that this phenomenon occurs as a result of an explanation put forth by Savičiūtė and Ambridge in their 2018 article. If a child has not yet learned the proper word form or morpheme in a language like Lithuanian that does not have 'bare' nouns or verbs, their only option besides saying nothing is to use one that they have learned, which often leads to errors (Savičiūtė & Ambridge, 2018).

Our findings provide support for the results previously reported by Bates and Goodman, (1997) in their study, which demonstrate that age is not a strong predictor of vocabulary or grammar skills within the 16-30-month age range, according to the results of their study involving a large group of typically developing English-speaking children. This confirmation of the previous research highlights the consistency and robustness of the relationship between age and vocabulary or grammar skills within the specified age range. Furthermore, this replication of results lends credibility to the generalizability of the findings, suggesting that the relationship observed is not specific to the sample or context of the original study. Age is not a strong predictor of vocabulary or grammar skills within the 16-30-month age range, according to the results of this study involving a large group of typically developing English-speaking children (Bates & Goodman, 1997).

The results of this research suggest that generative accounts are problematic because they predict an early error-free performance, while constructivist accounts are problematic because

they predict errors and maintain that the patterns of incorrect use of inflections are directly reflective of the input to which the child has been exposed.

In this research, it was observed that the majority of commission errors were near-misses in which only one feature was incorrect. Further analysis is required to determine whether this phenomenon can be attributed to analogy with frequent forms, and whether this is driven by the frequency of the target form in the input. It was also noted that children had a high number of omission errors, which resulted in the formation of morphologically illegal forms potentially indicating incorrect analogical formation. A more detailed examination revealed that a significant number of errors (i.e., using the illegal bare stem) could not be easily explained by either of the mechanisms mentioned earlier in this thesis. These errors likely result from the complex interaction of frequency, meaning, and phonological-related factors that cannot be easily explained by the current theories.

It is worth considering that naturalistic data may not be the most effective method for collecting error samples and analysing error patterns in children. This is because according to Savičiūtė & Ambridge (2018) it is unclear whether the child's early use of language, as seen in the corpus data, is due to rote learning or if it is just because the data did not capture all of the child's less frequent noun forms.

The acquisition of Persian morphology presents a significant challenge for children, particularly in the early stages of language development. To mitigate this challenge, children may utilize various strategies to simplify the task of mastering the intricate relationship between stem forms and affixes, which is further complicated by the presence of allomorphy in the Persian language. As evidenced in previous research studies, one such strategy may involve the production of stems without any prefixes or suffixes.

Conclusion

In summary, the findings of this research provide an insight into the earliest stages of morphological development of children in Persian. Based on these results, it seems that the child initially uses rote-learned inflected forms, such as /bijɑ/ 'IMP-come.2S', where the prefix is not easily recognizable, and then progresses to using inflectional forms more productively by the means of phonological analogy. For a word such as /bijɑ/ 'IMP-come.2S', it appears

that rote-learning or whole-word cognition is at play, as the word is stored as an unanalysable whole and no rule is applicable. Conversely, for a word like /ne-mi-x-ɑm/ 'NEG-DUR-want.1S', it appears that the only means by which a child could store it would be through morphological analysis.

These findings are consistent with the constructivist model of Pre-/Proto-morphology in which the child starts with rote-learnt inflected forms in the pre-morphology period and start using these inflectional morphemes contrastively which results in their productive use in the proto-morphology period. They are also in line with Marvasti (2014)'s study on the piecemeal fashion and gradual process of morphological development in Persian.

Moreover, the data presented in this study strongly supports the existence of a significant correlation between morphological and syntactic growth within the specified age range. However, the research failed to demonstrate a relationship between morphological development and phonological development, morphological development and lexical development, and morphological development and age.

Limitations

The analyses in this thesis have provided an opportunity to look into the first steps of inflectional morphology acquisition in Persian. However, because of the nature of the data collection used, and due to the individual differences between children, we are yet to have a very clear picture of the earliest stages of inflectional morphology acquisition in Persian. Despite the fact that there was strong evidence of a strong relation between morphological measures with lexical and representational measures, and a partial relation with phonological measures, it is still not clear which of these resources and sub-field knowledge contribute to the child's morphological development.

One of the limitations was due to the nature of the data collection. If followed for a longer period of time, the children would have potentially been able to show evidence of productive use for all the inflectional morphemes studied here. Also, more frequent recordings could potentially pinpoint the first emergence versus contrastive use of those inflectional morphemes that first emerged and used contrastively at the same time.

Although this thesis has managed to capture child production at a very young age situated between the one-word stage and the multi-word stage, it has been limited regarding the parental input. It would be illuminating to have parental input to see to what extent distributional characteristics of the children's morphological productions mirror those found in child directed speech for example, the effect of frequency of morphological forms in the input. It could also shed some light on the early error patterns when considered the frequency of the forms in the input. Analysing parental input can provide a great opportunity to have a closer look at the effect of child directed speech on the child's morphological development thus future research could investigate the role of child directed speech in the children's morphological development.

In order to investigate the lack of correlation between vocabulary size and morphological development in children, it is necessary to examine the possibility that the measures and/or methods utilized in this study may have been a contributing factor. According to Marchman and Bates (1994) records of free speech might not provide an accurate representation of the increase in vocabulary for children. The vocabulary size of children up to 30 months of age can be accurately determined through parental reports (Bates & Goodman, 1997). The measure used in this research may only best indicate production level during a specific session and may not accurately reflect the size of the vocabulary. It is therefore suggested that utilizing parental reports through the use of the CDI during the data collection process could have been used as an additional method for determining vocabulary size in children. However, due to the limitation of only administering the CDI to parents prior to the initiation of data collection, it was not possible to utilize this potentially valuable resource in the current study.

Directions for future research

The effect of homophony in inflectional morphology in Persian has not been analysed. Future research can focus on classifying the early inflectional morphemes based on their homophonous nature and provide an analysis based on their similarities (which are not accidental on most cases). This means that for instance IMPNEG and NEG can be categories together under the category NEGATION. Same goes for personal suffixes and copula enclitics, under the category of PERSON/NUMBER.

Another point to consider is as we mentioned earlier there is a certain degree of overlap

between phonological development and morphological development. Specifically, as a child produces more inflectional morphemes, their phonological scores also tend to increase using the measures introduced in this thesis. This may be due to the fact that some inflectional morphemes contain consonants, which can contribute to higher phonological scores. To address this overlap when using measures other than Cvar, it may be advisable to only consider the phonological scores of uninflected forms in future research.

Reference

Ackerman, F., & Malouf, R. (2013). Morphological organization: The low conditional entropy conjecture. *Language*, 429–464.

Aksu-Koç, A., & Ketrez, F. N. (2011). Early verbal morphology in Turkish: Emergence of inflections. In *Development of Verb Inflection in First Language Acquisition* (pp. 27–52). De Gruyter Mouton.

Aksu-Koç, A., & Slobin, D. (1985). The acquisition of Turkish. *The Crosslinguistic Study of Language Acquisition*, *1*.

Allen, S. E. M., & Behrens, H. (2019). Insights into understanding human language from children's acquisition of morphology and syntax: a historical and current perspective on central questions in the field. In *Human Language: From Genes and Brains to Behavior.* (pp. 127–145). MIT Press.

Ambridge, B. (2020). Against stored abstractions: A radical exemplar model of language acquisition. *First Language*, *40*(5–6), 509–559.

Ambridge, B., & Lieven, E. V. M. (2011). *Child language acquisition: Contrasting theoretical approaches*. Cambridge University Press.

Ambridge, B., Pine, J. M., Rowland, C. F., Chang, F., & Bidgood, A. (2013). The retreat from overgeneralization in child language acquisition: Word learning, morphology, and verb argument structure. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*(1), 47–62.

