INVESTIGATING FINNISH-SPEAKING CHILDREN’S NOUN MORPHOLOGY: HOW DO YOUNG CHILDREN ACQUIRE CASE MARKING?

Thesis submitted to the University of Manchester for the degree of Doctor in Philosophy in the Faculty of Medical and Human Sciences

2015

HENNA PAULIINA LEMETYINEN

SCHOOL OF PSYCHOLOGICAL SCIENCES
Table of Contents

LIST OF TABLES .................................................................................................................. 6
LIST OF FIGURES ................................................................................................................ 7
ABSTRACT ............................................................................................................................... 7
DECLARATION ......................................................................................................................... 8
COPYRIGHT STATEMENT ....................................................................................................... 9
ACKNOWLEDGEMENTS ......................................................................................................... 10

Chapter 1: General introduction to language acquisition research ........................................ 11
  1.1. Generativist approaches to child language ................................................................. 11
  1.2. Usage-based approaches to child language .............................................................. 14
  1.3. The acquisition of morphology ................................................................................ 20
  2. The English past tense debate .................................................................................... 25
  3. The Competition Model .............................................................................................. 29
  4. Finnish .......................................................................................................................... 34
  5. Overview of the thesis ............................................................................................... 35
References ............................................................................................................................. 38

Rationale for Study 1 ......................................................................................................... 48

Chapter 2: Investigating a Finnish child’s noun morphology acquisition: Productivity ....... 49

Abstract ............................................................................................................................ 49
  1. Introduction .................................................................................................................... 50
  1.1. Generativist approaches to productivity ................................................................ 50
  1.2. Usage-based approaches ....................................................................................... 52
  1.3. Issues with defining and measuring productivity .................................................... 55
  1.4. Most recent measures of productivity ................................................................... 58
  1.5. Earlier naturalistic studies on child Finnish ............................................................ 61
  1.6. The aim of the study .............................................................................................. 65
  2. Method ......................................................................................................................... 67
    2.1. Participants ............................................................................................................. 67
    2.2. Procedure .............................................................................................................. 67
    2.3. Transcription and coding ..................................................................................... 68
  3. Results .......................................................................................................................... 70
  4. Discussion .................................................................................................................... 73
References ............................................................................................................................. 78

Rationale for Study 2 ......................................................................................................... 83
Chapter 3: Naturalistic study 2: Case marking errors .............................................. 84
Abstract .................................................................................................................. 84
1. Introduction ......................................................................................................... 85
   1.1. The aim and the research questions of the study ........................................ 95
2. Methods ............................................................................................................ 99
   2.1. Participants and data coding .................................................................... 99
   2.2. Procedure ............................................................................................... 100
3. Results ............................................................................................................. 101
4. Discussion ....................................................................................................... 107
References ........................................................................................................... 116
Rationale for Study 3 ............................................................................................. 121
Chapter 4: How do Finnish children comprehend word order and case marking cues in
transitive sentences with novel verbs? ................................................................. 122
Abstract .............................................................................................................. 122
1. Introduction ....................................................................................................... 123
2. Study 1: Word order and case marking in Finnish child-directed speech .......... 133
   2.1. Method .................................................................................................... 134
   2.2. Results and discussion ............................................................................ 136
3. Study 2: Comprehension experiment ............................................................... 140
   3.1. Method .................................................................................................... 140
      3.1.1. Participants ...................................................................................... 140
      3.1.2. Materials ........................................................................................ 140
      3.1.3. Design ............................................................................................ 142
      3.1.4. Procedure ....................................................................................... 144
      3.1.4.1. Part one ...................................................................................... 144
      3.1.4.2. Part two ...................................................................................... 147
   4. Results .......................................................................................................... 148
   5. Discussion ..................................................................................................... 154
   6. Conclusion .................................................................................................... 164
Chapter 5: General discussion ............................................................................. 170
1. Theoretical issues ............................................................................................ 173
2. Methodological issues ..................................................................................... 178
3. Implications for future research .................................................................... 181
4. Conclusion ....................................................................................................... 187
LIST OF TABLES

Page 62 - Table 2.1. The Finnish noun case marking system.

Page 70 - Table 2.2. Descriptive statistics for the controlled child (Mailis) and mother samples.

Page 100 - Table 3.1. Error rates for each case and rates of different cases used in incorrect contexts (%).

Page 104 - Table 3.2. Input type and token frequencies correlated with the child’s number of incorrect case uses (measures 4 and 5).

Page 105 - Table 3.3. Absolute input frequencies of noun+case combinations correlated with the child’s accuracy with these combinations (measure 6). Relative (%) input frequencies of noun+case combinations correlated with the child’s accuracy with these combinations (measure 7).

Page 136 - Table 4.1. Descriptive statistics for each sentence type (i.e. condition) per age. Mean percentages of ‘first noun as agent’ responses in each sentence type per age group with standard deviations.
LIST OF FIGURES

Page 70 - Figure 2.1. Type frequencies distributed across numbers of inflections for the child and mother samples (averaged across five samples).

Page 102 - Figure 3.1. Absolute noun input frequencies correlated with the child’s accuracy per noun type (measure 1).

Page 103 - Figure 3.2. Absolute case input token frequencies correlated with the child’s accuracy per case (measure 2).

Page 103 - Figure 3.3. Absolute input type frequencies correlated with the child’s accuracy per case(measure 3).

Page 136 - Figure 4.1. Cue availability, reliability and validity values (%) for agent noun phrase (NP) frames in Finnish child-directed transitive sentences.

Page 137 - Figure 4.2. Cue availability, reliability and validity values (%) for patient noun phrase (NP) frames in Finnish child-directed transitive sentences.

Page 151 - Figure 4.3. Mean proportion of ‘first noun as agent’ interpretations of each age group in every condition. Performance above chance with a 2-tailed t-test, adjusted $p = .002^*$. 

Page 152 - Figure 4.4. Mean percentages of ‘first noun as agent’ interpretations for each sentence type (condition) per age group: Interaction.
Child language acquisition is a fundamental topic in cognitive sciences as understanding language development can aid our understanding of cognition in general. The aim of this thesis is to explore inflectional noun morphology acquisition of Finnish-speaking children. There are two major methodologies applied to achieve the aim. A corpus consisting of 18 hours of child and child-direct speech was collected for a productivity analysis and for analyses of different input frequency effects on the child’s accuracy with case marking and the likelihood of extending cases to incorrect contexts. Second, an experiment was designed to investigate how two-, four-, and six-year-old Finnish children comprehend transitive sentences.

Study 1 (Chapter 2) investigated what linguistic productivity means. The study compared the mean number of inflections per noun between controlled samples of child speech at the ages of 1;7-1;8, and adult speech to compare the child’s productivity with noun inflections against an adult speaker. The results showed a significant difference in productivity between the two samples, demonstrating the gradual emergence of linguistic productivity.

Study 2 (Chapter 3) utilised the same naturalistic corpus as Study 1 to examine the impact of seven input frequency measures on the child’s accuracy with noun case marking. The analysis demonstrates that while high type and token input frequencies did correlate with more accurate use of nouns compared to items with lower frequencies, high frequencies were also associated with increased likelihood of usage in incorrect contexts.