Ambridge, B., & Rowland, C. F. (2013). Experimental methods in studying child language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*(2), 149–168.

Audring, J., & Masini, F. (2018). *The Oxford handbook of morphological theory*. Oxford University Press.

Bates, E., Bretherton, I., & Snyder, L. S. (1991). *From first words to grammar: Individual differences and dissociable mechanisms* (Vol. 20). Cambridge University Press.

Bates, E., & Goodman, J. C. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia and real-time processing. *Language and Cognitive Processes*, *12*(5–6), 507–584.

Beeman, W. O. (2005). *Persian, Dari and Tajik in Central Asia*. National Council for Eurasian and East European Research.

Behrens, H. (2008). *Corpora in language acquisition research: History, methods, perspectives* (Vol. 6). John Benjamins Publishing.

Behrens, H. (2009). *Usage-based and emergentist approaches to language acquisition*.

Behrens, H. (2020). From grammatical categories to processes of categorization. *Current*

Perspectives on Child Language Acquisition: How Children Use Their Environment to Learn, *27*, 91.

Bijankhan, M., Sheykhzadegan, J., Bahrani, M., & Ghayoomi, M. (2011). Lessons from building a Persian written corpus: Peykare. *Language Resources and Evaluation*, *45*(2), 143–164.

Bittner, D., Dressler, W. U., & Kilani-Schoch, M. (2003). *Development of verb inflection in first language acquisition: A cross-linguistic perspective*. Walter de Gruyter.

Bloom, L. (1970). Language development: Form and function. *Emerging Grammurs*.

Bod, R., & Cochran, D. (2007). Introduction to exemplar-based models of language acquisition and use. *Proceedings of the ESSLLI Workshop "Exemplar-Based Models of Language Acquisition and Use", ESSLLI*.

Bónová, I., Slancová, D., & Mikulajová, M. (2005). The measurement of phonological acquisition in Slovak. *Poster Presented at the X. International Congress for the Study of Child Language. Berlin, Germany*.

Brown, R. (1973). *A first language: The early stages. Harvard U. Press*.

Carstairs-McCarthy, A. (2005). Basic terminology. In *Handbook of word-formation* (pp. 5–23). Springer.

Cazden, C. B. (1968). The acquisition of noun and verb inflections. *Child Development*, 433–448.

Cotterell, R., Kirov, C., Hulden, M., & Eisner, J. (2019). On the complexity and typology of inflectional morphological systems. *Transactions of the Association for Computational Linguistics*, *7*, 327–342.

Dabir-Moghaddam, M. (1997). *Compound verbs in Persian*.

Darzi, A., & Boroujerdi, H. (2015). Free Word Order of Verbs and its Arguements. *Language Research*, *6*(2), 21–40.

Deen, K. (2009). The morphosyntax interface. *The Cambridge Handbook of Child Language. Cambridge University Press, Cambridge*, 259–280.

Dehghan, M., & Kambuziya, A. K. Z. (2012). A Short Analysis of Insertion in Persian. *Theory & Practice in Language Studies*, *2*(1).

Demuth, K. (1992). *Acquisition of Sesotho. The cross-linguistic study of language acquisition, vol. by D. Slobin, 557–638*. Hillsdale, NJ: Lawrence Erlbaum.

Dressler, W. U. (2010). A typological approach to first language acquisition. *Language Acquisition across Linguistic and Cognitive Systems*, *52*, 109.

Dressler, W. U. (2003). Morphological typology and first language acquisition: Some mutual challenges. *Mediterranean Morphology Meetings*, *4*, 7–20.

Dressler, W. U. (ed. . (1997). *Studies in pre-and protomorphology*. Wien: Verlag der Oester-

reichischen Akademie der Wissenschaften.

Eisenbeiss, S. (2010). Production methods in language acquisition research. *Experimental Methods in Language Acquisition Research*, 11–34.

Engelmann, F., Granlund, S., Kolak, J., Szreder, M., Ambridge, B., Pine, J., Theakston, A., & Lieven, E. (2019). How the input shapes the acquisition of verb morphology: Elicited production and computational modelling in two highly inflected languages. *Cognitive Psychology*, *110*, 30–69.

Eslami, M. (2019). Morphological Structure of Persian Verb in Perfect Tenses. *Researches in Linguistics*, *11*(1), 1–14.

Eslami, M., & Lemjiri, S. A. (2009). *Inflectional structure of word in Persian*.

Fenson, L., Dale, P., Reznick, J. S., Thal, D., Bates, E., Hartung, J., Pethick, S., & Reilly, J. (1993). *The MacArthur communicative development inventories: user's guide and technical manual (San Diego, Singular/Thompson Learning)*.

Ferguson, C. A. (1957). Word stress in Persian. *Language*, *33*(2), 123–135.

Ferguson, C. A., & Farwell, C. B. (1975). Words and sounds in early language acquisition. *Language*, 419–439.

Gervain, J. (2022). The Early Acquisition of Morphology in Agglutinating Languages: The Case of Hungarian. *A Life in Cognition*, 109–123.

Ghaderniya, N., Alizadeh, M., Jalilevand, N., Jalaei, S., & MollaAghajani, S. (2019). Language Sample Length Effect on Mean Length of Utterance in Morpheme in Normal Persian Children. *Journal of Paramedical Sciences & Rehabilitation*, *8*(1), 17–25.

Ghelmanipour, M. (2006). *The study of some morpho-syntactic features of language development based on LARSP in 1:6 to 2:6 year old children in Tehran*. University of Welfare and Rehabilitation Sciences.

Ghomeshi, J. (1997). Non-projecting nouns and the ezafe: construction in Persian. *Natural Language & Linguistic Theory*, *15*(4), 729–788.

Ghomeshi, J. (2003). Plural marking, indefiniteness, and the noun phrase. *Studia Linguistica*, *57*(2), 47–74.

Ghomeshi, J. (2020). The additive particle in Persian. In R. K. Larson, S. Moradi, & V. Samiian (Eds.), *Advances in Iranian Linguistics* (Vol. 351, p. 57). John Benjamins Publishing Company.

Granlund, S., Kolak, J., Vihman, V., Engelmann, F., Lieven, E. V. M., Pine, J. M., Theakston, A. L., & Ambridge, B. (2019). Language-general and language-specific phenomena in the acquisition of inflectional noun morphology: A cross-linguistic elicited-production study

of {Polish}, {Finnish} and {Estonian}. *Journal of Memory and Language*, *107*, 169–194.

Gruber, F. A. (1999). Probability estimates and paths to consonant normalization in children with speech delay. *Journal of Speech, Language, and Hearing Research*, *42*(2), 448–459.

Guijarro-Fuentes, P., Larrañaga, M. P., & Clibbens, J. (2008). *First language acquisition of morphology and syntax: Perspectives across languages and learners* (Vol. 45). John Benjamins Publishing.

Hamedani, L. (2011). *The function of number in Persian*. University of Ottawa (Canada).

Haresabadi, F., Shirazi, T. S., Ebadi, A., Kazemi, M. D., Anaraki, Z. G., & Shahmahmood, T. M. (2018). Identification of high-frequency morphosyntactic structures in Persian-speaking children aged 4-6 years: A qualitative research. *Iranian Journal of Medical Sciences*, *43*(4), 386.

Hedberg, N., Görgülü, E., & Mameni, M. (2009). On definiteness and specificity in Turkish and Persian. *Proceedings of the 2009 Annual Meeting of the Canadian Linguistic Association*.

Hincha, G. (1961). *Beiträge zu einer Morphemlehre des Neupersischen*.

Ingram, D. (2002). The measurement of whole-word productions. *Journal of Child Language*, *29*(4), 713–733.

Jalilevand, N. (2011). *Roshde goftaro zaban dar kudakane Farsi[language and speech development in Persian speaking children]*. Danzheh.

Jalilevand, N. (2017). *Tahlil nemune goftar kudakane Farsi zaban ba ravesh emtiazdehi roshde jomle [speech sample analysis using the Persian developmental sentence scoring]*. Setayesh Hasti.

Jalilevand, N., Ebrahimipur, M., & Purqarib, J. (2012). Mean length of utterance and grammatical morphemes in speech of two Farsi-speaking children. *Audiology*, *21*(2), 96–108.

Jalilevand, N., Kamali, M., Modarresi, Y., & Kazemi, Y. (2016). The Persian developmental sentence scoring as a clinical measure of morphosyntax in children. *Medical Journal of the Islamic Republic of Iran*, *30*, 435.

Jasbi, M. (2020). Some unique semantic properties of Persian. In *The Routledge Handbook of Second Language Acquisition and Pedagogy of Persian* (pp. 134–146). Routledge.

Kahnemuyipour, A. (2003). Syntactic categories and Persian stress. *Natural Language & Linguistic Theory*, *21*(2), 333–379.

Kahnemuyipour, A. (2014). Revisiting the Persian Ezafe construction: A roll-up movement analysis. *Lingua*, *150*, 1–24.

Kalbasi, I. (1992). *Saxte eshteghaghi-ye vazhe dar Farsi-ye emruz [Word derivational structure in Modern Persian]*. Moassese-ye Motaleat and Tahghighat-e Farhangi.

Kalbasi, I. (2008). The derivational structure of word in modern Persian. *Tehran: Institute for Humanities and Cultural Studies*.

Karimi, S. (1994). Word-order variations in contemporary spoken Persian. *Persian Studies in North America*, 43–73.

Karimi, S., & Brame, M. (2012). A generalization concerning the EZAFE construction in Persian. *Linguistic Analysis*, *38*(1), 111–144.

Kazemi, Y. (2013). *Clinical assessment of Persian-speaking children with language impairment in Iran: exploring the potential of language sample measures*. Newcastle University.

Kazemi, Y., Nakhshab, M., & Maleki, T. (2016). *Persian child language assessment battery, 18 to 30 months (emerging stage)*.

Kazemi, Y., Nematzadeh, S., Hajian, T., Heidari, M., Daneshpajouh, T., & Mirmoeini, A. (2008). The validity and reliability coefficient of Persian translated McArthur-Bates Communicative Development Inventory. *Journal of Research in Rehabilitation Sciences*, *4*(1).

Kazemi, Y., Stringer, H., & Klee, T. (2015). Study of child language development and disorders in Iran: A systematic review of the literature. *Journal of Research in Medical Sciences: The Official Journal of Isfahan University of Medical Sciences*, *20*(1), 66.

Kazemi, Y., Taheri, A., Kiyanfar, F., Shafie, M., Eslamifard, R., Pirmoradian, M., Rezaei, Z. H., Daneshi, L., Delavari, A., & Zareei, M. (2012). Mean length of utterance (MLU) in typically-developing 2.5-5.5 year-old Persian-speaking children in Iran. *Journal of Research in Rehabilitation Sciences*, *8*(5), 928–937.

Kelly, B., Wigglesworth, G., Nordlinger, R., & Blythe, J. (2014). The acquisition of polysynthetic languages. *Language and Linguistics Compass*, *8*(2), 51–64.

Klee, T., & Fitzgerald, M. D. (1985). The relation between grammatical development and mean length of utterance in morphemes. *Journal of Child Language*, *12*(2), 251–269.

Laalo, K. (2009). Acquisition of case and plural in Finnish. *Development of Nominal Inflection in First Language Acquisition: A Cross-Linguistic Perspective*, 49–89.

Ladányi, E., Kovács, Á. M., & Gervain, J. (2020). How 15-month-old infants process morphologically complex forms in an agglutinative language? *Infancy*, *25*(2), 190–204.

Lee, L. L. (1974). *Developmental sentence analysis: A grammatical assessment procedure for speech and language clinicians*. Northwestern University Press.

Leonard, L. B., Caselli, M. C., & Devescovi, A. (2002). Italian children's use of verb and noun morphology during the preschool years. *First Language*, *22*(3), 287–304.

Lignos, C., & Yang, C. (2016). Morphology and language acquisition. *The Cambridge Handbook of Morphology*, 743764.

Macken, M. A., & Barton, D. (1979). *The Acquisition of the Voicing Contrast in Spanish: A Phonetic and Phonological Study of Word-Initial Stop Consonants. Papers and Reports on Child Language Development, No. 16*.

Mahmoodi-Bakhtiari, B. (2018a). Morphology. In *The Oxford Handbook of Persian Linguistics* (pp. 273–299). Oxford University Press Oxford.

Mahmoodi-Bakhtiari, B. (2018b). Spoken vs. written Persian: Is Persian diglossic? *Trends in Iranian and Persian Linguistics*, *313*, 183.

Mahootian, S. (1997). *Persian*. Routledge.

Majidi, S., & Mirdehghan, M. (2021). Agglutinative-Analytic Morphology of Persian: A Distributed Morphology Approach. *Language Research*, *11*(2), 147–172.

Marchman, V. A., & Bates, E. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language*, *21*(2), 339–366.

Mariscal, S., & Gallego, C. (2012). The relationship between early lexical and grammatical development in Spanish: Evidence in children with different linguistic levels. *The Spanish Journal of Psychology*, *15*(1), 112–123.

Marvasti, P. (2014). *The acquisition of verbal morphology in Persian*. University of York.

Meshkato-Dini, M. (2004). Inflectional morphology and structure links in child developing language. *Journal of Linguistics*, *19*(2), 10–26.

Meshkato-Dini, Mehdi. (2008). Persian grammar: the lexical categories and merge. *Tehran: Samt*.

Miller, J. F., & Chapman, R. S. (1981). The relation between age and mean length of utterance in morphemes. *Journal of Speech, Language, and Hearing Research*, *24*(2), 154–161.

Miyata, S., MacWhinney, B., Otomo, K., Sirai, H., Oshima-Takane, Y., Hirakawa, M., Shirai, Y., Sugiura, M., & Itoh, K. (2013). Developmental sentence scoring for Japanese. *First Language*, *33*(2), 200–216.

Nakipoğlu, M., Uzundağ, B. A., & Sarigül, Ö. (2022). Young minds' quest for regularity: Evidence from the Turkish causative. *Journal of Child Language*, *49*(6), 1214–1241.

Navab Safavi, Z., Fallahi, M. H., & Ghadimi Fomani, M. (2020). Acoustic Analysis of Glottal Stop Occurrence before Initial Vowels in Persian Words. *Language Science*, *7*(11), 197–234.

Newmeyer, F. J. (2021). Chomsky and Usage-Based Linguistics. *A Companion to Chomsky*, 287–304.

Oflazer, K. (1994). Two-level description of Turkish morphology. *Literary and Linguistic Computing*, *9*(2), 137–148.

Oryandi-Zanjani, M. M., Ghorbani, R., & Keikha, F. (2006). *Standardization of total numbers of*

word, mean length of utterance and mean length of 5 long sentences in normal Persian language children between 2 to 5 years old in Semnan city.

Paster, M. (2014). 219Allomorphy. In R. Lieber & P. Štekauer (Eds.), *The Oxford Handbook of Derivational Morphology* (p. 0). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199641642.013.0013

Penke, M. (2012). *The acquisition of inflectional morphology*. https://doi.org/10.13140/RG.2.2.21194.06085

Perry, J. R., & Kaye, A. S. (2007). Persian morphology. *Morphologies of Asia and Africa*, *2*, 975–1019.

Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency. *Frequency and the Emergence of Linguistic Structure*, *45*(137), 10–1075.

Pizzuto, E., & Caselli, M. C. (1994). The acquisition of Italian verb morphology in a cross-linguistic perspective. *Other Children, Other Languages: Issues in the Theory of Language Acquisition*, *137*, 187.

Plunkett, K., & Marchman, V. (1993). From rote learning to system building: Acquiring verb morphology in children and connectionist nets. *Cognition*, *48*(1), 21–69.

Pouladi, S., & Khoddam, A. (2003). Mean Length of Utterance in four to five years old Children. *Proceedings of the 7th Iranian Congress on Speech Therapy*, 20–22.

R Core, T. (2019). *R: A language and environment for statistical computing*. Vienna, Austria.

Räsänen, S. H. M., Ambridge, B., & Pine, J. M. (2016). An elicited-production study of inflectional verb morphology in child Finnish. *Cognitive Science*, *40*(7), 1704–1738.

Rasekh-Mahand, M. (2011). The rise of agreement; The case of Persian enclitics. *2011 International Conference on Languages, Literature, and Linguistics*, 530–534.

Rowland, C. F., Fletcher, S. L., & Freudenthal, D. (2008). How big is big enough? Assessing the reliability of data from naturalistic samples. In *Corpora in Language Acquisition Research*. John Benjamins.

Rytting, C. A. (2002). 5. Testing Analogical Modeling: The/k/~ Ø alternation in Turkish. In *Analogical Modeling* (pp. 123–137). John Benjamins.

Saaristo-Helin, K. (2009). Measuring phonological development: A follow-up study of five children acquiring {Finnish}. *Language and Speech*, *52*(1), 55–77.

Sadeghi, A. A. (2002). *Masaele tarikhiye zabane farsi [The historical problems of Persian language]*. Sokhan Publishing.

Samadi, H. (1997). *The acquisition of Persian: grammatically-based measures for assessing normal and abnormal Persian language development*. University of Sheffield.

Samadi, H., & Perkins, M. (2012). 10 {{Persian}}: Devising the P-LARSP1. *Assessing Grammar: The Languages of LARSP*, *7*, 167.

Samadi, H., & Perkins, M. R. (1998). P-LARSP: A developmental language profile for Persian. *Clinical Linguistics & Phonetics*, *12*(2), 83–103.

Samareh, Y. (1977). *The arrangement of segmental phonemes in Farsi*. University of Tehran Press.

Samareh, Y. (1990). Tahlili bar rade shenasie zaban: Vijegihaye radeshenakhti zabane Farsi [A typological analysis of the characteristics of Persian]. *Zabanshenasi [Linguistics]*, *7*(1), 61–80.

Sato, Y., & Karimi, S. (2016). Subject-object asymmetries in Persian argument ellipsis and the anti-agreement theory. *Glossa: A Journal of General Linguistics*, *1*(1).

Savičiūtė, E., & Ambridge, B. (2018). The roles of word-form frequency and phonological neighbourhood density in the acquisition of Lithuanian noun morphology. *Journal of Child Language*, *45*(3), 641–672.

Seidenberg, M. S., & Gonnerman, L. M. (2000). Explaining derivational morphology as the convergence of codes. *Trends in Cognitive Sciences*, *4*(9), 353–361.

Shriberg, L. D., Austin, D., Lewis, B. A., McSweeny, J. L., & Wilson, D. L. (1997). The percentage of consonants correct (PCC) metric: Extensions and reliability data. *Journal of Speech, Language, and Hearing Research*, *40*(4), 708–722.

Sloetjes, H., & Wittenburg, P. (2008). Annotation by category-ELAN and ISO DCR. *6th International Conference on Language Resources and Evaluation (LREC 2008)*.

Smoczynska, M. (1985). The acquisition of Polish. In D. Slobin (Ed.), *The crosslinguistic study of language acquisition* (Vol. 1, pp. 596–686). Lawrence Erlbaum.

Sorahi, M. A., & Alinezhad, B. (2013). A Typological Study of Clitics in Persian. *Journal of Linguistics & Khorasan Dialects*, *5*(8), 103–130.

Spearman, C. (1987). The proof and measurement of association between two things. *The American Journal of Psychology*, *100*(3/4), 441–471.

Speelman, D., Heylen, K., & Geeraerts, D. (2018). *Mixed-effects regression models in linguistics*. Springer.

Spencer, A., & Zwicky, A. M. (2001). *The handbook of morphology*. Wiley-Blackwell.

Staji, A., Namvar Fargi, M., & Keramati Yazdi, S. (2011). *An acoustic analysis of glottal stop and the study of the possibility of existing two adjacent vowel in successive syllables in rapid speech in Persian*.

Stephany, U., & Voeikova, M. D. (2009). *Development of nominal inflection in first language acquisition: A cross-linguistic perspective* (Vol. 30). Walter de Gruyter.

Stoel-Gammon, C, & Stone, J. R. (1991). Assessing phonology in young children. *Clinics in*

Communication Disorders, *1*(2), 25–39.

Stoel-Gammon, Carol. (1989). Prespeech and early speech development of two late talkers. *First Language*, *9*(6), 207–223.

Stolt, S., Haataja, L., Lapinleimu, H., & Lehtonen, L. (2009). Associations between lexicon and grammar at the end of the second year in Finnish children. *Journal of Child Language*, *36*(4), 779–806.

Szreder, M., De Ruiter, L. E., & Ntelitheos, D. (2021). Input effects in the acquisition of verb inflection: Evidence from Emirati Arabic. *Journal of Child Language*, 1–30.

Taelman, H., Durieux, G., & Gillis, S. (2005). Notes on Ingram's whole-word measures for phonological development. *Journal of Child Language*, *32*(2), 391–405.

Thordardottir, E. T., Weismer, S. E., & Evans, J. L. (2002). Continuity in lexical and morphological development in Icelandic and English-speaking 2-year-olds. *First Language*, *22*(1), 3–28.

Tomasello, M. (2003). *Constructing a language : A usage-based account of language acquisition*. Harvard university press.

Verbeke, G., Molenberghs, G., Fieuws, S., & Iddi, S. (2018). Mixed models with emphasis on large data sets. *Mixed-Effects Regression Models in Linguistics*, 11–28.

Vihman, M. M. (2014). *Phonological development: The first two years* (2nd ed.). Wiley-Blackwell.

Vihman, M. M. (2016). Prosodic structures and templates in bilingual phonological development. *Bilingualism: Language and Cognition*, 69–88.

Vihman, M. M. (2019). *Phonological templates in development* (Vol. 4). Oxford University Press, USA.

Vihman, M. M. (2022). The developmental origins of phonological memory. *Psychological Review*.

Vihman, M. M., DePaolis, R. A., & Keren-Portnoy, T. (2009). A dynamic systems approach to babbling and words. In *The Cambridge Handbook of Child Language*. Cambridge University Press.

Vihman, M. M., Ferguson, C. A., & Elbert, M. (1986). Phonological development from babbling to speech: Common tendencies and individual differences. *Applied Psycholinguistics*, *7*(1), 3–40.

Vihman, M. M., Keren-Portnoy, T., Whitaker, C., Bidgood, A., & McGillion, M. (2013). Late talking toddlers: Relating early phonological development to later language advance. *York Papers in Linguistics*, *2*(13), 47–68.

Vihman, M. M., & McCune, L. (1994). When is a word a word? *Journal of Child Language*, *21*(3),

517–542.

Vihman, M. M., & Miller, R. (1988). *Words and babble at the threshold of language: Acquisition.*

von Heusinger, K., & Sadeghpoor, R. (2020). The specificity marker-e with indefinite noun phrases in Modern Colloquial Persian. *Nominal Anchoring*, 115.