Study 3 (Chapter 4) presents an experiment comparing children’s comprehension of grammatical cues in transitive sentences. Case marking and word order were manipulated to establish how children of different ages utilise these cues in comprehension. The results demonstrated that all age groups performed most accurately when both cues supported each other whereas all age groups struggled with conflicting cues.
DECLARATION

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

COPYRIGHT STATEMENT

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trademarks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see http://www.manchester.ac.uk/library/aboutus/regulations) and in The University’s policy on Presentation of Theses.
ACKNOWLEDGEMENTS

I want to express my gratitude to my supervisors, Elena Lieven and Anna Theakston, for excellent guidance, never-ending patience and understanding my struggles during these challenging and intellectually stimulating four years. While I have learned a lot about child language, developmental psychology and research, nothing can replace the lessons I have learned in perseverance and hard work.

Special thanks must go to the Knitting Club, which was/is something I always look forward to. Anna, Helen, Rachel and Stacey – I cannot thank you enough for the support whether it was conveyed through humour, discussion or wine.

How could I forget my craziest, most lovable friend and mentor, Sanna. Most things I learned about studying, I learned from you.

Last but most definitely not least, I need to thank you, Wai, for everything you have done for me for the past four years. You always make me feel like my wellbeing is your number one priority. It is my turn to support you now.
Chapter 1: General introduction to language acquisition research

Language acquisition is a crucial part of a child’s development - not only because language is vital for communicating with fellow humans but also because it is a reliable predictor for development of numerous social and cognitive domains. A child’s language ability is linked to comprehension of false belief (de Villiers & Pyers, 2002; Milligan, Astington & Dack, 2007; Slade & Ruffman, 2005), school readiness, education outcomes and literacy (Fiorentine & Howe, 2004; Forget-Dubois et al., 2009; Levin, Ravid & Rapaport, 2001; Lyytinen, Eklund & Lyytinen, 2005) and it can even predict behaviour problems (Lonigan et al., 1999; Stowe, Arnold & Ortiz, 1999). Language can also reveal glimpses of cognition in general. For example, language is argued to structure or restructure spatial cognition (Majid, Bowerman, Kita, Haun & Levinson, 2004), to affect representations of number (Frank, Everett, Fedorenko & Gibson, 2008; Spelke & Tsivkin, 2001) and even shape how we perceive time (Boroditsky, 2001).

1.1. Generativist approaches to child language

One of the key debates in the language acquisition field is the question of what mechanisms lead to adult-like, abstract grammar. Generativist approaches are tackling this problem by assuming Universal Grammar (UG), which is a set of innate, language-modular principles that govern the acquisition process (Hoekstra & Hyams, 1998; Marcus, 1998; Ullman, 2001; Wexler, 1998). In fact, one of the core assumptions of the UG approach is that experience alone is not sufficient to acquire the abstract syntax that would enable children to comprehend and produce
sentences that they might have never heard before. Instead of relying purely on input and non-specific cognitive processes, infants have underlying competence to manipulate abstract symbols, i.e. grammar, and the language learner’s task is to link these symbols and variables to the particular language being acquired. Children are assumed to generalise a symbolic rule to all members of a grammatical category (e.g. the regular past tense marker –ed to all verbs) once they’ve learnt to apply the rule to one member, like a verb. Children are also assumed to have the same grammatical categories from birth as adults (Hoekstra & Hyams, 1998; Valian, Solt & Stewart, 2009; Wexler, 1998). Of course, many details about the specific language the child is hearing have to be learned (including phonology, lexicon and the specifics of the grammar) but the idea is that these can be ‘fitted’ into the pre-existing universal grammar. In addition, as all parts of language are not rule-based, some have to be rote-learnt, such as idioms and irregular morphology (Pinker & Ullman, 2002; Ullman, 2001).

Universal Grammar also contains general principles (‘rules’) that guide the innate linguistic categories with how sentences ought to be structured. Principles are also shared by all languages. For example, a transitive sentence in any language contains a subject, a verb and an object, even if one or some of the arguments can be omitted. These principles are adjusted by specific, detail-oriented parameters that ensure the principles fit the particular language a child is acquiring. The role of linguistic input is to turn appropriate parameters on in order for the child to acquire an appropriate grammar for the target language. For instance, the word order of an English transitive construction differs from Japanese and Turkish word orders,
which is accounted for by parameter setting. An underlying assumption following the idea of language processing being based on innate principles and parameters is that language acquisition should occur with relative ease, fairly few errors and sudden productivity. This would be a result of a parameter being either switched on or off – no medium between the two positions has been included in UG (Wexler, 1998). To tackle the problem of children mastering different aspects of grammar in different developmental stages, it has been proposed that not all parameters are available from birth. Some parameters, that children seem to acquire in later stages of acquisition, become available when children have passed the earliest stages of development (Radford, 1990). It is still important to note that even when a gradual availability of parameters is assumed, it does not exclude the assumption of productivity from the earliest stages of a parameter’s availability and consequently, low error rates. These assumptions contradict a body of research that suggests a lexically-specific system, that is partly productive at best at the early stages of acquisition. Yang (2002) and Legate & Yang (2007) addressed the problem with instantaneous acquisition by proposing that these parameters are set probabilistically based on the kinds of input a child is exposed to. When they conducted a cross-linguistic comparison of the proportion of verbs with overt tense marking in English, Spanish and French children's input, the theory predicted the order of verbal tense marking acquisition of these languages correctly based on input. While it is important that the role of input is acknowledged in this model, the issue is that innate, language-modular parameters are still assumed to be necessary for language acquisition in this probabilistic parameter setting account.
1.2. **Usage-based approaches to child language**

While generativist accounts approach language acquisition in terms of innate language universals, usage-based accounts aim to explain the acquisition process in terms of cognitive-general mechanisms that are not necessarily specific to language. Usage-based approaches account for language acquisition with mechanisms like statistical learning and pattern detection. There is considerable evidence showing that other aspects of cognition than language, including visual perception, auditory perception and problem solving, benefit from these mechanisms (Conway & Christiansen, 2005; Cosmides & Tooby, 1996; Fiser & Aslin, 2001; Kirkham, Slemmer & Johnson, 2002). Thus, an underlying assumption of usage-based approaches is that the mechanisms that are applied in language acquisition and learning in general are human-evolved cognitive processes but just are not specific to language. Consequently, input plays a critical role in these approaches compared to generativist approaches that argue that language input is in fact too impoverished for a child to ever achieve abstract, adult-like command of grammar. Relative input frequencies of specific lexical items and constructions can in fact account for the kinds of output patterns observed in naturalistic and experimental data (Abbot-Smith & Tomasello, 2006; Rowland, Pine, Lieven & Theakston, 2003; Theakston, Lieven, Pine & Rowland, 2001; Theakston, Lieven, Pine & Rowland, 2004; Theakston, Lieven & Tomasello, 2003; Tomasello, 2000a; Tomasello, 2000b). It is firmly established that error rates of an individual item (e.g. an inflected form) decrease and the relative frequency of the correct use in obligatory contexts increases in a child’s speech as a function of high input
frequency. For example, children’s errors with auxiliaries in WH-questions can be explained in terms of input frequencies of auxiliaries in these types of questions. A naturalistic investigation of two- and three-year-old children’s spontaneous speech revealed significantly different error rates for the auxiliaries DO, BE and HAVE but also for their different forms (e.g. IS and ARE for BE; HAVE and HAS for HAVE). Also uses of these verbs as copulas were analysed. The results showed that wh-questions with copula and auxiliary IS and auxiliary HAS were used correctly more often than copula and auxiliary ARE and HAVE. These rates of correct use and errors reflect the input frequency of the frames each verb occurs in. Initially, children acquire partly productive frames like ‘what’s X doing?’, which increases the rate of correct use of the auxiliary of these kinds of frames, like IS in the aforementioned example. Relatively high error rates with some of the auxiliaries and copulas occur when children don’t have appropriate wh-question frames to use them in yet (Rowland, Pine, Lieven & Theakston, 2005).