Wijnen, F., Kempen, M., & Gillis, S. (2001). Root infinitives in Dutch early child language: an effect of input? *Journal of Child Language*, *28*(3), 629–660.

Windfuhr, G. (2009). *The Iranian Languages*. Psychology Press.

Xanthos, A., Laaha, S., Gillis, S., Stephany, U., Aksu-Koç, A., Christofidou, A., Gagarina, N., Hrzica, G., Ketrez, F. N., Kilani-Schoch, M., & others. (2012). On the role of morphological richness in the early development of noun and verb inflection. *First Language*, *31*(4), 461–479.

Yalçin, P. (1996). *Turkish language and its place among world languages*.

Yazarlou, S. (2014). Glottal stop in hiatus: An acoustic investigation in Persian. *Procedia-Social and Behavioral Sciences*, *136*, 12–20.

Yousefi, S. (2018). *Persian: a comprehensive grammar*. Routledge.

Zarei Mahmood Abadi, M., Zarifian, T., & Teymouri, R. (2021). Survey Development of Compound Sentences and Mean Length of Utterance a Persian-speaking Child From 15 to 34 Months: a Longitudinal Case Report. In *HBI_Journals* (Issue 4, pp. 582-595 K1-Mean length of utterance K1-Pediatr). http://jams.arakmu.ac.ir/article-1-6491-en.html

Appendix

Abbreviations and symbols

ADD	additive
ADJ	adjective
AllWToken	all word tokens
AllWToken/m	all word tokens per minute
AllWType	all word types
AllWType/m	all word types per minute
AT	adult target
CD	child-directed word
CF	child form
CL	number classifier
CVar	consonant variegation score
Dem	demonstrative
DUR	durative
EZ	Ezafe particle
I	inflection
IMP	imperative
INDEF	indefinite
Lem	lemma
Lem/m	lemmas per minute
MLUm	mean length of utterance in morpheme
MLUw	mean length of utterance in words
NEG	negative
O	object
OM	object marker
P	pronoun
PCC-R	percentage consonant correct – revised
PDSS	Persian developmental sentencing score
PL	plural

pMLU	phonological mean length of utterance
POSS	possessive
PP	prepositional phrase
Prep	preposition
PSPT	past participle
PWC	proportion of whole-word correctness
PWP	percentage whole-word proximity
QW	question word
S	subject
SBJV	subjunctive
SpnWToken	spontaneous word tokens
SpnWToken/m	spontaneous word tokens per minute
SpnWType	spontaneous word types
SpnWType/m	spontaneous word types per minute
V	verb
WComb	word combination
WComb/m	word combinations per minute
WToken	word token
WType	word type
1S	first person singular
2S	second person singular
3S	third person singular
1P	first person plural
2P	second person plural
3P	third person plural
1SPC	first person singular pronominal clitic
2SPC	second person singular pronominal clitic
3SPC	third person singular pronominal clitic
1PPC	first person plural pronominal clitic
2PPC	second person plural pronominal clitic
3PPC	third person plural pronominal clitic

-	(hyphen)	morpheme boundary (Only when there is one-to-one correspondence between Persian morphemes and their English gloss)
.	(period)	morpheme boundary (For more than one-to-one correspondence between Persian morphemes and their English gloss)
#	(hashtag)	a short pause

Appendix 1 – social media participant recruitment poster

Figure 50 Social media participant recruitment poster

Appendix 2 – Information Sheet (English): data collection 1

UNIVERSITY of York

Sarvenaz Moradi
DEPARTMENT OF LANGUAGE AND LINGUISTIC SCIENCE
Heslington, York, YO10 5DD, UK
Mobile : (+44)XXXXXX(UK)
(+98)XXXXXX(Iran)
Email : sm1833@york.ac.uk

Participant Information Sheet

You are invited to take part in a research study. Before you decide whether to participate it is important for you to understand why the research is being done and what it will involve. Please take the time to read the following information carefully. If there is anything you do not understand, or if you want more information, please ask the researcher.

What is the research about?

This is a study related to my PhD. I am investigating the early vocabulary of Persian and learning infants and their use of different sound patterns.

Who is carrying out the research?

The research is being carried out by Sarvenaz Moradi, under the supervision of Professor Marilyn Vihman in the Department of Language and Linguistic Science at the University of York.

Who can participate?

The restrictions as to who can participate in this study are as follows:

(i) Infants who produce about 35 to 50 words according to parental report in May 2017.

(ii) The parents are native speakers of Persian.

(iii) The infant is being brought up in a household where only Persian is spoken.

(iv) The infant has no hearing/medical problems.

(v) The infant is full-term.

What does the study involve?

If you agree to participate, I will visit your home and video record you playing with your child. You will not have to do anything special, just interact as usual with your baby. I will visit you every two weeks for four months to make a recording. In this way, your child will participate in 8 video recording sessions over all. Each session will go on for about 40 minutes.

I will ask you to fill out two forms before the first recording session.1) the language background information form and 2) the CDI (Communicative Developmental Inventory). The CDI is a parental report where parents record the words their children know and say.

If you need to stop the session or cancel it due to your baby being unwell, upset, or for any other reason, I will try to rebook the session.

Once I have your recordings, I will transcribe and analyse your speech and that of your child.

Do my child and I have to take part?

You and your child do not have to take part in the study. If you do decide to take part, you will be given this information sheet to keep and will be asked to sign two copies of the consent form (one copy is for you to keep). If you decide to take part, you will still be free to withdraw at any time without giving a reason. If you withdraw from the study, we will destroy your and your child's data in all forms and will not use it in any way.

What are the possible risks of taking part?

The study poses no foreseeable risk to you or your child. If at any stage your child becomes distressed, we will end the session.

Are there any benefits to participating?

You will be participating in an exciting new piece of linguistic research that may help linguists better understand early vocabulary and speech-sound development in Persian-learning infants. There has been no study to date of the early development of vocabulary or speech production in Persian-learning infants. You will be helping me to fill that gap in the field of research and also to publish a study on infants from Iran. There will be no direct benefits to yourself, however.

What kind of information do I have to give?

You and your spouse will be asked to fill out a sheet with your contact details, education level, occupation and language background. The information provided by you will not be revealed in any publication or presentation.

What will happen to the data I provide?

The data (audio and video files) you and your infant provide will be kept safely on a University of York computer. The data will also be backed up to a secure computer. The results will be reported in my PhD thesis, in presentations and in one or more academic papers. The data may also be kept after the duration of the current project, to be used in future research. Your faces may be visible in the recordings. However, you can opt out of allowing your videos to be used in lectures and presentations.

What about confidentiality?

Your and your child's identity will be kept strictly confidential. No real names will be used in any presentations or publications. Your data will be stored securely in the Department of Language and Linguistic Science, University of York, the UK.

Will I know the results?

I may not be able to give you feedback on your child's individual results. However, after the study has been completed, I will email the summary of the results if you provide your email address.

Appendix 3 – Information Sheet (English): data collection 2

UNIVERSITY of York

Sarvenaz Moradi
DEPARTMENT OF LANGUAGE AND LINGUISTIC SCIENCE
Heslington, York, YO10 5DD, UK
Mobile : (+44)XXXXXX(UK)
(+98)XXXXXX(Iran)
Email : sm1833@york.ac.uk

Participant Information Sheet

You are invited to take part in a research study. Before you decide whether to participate it is important for you to understand why the research is being done and what it will involve. Please take the time to read the following information carefully. If there is anything you do not understand, or if you want more information, please ask the researcher.

What is the research about?

This is a study related to my PhD. I am investigating the early vocabulary of Persian and learning infants and their use of different sound patterns.

Who is carrying out the research?

The research is being carried out by Sarvenaz Moradi, under the supervision of Professor Marilyn Vihman in the Department of Language and Linguistic Science at the University of York.

Who can participate?

The restrictions as to who can participate in this study are as follows:

(i) Infants who produce about 35 to 50 words according to parental report in May 2018.

(ii) The parents are native speakers of Persian.

(iii) The infant is being brought up in a household where only Persian is spoken.

(iv) The infant has no hearing/medical problems.

(v) The infant is full-term.

What does the study involve?

If you agree to participate, I will visit your home and video record you playing with your child. You will not have to do anything special, just interact as usual with your baby. I will visit you every two weeks for four months to make a recording. In this way, your child will participate in 8 video recording sessions over all. Each session will go on for about 40 minutes.

I will ask you to fill out two forms before the first recording session.1) the language background information form and 2) the CDI (Communicative Developmental Inventory). The CDI is a parental report where parents record the words their children know and say.

If you need to stop the session or cancel it due to your baby being unwell, upset, or for any other reason, I will try to rebook the session.

Once I have your recordings, I will transcribe and analyse your speech and that of your child.

Do my child and I have to take part?

You and your child do not have to take part in the study. If you do decide to take part, you will be given this information sheet to keep and will be asked to sign two copies of the consent form (one copy is for you to keep). If you decide to take part, you will still be free to withdraw at any time without giving a reason. If you withdraw from the study, we will destroy your and your child's data in all forms and will not use it in any way.

What are the possible risks of taking part?

The study poses no foreseeable risk to you or your child. If at any stage your child becomes distressed, we will end the session.

Are there any benefits to participating?

You will be participating in an exciting new piece of linguistic research that may help linguists better understand early vocabulary and speech-sound development in Persian-learning infants. There has been no study to date of the early development of vocabulary or speech production in Persian-learning infants. You will be helping me to fill that gap in the field of research and also to publish a study on infants from Iran. There will be no direct benefits to yourself, however. Upon taking part in at least 8 session, you will receive £50 as a thank you for taking part in this study.

What kind of information do I have to give?

You and your spouse will be asked to fill out a sheet with your contact details, education level, occupation and language background. The information provided by you will not be revealed in any publication or presentation.

What will happen to the data I provide?

The data (audio and video files) you and your infant provide will be kept safely on a University of York computer. The data will also be backed up to a secure computer. The results will be reported in my PhD thesis, in presentations and in one or more academic papers. The data may also be kept after the duration of the current project, to be used in future research. Your faces may be visible in the recordings. However, you can opt out of allowing your videos to be used in lectures and presentations.

What about confidentiality?

Your and your child's identity will be kept strictly confidential. No real names will be used in any presentations or publications. Your data will be stored securely in the Department of Language and Linguistic Science, University of York, the UK.

Will I know the results?

I may not be able to give you feedback on your child's individual results. However, after the study has been completed, I will email the summary of the results if you provide your email address.

Appendix 4 – Consent Form (English)

UNIVERSITY of York

Sarvenaz Moradi
DEPARTMENT OF LANGUAGE AND LINGUISTIC SCIENCE
Heslington, York, YO10 5DD, UK
Mobile : (+44)XXXXXX(UK)
(+98)XXXXXX(Iran)
Email : sm1833@york.ac.uk

Title of project: Templatic Approach to Persian Child Phonology

Researcher: Sarvenaz Moradi

Consent form for infants

This form is for you to state that you agree to take part in the study together with your child. Please read and answer every question. If there is anything you do not understand, or if you want more information, please ask the researcher.

Have you read and understood the information we have given you about the study?	Yes ❒ No ❒
Have you had an opportunity to ask questions about the study and have these been answered satisfactorily?	Yes ❒ No ❒
Do you understand that the information provided will be held in confidence by the research team, and your or your child's name or identifying information will be withheld or masked in any publication?	Yes ❒ No ❒
Do you understand that you may withdraw your agreement to take part in the study at any time before the video recording session without giving any reason, and that in such a case all your data will be destroyed?	Yes ❒ No ❒
Do you understand that the information you and your child provide may be kept after the duration of the current project, to be used in future research on language?	Yes ❒ No ❒
Do you agree to take part in the study?	Yes ❒ No ❒

Do you agree to your child and yourself being recorded on audio and video in the play recordings? If not, will some other member from your family be playing with your child during the recording sessions?	Yes ❒ No ❒
Do you agree to excerpts from your and your child's audio/video recordings being used in presentations or in teaching by the researcher, without disclosing your or your child's real name? *(You may take part in the study without agreeing to this).*	Yes ❒ No ❒
Do you agree to the researcher's keeping your contact details after the end of the current project, in order that s/he may contact you in the future about possible participation of your child in other studies? *(You may take part in the study without agreeing to this).*	Yes ❒ No ❒

Your name: _____

Your child's name: _____

Email address (Please write your email address if you want to receive a summary of the results):

_____

Your signature: _____

Researcher's name: _____
Researcher's signature:
_____

Date: _____

Appendix 5 – Information Sheet (Persian): Data collection 1

UNIVERSITY of York

Sarvenaz Moradi
DEPARTMENT OF LANGUAGE AND LINGUISTIC SCIENCE
Heslington, York, YO10 5DD, UK
Mobile : (+44)XXXXXX(UK)
(+98)XXXXXX(Iran)
Email : sm1833@york.ac.uk

برگه اطلاعات

شما به شرکت در این پروژه تحقیقاتی دعوت شده اید. قبل از تصمیم گیری در مورد شرکت در این پروژه لازم است از علت انجام گیری این پروژه تحقیقاتی و آنچه قرار است رخ دهد مطلع شوید. لطفا مطالب ذیل را با دقت بخوانید. اگر نکته ای وجود دارد که متوجه آن نمی شوید لطفا حتما از پژوهشگر سوال نمایید.

موضوع پروژه تحقیقاتی چیست؟

این پروژه تحقیقاتی بخشی از تز دکترای من است. پژوهش من بر روی واژه های اولیه کودک و فرم های آوایی مختلفی است که کودک تولید می کند.

پژوهشگر این پروژه کیست؟

این پژوهش توسط سروناز مرادی و تحت نظارت پرفسور مرلین ویمان (Prof. Marylin Vihman) از دپارتمان زبان و علوم زبانشناختی دانشگاه یورک انگلستان انجام می پذیرد.

چه کسی میتواند در آن شرکت کند؟

1. کودکی که فروردین ۱۳۹۶ حدود ۵۰ کلمه بر اساس گزارش والدین تولید کند (یا بین یک و نیم تا دو سال باشد)
2. زبان مادری پدر و مادر فارسی باشد.
3. کودک در خانه ای که تنها فارسی در آن صحبت میشود زندگی کند.
4. کودک هیچ گونه مشکل شنوایی/پزشکی گزارش شده نداشته باشد.
5. کودک ۹ ماهه به دنیا آمده باشد.

پژوهش شامل چه چیزهایی می شود؟

اگر که شما با شرکت کودکتان در این پژوهش موافقت کنید، من به منزل شما خواهم آمد و از بازی شما با کودکتان فیلم و صوت تهیه خواهم کرد. نیازی نیست کار خاصی انجام دهید، صرفا بازی و صحبت معمول شما با کودکتان کافی است. این کار را هر دو هفته یک بار و به مدت ۴ ماه ادامه خواهیم داد. مجموع این جلسات فیلم برداری ۸ جلسه خواهد بود. هر جلسه حدود ۳۰ تا ۴۰ دقیقه به طول خواهد انجامید.
قبل از جلسه اول من از شما می خواهم که دو فرم پر کنید: ۱. فرم پیشینه زبانی ۲. فرم مک آرتور (این فرم گزارش والدین از تعداد کلماتی است که کودک میداند و/یا به زبان می آورد.
اگر به هر دلیلی (مثل احساس ناراحتی توسط کودکتان و غیره) نیاز به متوقف کردن ضبط و یا کنسل کردن جلسه داشتید، ضبط را به جلسه دیگری موکول میکنیم.
بعد از ضبط، من گفتار کودکتان را آوانگاری کرده و آنالیز میکنم.