The frequency effects literature distinguishes two kinds of frequencies: token frequency and type frequency. Token frequency refers to the number of times an item is attested in input (e.g. mother), while type frequency refers to the number of different items attested in a particular pattern, such as the number of different nouns used with the regular plural ending –s in input. It is suggested that high type frequency is particularly crucial to productivity by opening constructions up for a wider variety of new types. The more types a construction is used with, the easier making analogies between lexically specific constructions becomes (Bybee, 1995; Tomasello, 2000b). When applied to inflectional morphology acquisition,
children should be more productive with suffixes with high type frequency compared to those with lower type frequencies and/or a high token frequency instead. Modelling in particular has provided support for this but due to the close connection between inflectional morphology and phonology, many studies struggle to separate the effects of the two (Bybee, 1995; Forrester & Plunkett, 1994; Plunkett & Nakisa, 1997), making it harder to demonstrate precise type frequency effects. While it is known that type and token frequencies are linked to productivity, abstraction and error patterns, there are gaps in the literature concerning how both kinds of frequencies interact to impact on those aspects of child language. For example, at present it is unclear how accurately children comprehend and use high token frequency lemmas in combination with low type frequency affixes, or a low type frequency lemma with a high token frequency affix. Both kinds of frequencies support acquisition in different ways, but it is also known that high token frequencies can cause error, too, by substitution in place of lower frequency forms (Aguado-Orea & Pine, 2015). Another underexplored interaction is that of between absolute and relative frequencies. Absolute frequency is the actual number of items in a sample, and relative frequency is the competition between items, such as the competition between two inflected forms of the same lemma. High absolute frequency has been shown to promote the correct usage of irregular forms instead of an erroneous regular form (e.g. feet instead of *foots) (Maslen, Theakston, Lieven & Tomasello, 2004). But high relative frequency in relation to a potential competitor also increases children’s accuracy (Matthews & Theakston, 2006). Currently it is unclear how these frequency effects interact in different contexts. Children most likely perform best when an item of the highest relative frequency
also has a high absolute frequency, but it is currently challenging to predict performance in contexts with items of mixed frequencies. The field lacks an account that can make reliable predictions regarding children’s performance in such contexts. (For further discussion, please see Chapter 3.)

The importance of frequency effects on language acquisition has been captured by studies investigating what children do with unusual word orders. These studies typically involve teaching children verbs, novel and familiar, in a highly frequent word order (e.g. SVO for English) and in a weird word order to examine effects of input on syntax acquisition. A classic study of this nature was conducted by Akhtar (1999), who taught novel verbs to 2-, 3-, and 4-year-olds either in SVO, SOV or VSO. Two- and three-year-olds happily used the two ‘weird’ word orders (SOV and VSO) with novel verbs whereas the oldest group was resistant and corrected the experimenter when other orders than SVO were used. The effect was not observed in control conditions for any age group where a familiar verb was used instead. These results illustrate the interaction between a construction and specific lexical items. This methodology was adapted for a study examining children’s (2;4, 3;9) representations of the intransitive construction (Abbot-Smith, Lieven & Tomasello, 2001). They compared children’s performance in the more frequent SV order against the less frequent VS with novel verbs as well as VS with a familiar verb. Three-year-olds showed a clear preference to correct the low frequency word order to the high frequency SV with both familiar and novel verbs. Two-year-olds’ performance, on the other hand, depended on verb frequency. They used novel verbs in the word order they had been taught and used familiar verbs
more often than novel ones in SV. These studies by Akhtar (1999) and Abbot-Smith et al. (2001) show that 2-year-old children’s representations of constructions are more lexically specific than older children’s representations, which is an effect of item frequency. Overall, findings such as these illustrate how linguistic categories emerge gradually. A 2-year-old child’s schema of a construction, e.g. the transitive, may be weaker or more partial compared to an older child’s schema which is more entrenched due to experience. In fact, a recent study on young children’s flexibility and creativity with the transitive schema established a novel method for measuring the relative degrees of abstraction. This naturalistic study aimed to compare the levels of abstraction between child and adult samples to examine whether children show adult-like productivity within a construction. Two, and three-year-old children’s flexibility, measured as the absolute and relative type frequencies of verbs within noun slots (e.g.” __ want __”) was compared against flexibility observed in child-directed speech. Similarly, the children’s creativity, defined as two measures of lexical specificity and overlap of lexical items, were compared against adult creativity. Importantly, the study captured significant differences between children’s and adults’ degrees of abstraction, which supports the idea of graded representations (Theakston, Ibbotson, Freudenthal, Lieven & Tomasello, 2015). In contrast with measures of production, especially those taken in experimental contexts, it is important to bear in mind that comprehension studies adopting a preferential looking task or the forced-choice pointing method may be less challenging, which is why children show earlier ability with these methodologies. There are asymmetries between children’s comprehension and production, which can become evident in contexts with heavy cognitive load (Brandt-Kobele & Höhle,
Methodologies that have comprehension and production components, such as act-out or sentence repetition, are more demanding than preferential looking or forced-choice pointing due to additional working memory demands – especially if participants are not trained with the test items. However, this does not provide evidence for children possessing abstract representations of syntax and just being unable to perform under increased cognitive load. Abbot-Smith, Lieven & Tomasello (2008) proposed that as young as two-year-old children have graded representations of constructions, such as the transitive. The strength of these representations increases as a function of age and linguistic experience via input. As children struggle to actively access representations that are still fairly weak (e.g. at the age of two), different experimental methodologies yield different results for children’s performance with the same construction. Children as young as 1;11, with a mean of 2;3 (Noble, Rowland & Pine, 2011), have been found to identify a matching scene out of two simultaneous options in a forced choice pointing paradigm task with novel verbs in the transitive construction. Dittmar, Abbot-Smith, Lieven & Tomasello (2008a) demonstrated 21-month-old German children to have abstract but weak knowledge of transitives as long as they were trained with the same verbs before testing.

In addition to frequency, another powerful general learning mechanism that also applies to language learning is analogical processing. An essential cognitive tool for becoming a productive language user is making analogies between patterns in language. In fact, usage-based approaches assume that children acquire all aspects of language by making analogies between stored patterns and novel items, whereas
generativist approaches utilise analogy as only one part of language development. Once a child has acquired a number of lexically-specific constructions, these constructions or schemas are generalised based on functional or formal similarity of two schemas. For example, a child may rote learn a phrase like ‘Where’s Mummy?’ as one item, but after hearing multiple instances of other utterances beginning with ‘Where’s…[object]’ this will develop into a partly productive ‘Where’s X’ frame, and eventually into an abstract WH-question schema. It is unclear what the critical step is that must occur between the acquisition of multiple lexically-specific schemas and abstraction between them (Tomasello, 2000b).

Many world languages have a much richer inflectional system than English, which allows more flexibility with word orders. The child language field must study underrepresented languages in order to have a full view of the acquisition process for a universal account that can explain the acquisition patterns in any language. Studying languages with rich morphology and alternative word orders might reveal, for instance, more complex frequency effects or different interactions between meaning and form compared to languages, e.g. English and French, that have been studied so far.

1.3. The acquisition of morphology

Early studies on inflectional morphology acquisition claim that even the earliest stages of children’s morphology use is almost completely error-free. Cross-linguistic data reviewed by Hoekstra & Hyams (1998) revealed an overall error rate of smaller than 5% with person/number agreement marking, leading to their conclusion that children possess innate knowledge of verb inflection. These data suffer from
sampling issues as the data were not dense enough to capture low frequency forms and very early usage of inflections. The data also lack controls to truly measure the children’s productivity against, such as a baseline for inflection usage and controlling vocabulary and sample size. Finally, when children’s error rates are collapsed altogether across all lexical items and inflections, as they have been in these studies, the *overall* error rate does remain low but misses out on parts of the system children make more errors with (Maslen, Theakston, Lieven & Tomasello, 2004; Maratsos, 2000; Rubino & Pine, 1998). Lack of these control measures has misled researchers to conclude children to have adult-like productivity from the earliest stages of inflectional morphology acquisition.

More recent studies on children’s productivity with morphology demonstrate the methodological weaknesses in older research. Two naturalistic studies on child productivity demonstrated the importance of the control measures that the earlier studies lack. Krajewski, Lieven & Theakston (2012), examining a Polish child’s productivity with noun morphology, and Aguado-Orea & Pine (2015), investigating two Spanish children’s verb inflection, compared the children’s speech to their adult counterparts to establish a baseline for productivity. The baseline is essential for detecting the level of lexical specificity in the child sample. Aguado-Orea & Pine (2015) point out that even the adult speakers in their Spanish corpus are somewhat restricted with their verb + inflection combinations, which shows that ‘adult-like productivity’ is not evenly distributed use of lexical items and inflections. These two studies controlled for vocabulary, inflections and sample size to ensure the differences between the adult samples and the child samples cannot
be explained by differences in lexical knowledge or acquired suffixes, or skewed data due to the adults speaking more than the child participants. The results from both studies support the usage-based claim of gradually emerging productivity, as the children were significantly less flexible in their lemma + suffix combinations. (For more details, see Chapter 2.)

Morphology is an important subject for language acquisition research as understanding the developmental trajectory and the means of morphology acquisition can provide insight into the language acquisition process in general. Investigating morphology acquisition provides a useful testbed for understanding how children achieve full linguistic productivity. Generally speaking, the ability to create novel, grammatically correct utterances is considered a prime measure of productivity (Mirkovic, Seidenberg & Joanisse, 2011; Pinker, 1999). Morphology, especially case marking, is guided by syntax (Dressler, 2012). Firstly, it is known that rich morphological systems promote more rapid acquisition of inflections in terms of how early a child’s earliest forms emerge compared to languages that have less morphology. Even though theoretical morphological richness - the maximal use of all possible inflections of a morphological system - is not always attested in real, naturalistic speech samples, studies of child-directed speech (CDS) have shown that CDS samples of morphologically richer languages do show more evidence for more complex morphology compared to languages with less morphological richness (Xanthos et al., 2011). It has also been found that an aspect affecting the acquisition of each inflectional marker is its formal complexity. It comprises three subcomponents: degree of reliability, transparency and iconic affixation, described
by Clark (2003). Degree of reliability measures the extent of uniqueness of single morphemes. For example, the English –s is less unique, and thus less reliable, than the –st morpheme. The morpheme –s denotes both plurality in nouns and verbal third singular forms, while the –st morpheme denotes only superlative adjective forms. Transparency refers to the clarity of affixes and any stem changes a lexical item is affected by in the process of inflection. Lastly, affixation by adding a morpheme to the end or beginning of a word is acquired faster than affixation that is less common, like adding an affix in the middle of the a word.

Mirkovic, Seidenberg & Joanisse (2011) provided convincing evidence for the idea that general cognitive principles govern acquisition of inflectional morphology. They trained five models of a connectionist network on a large corpus consisting of more than 3,000 Serbian nouns that are inflected for case, gender and number. The aim of the study was to examine how well the network will generalise its knowledge of morphology to produce noun and suffix combinations it had not encountered before, even though the network had encountered the particular inflections and nouns separately – just like a child would have. The five models, i.e. ‘participants’, had different initial weights in training and the final results are an average of the results from each simulation. Following the training, the authors analysed inflection errors and correctly produced forms. The network’s behaviours resembled results gained from earlier studies with human participants. The errors analyses demonstrated that low frequency forms and forms with multiple possible suffixes were most challenging to acquire, which is predicted by constructivist and usage-based approaches due to an increased number of competing items. Even
though the models made errors with noun suffixes, it is important to note that the overall error percentage was low. Analyses of correct morphological forms revealed that a relatively simple network could acquire a complex morphological system, showing also generalisation of these newly learnt inflections. Overall, the analyses showed that this kind of a model learnt complex morphophonological noun forms in a way that emulated human learning. A key finding was the importance of consistency between the input and output relationships, which was found to facilitate production of correct inflections. The study also showed that the cognitive-general principles shown to underlie language acquisition in English-speaking children are equally applicable when a morphologically more complex language, like Serbian, is learnt. It is important to bear in mind that current computational models do not include the whole complexity of language acquisition. For example, many lack the impact of semantics and phonology, but nevertheless, they do offer a fairly clear picture of some of the cognitive principles underlying the acquisition process. Another widely investigated domain concerning morphological productivity, and what it can tell us about the language acquisition process more generally, comes from the English past tense debate. Even though English is a morphologically impoverished language, the interest arises because the acquisition of the regular \(-ed\) past tense suffix provides an ideal testbed for whether language is operated via abstract variable manipulation, such as ‘add \(-ed\) to the end of all regular verbs’ or by analogy formation across stored past tense exemplars.
2. The English past tense debate