آیا من و کودکم مجبور به شرکت در این پژوهش هستیم؟

خیر. شما و کودکتان مجبور به شرکت در این پژوهش نیستید. اگر شما تصمیم به شرکت بگیرید، این برگه به شما تحویل داده میشود تا آن را نزد خود نگه دارید. همچین نیاز است دو نسخه برگه رضایت نامه شرکت در پژوهش را امضا کنید و یک نسخه را نزد خود نگه دارید.
اگر تصمیم به شرکت در این پژوهش بگیرید، باز هم این حق را خواهید داشت که در هر مرحله ایی از پژوهش بدون نیاز به ارائه دلیلی از شرکت در آن انصراف نمایید. اگر تصمیم به کناره گیری بکنید، اطلاعات جمع آوری شده از کودک شما، به طور کامل نابود میشود و به هیچ طریقی از آن استفاده نخواهد شد.

آیا شرکت در این پژوهش خطراتی به همراه دارد؟
خیر. شرکت در این پژوهش هیچ خطری برای کودک شما نخواهد داشت. اگر در هر مرحله ایی کودک شما احساس ناراحتی کرد ضبط متوقف خواهد شد.

فوائد شرکت در این پژوهش چیست؟
شما در یک پژوهش زبانشناختی هیجان انگیز شرکت خواهید کرد که به زبانشناسان در درک و تحلیل واژه های اولیه و رشد زبانی کودکان فارسی زبان کمک خواهید کرد. تا کنون هیچ پژوهشی در زمینه تولید گفتار و واژه های اولیه کودک در زبان فارسی به این شکل انجام نشده است. شما به پر کردن این شکاف علمی و گرداوری اطلاعات از کودکان ایرانی کمک خواهید کرد.

اطلاعاتی که در اختیار این پژوهش میگذارید چیست؟
از شما و یا همسرتان خواسته میشود یک فرم مربوط به اطلاعات تماس، میزان تحصیلات تان و شغل و پیشینه زبانی خود و همسر خود پر کنید. این اطلاعات به هیچ نحو و در هیچ کجا در اختیار هیچ کس قرار گرفته نخواهد شد.

بر سر اطلاعاتی که از کودک من جمع میشود چه می آید؟
فایل های صوتی و تصویری که از کودک شما جمع آوری میشود در یکی از کامپیوترهای دانشگاه یورک محفوظ خواهد بود. یک نسخه پشتیبانی هم از آن تهیه و نگهداری خواهد شد. نتایج در رساله دکتری من، ارائه گزارش های شفاهی و یک یا چند مقاله علمی ارائه خواهد شد. فایل های صوتی و تصویری ممکن است بعد از اتمام دوره دکترای من نگهداری و برای پژوهش های آتی مورد استفاده قرار گیرد. صورت شما و کودک شما در فایل های تصویری واضح خواهد بود. با این حال شما قادر به تصمیم گیری در مورد استفاده و یا عدم استفاده از فایل های تصویری در ارائه گزارشات خواهید بود.

چه بر سر هویت من و کودک من میاد؟
هویت شما و کودک شما به جد محفوظ خواهد ماند. اسم واقعی کودک شما به هیچ وجه و در هیچ نوع مقاله و یا ارائه گزارشی استفاده نخواهد شد. اطلاعات شما در زمینه هویت شما و کودکتان در دپارتمان زبان و علوم زبانشناختی دانشگاه یورک انگلستان نگهداری خواهد شد.

آیا از نتایج با خبر خواهم شد؟
ممکن است من نتوانم نتایج فردی کودک شما را در اختیار شما بگذارم. اما پس از اتمام پژوهش، در صورت تمایل شما و ارائه آدرس ایمیل تان میتوانم خلاصه نتایج را در اختیار شما قرار دهم.

Appendix 6 – Information Sheet (Persian): Data collection 2

Sarvenaz Moradi
DEPARTMENT OF LANGUAGE AND LINGUISTIC SCIENCE
Heslington, York, YO10 5DD, UK
Mobile : (+44)XXXXXX(UK)
(+98)XXXXXX(Iran)
Email : sm1833@york.ac.uk

برگه اطلاعات

شما به شرکت در این پروژه تحقیقاتی دعوت شده اید. قبل از تصمیم گیری در مورد شرکت در این پروژه لازم است از علت انجام گیری این پروژه تحقیقاتی و آنچه قرار است رخ دهد مطلع شوید. لطفا مطالب ذیل را با دقت بخوانید. اگر نکته ای وجود دارد که متوجه آن نمی شوید لطفا حتما از پژوهشگر سوال نمایید.

موضوع پروژه تحقیقاتی چیست؟

این پروژه تحقیقاتی بخشی از تز دکترای من است. پژوهش من بر روی واژه های اولیه کودک و فرم های آوایی مختلفی است که کودک تولید می کند.

پژوهشگر این پروژه کیست؟

این پژوهش توسط سروناز مرادی و تحت نظارت پرفسور مرلین ویمان (Prof. Marylin Vihman) از دپارتمان زبان و علوم زبانشناختی دانشگاه یورک انگلستان انجام می پذیرد.

چه کسی میتواند در آن شرکت کند؟

1. کودکی که فروردین ۱۳۹۷ حدود ۵۰ کلمه بر اساس گزارش والدین تولید کند (یا بین یک و نیم تا دو سال باشد)
2. زبان مادری پدر و مادر فارسی باشد.
3. کودک در خانه ای که تنها فارسی در آن صحبت میشود زندگی کند.
4. کودک هیچ گونه مشکل شنوایی/پزشکی گزارش شده نداشته باشد.
5. کودک ۹ ماهه به دنیا آمده باشد.

پژوهش شامل چه چیزهایی می شود؟

اگر که شما با شرکت کودکتان در این پژوهش موافقت کنید، من به منزل شما خواهم آمد و از بازی شما با کودکتان فیلم و صوت تهیه خواهم کرد. نیازی نیست کار خاصی انجام دهید، صرفا بازی و صحبت معمول شما با کودکتان کافی است. این کار را هر دو هفته یک بار و به مدت ۴ ماه ادامه خواهیم داد. مجموع این جلسات فیلم برداری ۸ جلسه خواهد بود. هر جلسه حدود ۳۰ تا ۴۰ دقیقه به طول خواهد انجامید.
قبل از جلسه اول من از شما می خواهم که دو فرم پر کنید: ۱. فرم پیشینه زبانی ۲. فرم مک آرتور (این فرم گزارش والدین از تعداد کلماتی است که کودک میداند و/یا به زبان می آورد.
اگر به هر دلیلی (مثل احساس ناراحتی توسط کودکتان و غیره) نیاز به متوقف کردن ضبط و یا کنسل کردن جلسه داشتید، ضبط را به جلسه دیگری موکول میکنیم.
بعد از ضبط، من گفتار کودکتان را آوانگاری کرده و آنالیز میکنم.

آیا من و کودکم مجبور به شرکت در این پژوهش هستیم؟

خیر. شما و کودکتان مجبور به شرکت در این پژوهش نیستید. اگر شما تصمیم به شرکت بگیرید، این برگه به شما تحویل داده میشود تا آن را نزد خود نگه دارید. همچین نیاز است دو نسخه برگه رضایت نامه شرکت در پژوهش را امضا کنید و یک نسخه را نزد خود نگه دارید.