One of the fiercest debates in the study of inflection is the acquisition of regular and irregular forms – particularly those of the English past tense. Generativist researchers tend to account for the phenomenon with two mechanisms: a formal rule for regular forms (“add –ed to the end of the verb”) and storing and retrieving whole words for irregular forms. Irregular forms are stored based on phonological neighbourhood: items with similar phonological properties are grouped together in the lexicon whereas phonology is argued not to have an effect on regular past tense form storage or retrieval. For example, Prasada & Pinker (1993) argue for this kind of a hybrid model. They conducted a series of experiments to study the mechanisms of acquiring verb inflections. Adult speakers of English completed tasks that involved inflecting sixty novel verbs in grammatically correct sentence frames. One experiment utilised also novel nouns and adjectives with definitions to encourage the participants to treat them like real words instead of sounds that resemble English. It was found that when a novel verb was phonologically similar to a familiar irregular verb, the participants were more likely to use an irregular verb inflection paradigm than use the regular –ed ending. However, an analogy formation mechanism was not associated with novel verbs that resembled familiar regular verbs. Following the results, the authors argued that the English past tense is formed by two different mechanisms: regular inflections are formed with an ‘add the suffix –ed to the verb stem’ rule and irregular forms are retrieved from the associative memory. Ullman (2001) even argues that the two routes for past tense formation are orchestrated by two different brain structures: association formation,
i.e. irregular past tense forms, are claimed to be based in the temporal lobe whereas rule-based systems, such as the regular past tense forms, are claimed to originate from the frontal lobe and basal ganglia regions.

An issue with the dual route approach is that it seems tailored to explain the acquisition of the English past tense, which is a system of clear-cut regular and irregular verb categories. This is not the case with some other languages, especially those with more complex morphology. Finnish verbs, for instance, are not divided into regular and irregular forms, but into six different conjugation categories based on phonological properties of the verb (Karlsson, 2008). An investigation into the past tense acquisition of Finnish children aged four to six years does suggest that superficial rule-like properties of inflectional morphology acquisition arise due to children’s sensitivity to lemma/base frequency and phonological similarity across lexical items rather than an application of an actual ‘add a past tense suffix’ type of rule (Kirjavainen, Nikolaev & Kidd, 2012). Another issue with the dual route model is that it assumes mechanisms that are language-modular, even though there are alternatives to accounting for language processing in cognitively general terms, such as analogy and frequency. Usage-based approaches argue for a system that is psychologically and biologically more economical compared to the dual-route model. Instead of having two separate pathways, that may be controlled by different brain regions, usage-based approaches propose that both regular and irregular past tense forms are generalised based on interaction between phonology and frequency of the novel and familiar item. If a novel verb resembles a familiar verb that is already stored in the speaker’s system, the novel verb will be inflected
according to the familiar verb’s inflection paradigm. For example, if an adult speaker of English is asked to provide a past tense form for the novel verb *clow*, they could inflect it according to a phonologically similar regular verb ‘*show+ed*’, or inflect it according to a phonologically similar irregular verb, ‘*know/knew*’. The preferred inflection paradigm depends on token frequency of the familiar verb and how easily each applicable exemplar is retrieved from memory (Ambridge, 2010).

There is further evidence supporting the effect of phonology on the acquisition of both regular and irregular past tense forms. An extensive investigation into the past tense acquisition of school-aged children between the ages of 3;8 and 13;5 revealed the importance of phonological properties of verb stems and past tense neighbourhood, i.e. the phonological resemblance, or rhyme, between verbs when they are inflected in past tense. The participants were prompted to produce a past tense form of fifty verbs with a picture-book elicitation task. The task entailed looking at images of intransitive actions and prompting the children to complete sentences of ‘what happened yesterday’, such as *This boy is walking. He walks everyday. Yesterday, he ….*. Null-marking and vowel changes were the most common error types across all verbs and age groups, but as expected, the error rates decreased as a function of age. Both regular and irregular past tense forms were more prone to errors when they ended with particular consonants (s, t, n, and l), had a low token frequency or had a low count of phonologically similar verbs, that could be utilised in analogy formation. These effects were evident with both regular and irregular past tense forms, supporting the single-route model of past tense acquisition (Marchman, 1997).
Other factors in addition to item-based learning, frequency and analogies have been identified to affect language development, and therefore the acquisition of past tense forms. Ramscar (2002) showed how semantic similarity plays a significant role in how past tense morphology is generalised to novel items. Adult English-speakers were presented novel verbs, e.g. *frink*, in a reading task. The novel verbs had phonologically similar regular and irregular verbs, such as *blink* and *drink*, respectively. It was argued that if semantics affect the inflection paradigm participants choose for the novel verbs, they are more likely to use the paradigm of that phonologically corresponding verb that is perceived to be semantically closest with the novel verb. When the semantic context of the novel verbs was manipulated to resemble a regular verb, more than 70% of the participants inflected the novel items as regular verbs. Equivalent findings were obtained when the novel verbs were presented semantically closer to an irregular verb – vast majority of participants inflected the novel items as irregular verbs. This finding is problematic for the dual-route account as it does not predict an effect of phonology with regular past tense forms. It also cannot explain the effect of semantics. The single-route account does predict an effect of phonology with both regular and irregular forms, which these findings are in accordance with, but the approach often does not *explicitly* include semantics in predictions despite having underlying assumptions regarding the impact of semantics in terms of form-function mappings. The study shows that a model of past tense acquisition must include the role of semantics, which both approaches lack, and the role of phonology in the acquisition of all past tense forms, which the dual-route model lacks.
3. The Competition Model

The competition model is a method to examine cross-linguistic differences in how grammatical cues such as word order or case marking are processed to obtain a meaning in sentence comprehension. It is also used to account for patterns of acquisition, such as why English-speaking children show a strong agent-initial bias (MacWhinney, Bates & Kliegl, 1984; Slobin & Bever, 1982) and that adult Chinese speakers rely more on the Chinese passive marker and animacy than word order (Li, Bates & MacWhinney, 1993). The model’s purpose is to map between two levels of information: the functional level and the formal level. The former conveys meanings and intentions of an utterance, e.g. the agent, the patient and what is happening. The latter contains technical information used in the particular language being processed, e.g. inflection and argument structure. The model describes the mapping of these levels as straightforward even though multiple forms can map onto more than one function. For instance, in most cases an English transitive sentence consists of at least one agent, an action, and at least one patient, in the respective order: A boy hugs a girl. In this example, the pre-verbal noun is an agent, allowing a “preverbal noun equates to agent” mapping. Even though the English word order does not allow a post-verbal agent, a preverbal noun does not always map onto the role of an agent. In a sentence like ‘a boy fears a girl’, the preverbal noun is an experiencer rather than an agent. This problem extends to morphological marking: the same affix can denote multiple meanings and multiple affixes can denote the same meaning. In English, the morpheme –s marks third person singular present tense verbs as well as contracted auxiliary and copula verbs.
and regular plural and possessive noun forms. This problem is even more widespread in languages with complex morphological systems like German and Polish that mark for number, case and gender in nouns at any one time, and not all meanings have a unique, single morpheme associated only with that meaning. The same meaning can be achieved with different morphemes, too. In English, regular and irregular plural forms are marked with different endings, as well as regular and irregular past tense forms. There are even five different plural markers in Swedish (Leonard, Salameh & Hansson, 2001). Complexity of these mappings makes investigating the acquisition of grammatical cues challenging. The more there is overlap between cues, e.g. one suffix mapping onto multiple different uses, the more complex the mappings are as additional cues are required to arrive at an accurate interpretation of a sentence.

A strength of the model is that it makes quantifiable predictions of how each grammatical cue is treated individually in comprehension. It is achieved by calculating cue validity, which consists of two subcomponents:

1) **Cue availability**: a measure of how frequently a cue is present in relevant utterances, e.g. how often the first of two arguments is marked with the accusative case as opposed to a different case, lack of case marking or ambiguous case marking.

2) **Cue reliability**: a measure of how frequently a cue maps onto a particular function, e.g. how often accusative case marking denotes agency in a transitive sentence.
To test the competition account, MacWhinney, Pléh & Bates (1985) conducted a series of experiments on Hungarian pre-schoolers. Hungarian is known as a language with rich noun case marking and flexible word order which made it a novel language to investigate from the perspective of cue competition. It is argued that cues with highest validity will be acquired first, which in Hungarian was predicted to be case marking over word order. Across the experiments, children aged three, four and five and adults were tested on transitive sentence comprehension with act-out tasks. Cues in the test sentences, including case marking, word order, animacy and stress, were manipulated to study the effect of these cues on which noun the participants interpreted as the agent of the actions. It is worth noting that the test sentences comprised familiar nouns and verbs, which made their interpretation easier than if novel items were used. The first experiment examined how the aforementioned cues affect the participants’ comprehension in comparison to one another. The results showed that the competition model was fairly accurate at predicting which cues were most valid in Hungarian transitive sentence interpretation. As case marking was a completely reliable and available cue, all age groups relied on it the most. Additionally, the youngest group considered animacy and word order as equally important cues but the latter declined consistently up to the mean age of 5;7.

The second experiment revealed that the participants’ comprehension may be affected by morphophonology. In this case marking was replaced with possessive marking to investigate the validity of the competition model further. Unlike in English, a possessive suffix can be added to nouns in Hungarian, for
example the meaning of *your dog* could be conveyed with a structure like *dog+2nd singular possessive*. The participants acted out sentences that were manipulated for possessive suffixes. The authors observed that the participants tended to interpret the second possessive form (*your*), as an accusative form as they are phonetically similar: the second singular possessive marker is a –*d* while the accusative marker is a –*t*. This observation inspired the third experiment, where the phonological detectability of suffixes was manipulated. The third study was otherwise almost identical to their second study. The only difference was that while in experiment two the second possessive marker resembled that of the accusative marker, in experiment three neither of the noun inflections used were similar to case markers. This forced the participants to rely on other cues than case marking to determine the agents and the patients. The results showed that children relied increasingly on animacy with increasing age. An issue with this series of experiments is that the test sentences comprised of familiar nouns and verbs. Test items that consist entirely of familiar items measures lexically-specific knowledge of the used items instead of *abstract* knowledge of the underlying grammar. Using familiar items instead of novel ones will boost children’s comprehensions purely by the elevated level of lexical knowledge as opposed to using e.g. novel verbs.

Newer studies utilising novel verbs have found that children learn to utilise grammatical cues in the transitive construction fairly late compared to the earlier studies outlined above. Even though one strength of the model is that it attempts to define how to quantify cue validity, i.e. cue weights, a major weakness of the model is also related to this. This problem is demonstrated in an investigation on
German children’s cue comprehension in transitive sentences (Dittmar, Abbot-Smith, Lieven & Tomasello, 2008). It was found that German 4-year-olds interpreted transitive sentences based on word order only. The four-year-olds’ performance was at chance with sentences where word order and case marking cues conflicted. The problem with the finding was that cue validity calculations for German child-directed speech did not show word order to be a more valid cue than case marking. The authors had two suggestions to account for the discrepancy between the prediction and the finding. First, the authors assumed that children use word order as a cue to spatial relationships between two noun phrases, agent and patient, in a transitive sentence. It may be that children use it as a cue to detect the position of a noun phrase in relation to a verb instead. Another way of explaining the contradiction between the result and the prediction is related to the item-based nature of language acquisition. German has three grammatical genders which interact with case marking, i.e. German cases have different endings depending on the grammatical gender of the noun. The usage-based approach has shown that young children do not generalise morphology as readily as some generativist approaches argue (Dąbrowska, 2004; Dąbrowska, 2006; Dąbrowska & Lieven, 2005; Krajewski, Lieven & Theakston, 2012, Räsänen, Ambridge & Pine, 2014). Thus four-year-old German speakers might treat various endings marking the same case differently from each other. When the authors re-calculated cue validities separately for individual markers instead of including all markers of one case together, validity for case marking was considerably lower than for word order (21% and 87%, respectively). These examples demonstrate how the model has been tested on multiple languages, especially in relation to comprehension of the simple
transitive construction. A more detailed review can be found in the introduction to Chapter 4.

4. Finnish

Finnish belongs to the Uralic language family and is related to Karelian and Estonian. It is an agglutinative language, meaning that multiple unchangeable affixes can be added to one stem instead of fusing affixes together. For example, a plural partitive form of ‘dog’ consists of the noun stem (*koir*), the plural suffix (*-i*) and the partitive suffix (*-a*); *koir+i+a*, meaning ‘some dogs’. Finnish is noted for the relatively high transparency in morphology as, with the exceptions discussed below, most case markers are unique to one case. Depending on the definition of a case, there are either fourteen or fifteen cases in Finnish. The controversy is caused by the phonological overlap between the traditional accusative case, the traditional genitive case and the nominative case. The singular accusative has two forms: an unmarked accusative, which is identical to the singular nominative form, and a marked form, which is identical to the singular genitive. In addition, the plural accusative form is identical to the nominative plural form. Because of this overlap, the accusative and the genitive are treated as one case in some sources and as two separate cases in others. Despite the similarity, the accusative and genitive do mark for different roles in sentences – direct object and possession, respectively, which is why we treat them as individual cases in this thesis. Nevertheless, Finnish still has a relatively high number of cases compared to the majority of world languages, which makes it an interesting test for language acquisition research. The final aspects to consider are general processing and memory effects specifically on case marking.
Recency effects have been documented to be greater than primacy effects in young children (Thurm & Glanzer, 1971). This creates a natural cognitive bias to acquiring case markers in a language where case suffixes are attached to the very end of the word, instead of e.g. the middle (Stephany & Voeikova, 2009). This directly applies to Finnish case marking, which is done by adding a suffix to the end of nominals: if there is more than one suffix attached to a word, young children might be more likely to store the last suffix in their long-term memory compared to other suffixes that are positioned before the last one, promoting the acquisition of the last suffix over the others. In Finnish, this effect is beneficial for case marking acquisition as case markers are the last attached morphemes.