اگر تصمیم به شرکت در این پژوهش بگیرید، باز هم این حق را خواهید داشت که در هر مرحله ایی از پژوهش بدون نیاز به ارائه دلیلی از شرکت در آن انصراف نمایید. اگر تصمیم به کناره گیری بکنید، اطلاعات جمع آوری شده از کودک شما، به طور کامل نابود میشود و به هیچ طریقی از آن استفاده نخواهد شد.

آیا شرکت در این پژوهش خطراتی به همراه دارد؟

خیر. شرکت در این پژوهش هیچ خطری برای کودک شما نخواهد داشت. اگر در هر مرحله ایی کودک شما احساس ناراحتی کرد ضبط متوقف خواهد شد.

فوائد شرکت در این پژوهش چیست؟

شما در یک پژوهش زبانشناختی هیجان انگیز شرکت خواهید کرد که به زبانشناسان در درک و تحلیل واژه های اولیه و رشد زبانی کودکان فارسی زبان کمک خواهید کرد. تا کنون هیچ پژوهشی در زمینه تولید گفتار و واژه های اولیه کودک در زبان فارسی به این شکل انجام نشده است. شما به پر کردن این شکاف علمی و گرداوری اطلاعات از کودکان ایرانی کمک خواهید کرد.

پس از اتمام دوره، به شرط حضور در کلیه جلسات (۸ جلسه) ۵۰ پوند به عنوان تشکر تقدیم شما خواهد شد.

اطلاعاتی که در اختیار این پژوهش میگذارید چیست؟

از شما و یا همسرتان خواسته میشود یک فرم مربوط به اطلاعات تماس، میزان تحصیلات تان و شغل و پیشینه زبانی خود و همسر خود پر کنید. این اطلاعات به هیچ نحو و در هیچ کجا در اختیار هیچ کس قرار گرفته نخواهد شد.

بر سر اطلاعاتی که از کودک من جمع میشود چه می آید؟

فایل های صوتی و تصویری که از کودک شما جمع آوری میشود در یکی از کامپیوترهای دانشگاه یورک محفوظ خواهد بود. یک نسخه پشتیبانی هم از آن تهیه و نگهداری خواهد شد. نتایج در رساله دکتری من، ارائه گزارش های شفاهی و یک یا چند مقاله علمی ارائه خواهد شد. فایل های صوتی و تصویری ممکن است بعد از اتمام دوره دکترای من نگهداری و برای پژوهش های آتی مورد استفاده قرار گیرد. صورت شما و کودک شما در فایل های تصویری واضح خواهد بود. با این حال شما قادر به تصمیم گیری در مورد استفاده و یا عدم استفاده از فایل های تصویری در ارائه گزارشات خواهید بود.

چه بر سر هویت من و کودک من میاد؟

هویت شما و کودک شما به جد محفوظ خواهد ماند. اسم واقعی کودک شما به هیچ وجه و در هیچ نوع مقاله و یا ارائه گزارشی استفاده نخواهد شد. اطلاعات شما در زمینه هویت شما و کودکتان در دپارتمان زبان و علوم زبانشناختی دانشگاه یورک انگلستان نگهداری خواهد شد.

آیا از نتایج با خبر خواهم شد؟

ممکن است من نتوانم نتایج فردی کودک شما را در اختیار شما بگذارم. اما پس از اتمام پژوهش، در صورت تمایل شما و ارائه آدرس ایمیل تان میتوانم خلاصه نتایج را در اختیار شما قرار دهم.

Appendix 7 – Consent Form (Persian)

UNIVERSITY of York

Sarvenaz Moradi
DEPARTMENT OF LANGUAGE AND LINGUISTIC SCIENCE
Heslington, York, YO10 5DD, UK
Mobile : (+44)XXXXXX(UK)
(+98)XXXXXX(Iran)
Email : sm1833@york.ac.uk

نام پژوهش: رویکرد تمپلاتیک به واج شناسی کودک فارسی زبان

نام پژوهشگر: سروناز مرادی

رضایت نامه کودک

این فرم برای ثبت رضایتمندی شرکت شما و کودکتان در این پژوهش است. لطفا هر سوال را با دقت بخوانید و به آن پاسخ دهید. اگر متوجه قسمتی نمی شوید و یا نیاز به اطلاعات بیشتری دارید، از پژوهشگر سوال کنید.

۱. آیا برگه اطلاعات را کامل خوانده و مفهوم آن را کامل متوجه شده اید؟	بله	خیر
۲. آیا فرصت پرسش سوال در مورد پژوهش را داشتید و آیا این پاسخ ها به طور رضایت بخشی پاسخ داده شدند؟	بله	خیر
۳.آیا کامل متوجه این امر شدید که اطلاعات شما نزد پژوهشگر محفوظ خواهد ماند و اسم و تصویر شما و کودک شما مخفی خواهد ماند؟	بله	خیر
۴. آیا کامل متوجه این امر شدید که در هر مرحله ایی میتواند بدون نیاز به ارائه دلیلی از شرکت در پژوهش انصراف دهید و در این صورت اطلاعات شما کاملا نابود میشود؟	بله	خیر
۵. آیا کامل متوجه این امر شدید که اطلاعاتی که شما و کودکتان ارائه میکنید ممکن است در پژوهش های آتی مورد استفاده قرار گیرد؟	بله	خیر
۶. آیا موافق شرکت در این پژوهش هستید؟	بله	خیر
۷. آیا رضایت میدهید که شما و کودکتان در هنگام بازی در ضبط ویدیویی و صوتی شرکت کنید؟ اگر خود تمایل به شرکت ندارید آیا فرد دیگری از اعضای خانوادتان شرکت خواهد کرد؟	بله	خیر
۸. آیا رضایت میدهید در صورت نیاز گزیده ایی از تصاویر و صوت شما و کودک شما در ارائه گزارشات، بدون فاش شدن هویت شما، مورد استفاده قرار گیرد؟ (در صورت عدم رضایت در این مورد، همچنان میتوانید در پژوهش شرکت کنید)	بله	خیر
۹. آیا رضایت میدهید پژوهشگر اطلاعات تماس شما را بعد از اتمام پژوهش نگهداری کند تا در صورت نیاز برای شرکت در پژوهش های احتمالی آتی از شما دعوت کند؟ (در صورت عدم رضایت در این مورد، همچنان میتوانید در پژوهش شرکت کنید)	بله	خیر

نام شما:
نام کودک شما:
آدرس ایمیل شما (ایمیل خود را در صورتی که تمایل به دریافت خلاصه نتایج دارید بنویسید):
امضا شما:

نام پژوهشگر:
امضای پژوهشگر:
تاریخ:

Appendix 8 – Linguistics background form (Persian)

Sarvenaz Moradi
DEPARTMENT OF LANGUAGE AND LINGUISTIC SCIENCE
Heslington, York, YO10 5DD, UK
Mobile : (+44)XXXXXX(UK)
(+98)XXXXXX(Iran)
Email : sm1833@york.ac.uk

تاریخ: .....

درباره کودک

نام کودک شما: .....
جنسیت کودک شما: .....
تاریخ تولد کودک شما: .....
زبان‌هایی که کودک شما در معرض آن قرار می‌گیرد: .....

درباره مادر

تحصیلات مادر: .....
شغل مادر: .....
وضعیت شغلی مادر:
تمام وقت □
نیمه وقت □

زبان مادری مادر: .....
زبان‌های دیگر مادر: .....

درباره پدر

تحصیلات پدر: .....
شغل پدر: .....
وضعیت شغلی پدر:
تمام وقت □
نیمه وقت □

زبان مادری پدر: .....
زبان‌های دیگر پدر: .....