There has not been much research on the typical acquisition of Finnish morphology with the exception of an early study by Bowerman (1973) and a few later naturalistic studies all of which are discussed in the introduction to Chapter 2.

5. **Overview of the thesis**

This thesis investigates how young children acquire a complex case marking system. More specifically, we first tackle productivity, followed by accuracy with case marking in the early stages of acquisition (i.e. case marking errors), and, finally, how and when children begin to appreciate case marking as a significant cue amongst others in sentence comprehension. The main research questions of the thesis are:

1. What is a young Finnish child’s productivity with case marking like and how does it compare to adult productivity?
2. What do possible relationships between different input frequency measures and a child’s accuracy with case marking reveal about the nature of morphology acquisition?

3. How do children acquire the understanding of how grammatical cues like case marking and word order interact to convey agent-patient relations?

The first study of this thesis (Chapter 2) addresses the question of whether children show early adult-like productivity with noun case marking. This is a theoretical issue that has been a topic of debate between the different approaches outlined above. To examine this, one Finnish mother-child dyad was audio-recorded for 18 hours in naturalistic everyday situations in their home. Strict controls for the child and adult samples were administered to allow careful comparisons of case marking productivity between the input and output.

In the second study (Chapter 3) I examined what kinds of input are correlated with a child’s accuracy with case marking. Previous literature has confirmed that frequency has a crucial role in predicting patterns of acquisition cross-linguistically. It is still unclear what types of frequency the child is sensitive to when acquiring case marking. We examined this by comparing seven frequency measures against a child’s accuracy with case marking to find out whether children attend to relative frequencies or competition between frequencies.

My third study (Chapter 4) investigates cue competition between two grammatical cues that affect Finnish speakers’ comprehension of agent-patient relationships: case marking and word order. Even though case marking theoretically
allows any word order in a transitive sentence with two arguments and a verb, some word orders are used more frequently than others in both child-directed and adult-directed speech. Therefore, word order frequencies might initially lead children acquiring Finnish to treat word order as a grammatical cue that can change the meaning of a sentence. Finnish children’s comprehension, aged two, four and six, was tested with causative transitive sentences in various word orders with unambiguous and ambiguous case marking.

The thesis ends with a general discussion of the theoretical and methodological implications of the results, together with suggestions for future research.
References


Rationale for Study 1

Productivity is the ability to generate novel utterances that the speaker has not necessarily heard before within the grammatical boundaries of the target language. As discussed in General Introduction, examining children’s morphological productivity is considered to reveal insights into the language acquisition, and cognitive processing more generally. The extent of young children’s morphological productivity interests researchers not only from the developmental perspective, but also for understanding the very nature of the language acquisition process considering the end-state of the entire typical child language acquisition process is adult-like full productivity.

Naturalistic research methods, such as conducting observations, keeping a diary of a child’s utterances and audio-recording a child’s speech, have been the core of language acquisition research since its beginning. Despite the popularity of naturalistic research, these kinds of studies have not evolved methodologically as swiftly as experimental studies. Data collection for naturalistic studies is time consuming and it can be fairly difficult to avoid participants opting out in longitudinal studies. In the past, naturalistic samples have not been sufficiently dense to allow reliable, controlled analyses of productivity to be conducted. Most importantly, child productivity did not used to be compared directly against a baseline sample, i.e. adult productivity. In this study we explore what productivity really means by examining a young Finnish child’s (1;7-1;8) noun inflection productivity and comparing her productivity, measured as the number of different inflections per noun, to her caregiver’s speech.
Chapter 2: Investigating a Finnish child’s noun morphology acquisition: Productivity

Abstract

Current accounts of child language acquisition debate the extent of young children’s grammatical productivity. Generativist approaches argue for relatively early adult-like productivity whereas constructivist and usage-based approaches assume that children attain adult-like competence much later, and only gradually. The current paper takes the acquisition of Finnish noun morphology as a test case and, using densely collected corpus data, reports a comparison of a child’s input and output to test what productivity really means and how it can be measured. Eighteen hours of audio recordings were collected from a monolingual Finnish mother-child dyad over the course of four weeks in naturalistic play situations at home when the child was aged 1;7-1;8. The recordings were transcribed and coded for nominal case, number and errors. The results revealed that the child was significantly less productive with case marking compared to the mother, even when strict vocabulary and sample size control were applied, providing evidence for the usage-based accounts. The findings support the usage-based prediction of a gradual acquisition process even though the child’s use of noun inflections seemingly looks almost adult-like. The study also showed that naturalistic data are very informative of productivity when appropriate controls are applied.
1. Introduction

One question the language acquisition field is trying to solve is how children acquire sophisticated grammatical systems such as morphology. The study of inflectional morphology is probably one of most researched topics in the field, particularly in terms of productivity. It is still unclear, though, what being a productive speaker really means and what the underlying mechanism to achieve productivity is. Morphology research usually adopts one of two theoretical perspectives: a generativist approach, that argues for a language-specific learning system and promotes relatively high levels of productivity at the very early stages of acquisition; and a usage-based approach, that is more conservative with its assumptions regarding children’ productivity by arguing for gradual, slowly emerging acquisition.

1.1. Generativist approaches to productivity

As mentioned, one approach is to assume some degree of innateness in the language faculty, inspired by humans’ ability to produce (i.e. generate) and comprehend an non-finite number of novel utterances without ever having heard them previously. With respect to morphological systems, generativist approaches assume that once a morpheme has been acquired, children can apply it to all or most relevant contexts without significant omission rates or replacements with another morpheme (de Villiers & de Villiers, 1973; Hoekstra & Hyams, 1998; Wexler, 1998). An important implication following the assumption of early productivity is that even young children are expected to show very low error rates with those inflections they have already learnt. Another implication is that very young children are expected to generalise the suffixes familiar to them to lexical items, e.g. to
verbs or nouns, that they may have not heard inflected in the required target form before. Potential errors are regarded as insignificant processing errors due to high cognitive load or just as noise in data (Wexler, 1998).

In order to investigate the prediction of children’s error rates, Hoekstra & Hyams (1998) reviewed children’s subject-verb agreement marking (e.g. *She* drives a car; *They are* playing football) cross-linguistically in e.g. Spanish (Serra & Solé, 1992) and Italian (Pizzuto & Caselli, 1992). It was found that children’s error rates were very low, remaining under 5%. These findings were concluded to support the generativist approach as the error rates were considered low enough to demonstrate children’s productivity with subject-verb agreement marking. Also Marcus (1995) found that children’s (aged 1;3-5;2) overall error rates were very low. The overall average across ten children’s plural noun marking was 8.5% in naturalistic data.

There are differences amongst generativist approaches in how much of an impact input is considered to have. Yang (2002) argues that children are innately equipped with the same parameters (e.g. plural marking or verbal tense marking) independently of the language they are going to learn. These parameters are set appropriately during the acquisition process based on input: the more there is evidence for a particular feature of grammar in input, the more likely it is that the mental parameter for that kind of grammar will be switched ‘on’. If evidence for a grammar is lacking, the parameter for it will be switched ‘off’ by probabilistically losing to the ‘on’ setting. For example, Legate & Yang (2007) compared children’s acquisition of tense marking in Spanish, French and English by investigating the
proportion of verbs with overt verbal tense marking in input. It was predicted that the more evidence there is for 3SG marking in the input language, the faster the child’s parameters set accurately. The model predicted the order of languages in terms of speed of 3SG acquisition correctly. An issue, however, is that it cannot explain variation in error rates between different verbs (Freudenthal, Pine & Gobet, 2010), which has been ignored by the other generativist studies mentioned above. The findings also do not adequately explain why a language-specific parameter setting mechanism is required over a cognitive-general statistical learning mechanism.

1.2. Usage-based approaches

A different approach to studying child language is to assume that cognitive-general mechanisms such as pattern detection and cross-situational learning provide the basis for language learning. This approach is motivated by the apparently limited knowledge of grammatical systems displayed in children’s early utterances and comprehension. With respect to morphology, usage-based approaches argue that productivity cannot be assumed from the earliest uses of morphemes, even when children seem competent at using a morpheme in a particular context. Children initially learn frozen phrases that contain inflected forms, e.g. “it goes bang!” directly from input. Increasing language experience will enable children to make generalisations across these lexically-specific utterances by making analogies between utterances that share similar properties, e.g. substituting the pronoun ‘it’ with a noun. Some of the nouns the child is hearing most frequently might be in plural more often than in singular and thus the child
might use the plural form as a chunk. Hearing a number of plural forms, e.g. dog+s and flower+s will contribute towards learning a general [NOUN]+s construction. Sufficient exposure will eventually develop a partly-productive system into a fully productive one, that is equivalent to adult productivity (Tomasello, 2000a; Tomasello, 2000b).

As the approaches emphasise the gradual nature of the acquisition process, a central assumption is that children do not show full or almost full productivity from the early stages of inflectional morphology acquisition. Children’s inflected forms are considered to be at least somewhat lexically specific: even when they use a particular inflection with a few lexical items, a fully productive system should not be assumed based on these instances. A logical prediction based on these assumptions is that children have higher error rates than suggested by generativist research. Additionally, when young children seem to go through an error-free stage, it cannot be taken as evidence for full productivity: children may simply be retrieving rote-learnt lemma + suffix combinations directly from memory (Tomasello, 2000a; Tomasello, 2000b). There is a fair amount of research supporting the claims of gradual acquisition, partial productivity and errors that are an indication of restricted productivity rather than just increased cognitive load. Wilson (2003) investigated in what contexts English-speaking children omit grammatical words and morphemes. This was achieved by studying five English-speaking children’s spontaneous, naturalistic speech between the ages of 1;6-3;5. He examined three types of forms marked for tense and/or agreement: copula BE, which is a verb type linking the subject to the predicate, describing the subject (e.g.
I am your friend; The cat was under the table); auxiliary BE; and finally, third person singular agreement. He found that there were significant individual differences between the children and between different inflections themselves in rates of provision. For example, one of the children used copula BE correctly for 79% of the time, while the correct rate of using the third singular –s was only 13%. This illustrates the gradual nature of the language acquisition process and how children first begin the process from lexically-specific constructions before abstraction.

Supporting findings for the piecemeal progress of language development have been reported in a number of other studies. For example, Spanish children’s verb morphology is significantly more restricted compared to their caregivers (Aguado-Orea & Pine, 2015). They begin the inflection acquisition process with individual lexically-specific inflected forms (Gathercole, Sebastián & Soto, 1999) and similarly, Italian children aged 1;5-3;0 use 47% of the verbs in their speech with only one inflection despite five other possible forms (Pizzuto & Caselli, 1992). Further strong evidence for the gradual, lexically specific nature of language learning is provided by a cross-linguistic study that utilised computational modelling to examine children’s tendency to omit third singular verbal marking in compulsory contexts. As the usage-based approach assumes that children initially rote-learn chunks of language directly from input, the authors investigated whether omissions of tense agreement marking result from adult compound finite forms (e.g. ‘He can play’, ‘Will he play?’). Children’s working memory is particularly susceptible to a recency effect (Jarrold et al., 2015; Roome, Towse & Jarrold, 2014) which in turn influences the kinds of language strings children tend to learn from input. Based on these
principles, Freudenthal, Pine & Gobet (2010) compared third singular marking omissions (i.e. root infinitive errors) in English, Dutch, German, Spanish and French, by creating a computational model that learns utterances gradually from the end of utterances to the beginning. When the model is presented with real child-directed speech from naturalistic corpora, it starts by learning the end of an utterance, such as ‘play’ from the utterance ‘Can he play?’. Over time, when the utterance is repeated multiple times, the model will learn the next item, which in this case is ‘he’, leading to a stored chunk with a root infinitive error *he play. The model successfully mirrored the rates of infinitive root errors in four of the studied languages. Importantly, the model also showed a correlation between specific verbs that were used in compound finite forms in adult speech and the rate of infinitive root errors in child speech. The model shows the importance of accounting for morphology learning in terms of input and cognitive processing limitations.

1.3. Issues with defining and measuring productivity

A critical issue older research on inflectional morphology productivity has suffered from is the lack of a clear definition of what productivity truly means. An intuitive approach would be to claim that morphological productivity equates to error-free adult performance with inflected word forms – the ability to reliably provide target inflections in obligatory contexts. The underlying problem with this assumption is that it is known what adult productivity is like, when in reality not that much is known about the extent of an adult speaker’s spontaneous inflectional morphology use. Do adult speakers use all available suffixes? Do adults, like
children, show lexically-specific use of inflectional morphology? To control for this issue of lack of a baseline for inflectional morphology productivity, child productivity studies need an adult control sample as a more reliable way of measuring productivity. For example, it has been shown that even adults use different inflections at different rates and somewhat restrictively (Aguado-Orea & Pine, 2015). Instead of asking whether young children are productive or not, the question should be whether young children are as productive as adults.

Comparing a child speaker’s use of inflectional morphology to that of an adult speaker also eliminates sampling bias, but only when both samples are carefully controlled for a) types b) tokens and c) suffixes. Controlling for all three is important in order to make a meaningful comparison of child and adult productivity. Yang (2013) has argued that the previously reported observations of lexical specificity (e.g. Pizzuto & Caselli, 1992) merely demonstrate Zipfian distribution of words in naturalistic speech – the fact that relatively few items occur with high frequency and most items occur with low frequency - instead of lexically-specific knowledge of inflectional morphology. Controlling for vocabulary and sample size in both child and adult samples controls for this lexically specific appearance of naturalistic speech samples.

When only the same types (e.g. nouns or verbs) shared by both speakers are included, the data is not skewed due to the fact that an adult speaker is much more likely to use a wider variety of types, some of which a child might not even know yet and thus would not use spontaneously. This is related to point b), controlling for token frequency. Even when type frequency is controlled for, an adult sample
would be unfairly skewed simply because adults tend to speak more than young children, which results in more opportunities to use particular types as well as suffixes. Finally, controlling for suffixes ensures that the comparison occurs between those exact suffixes that both samples have in common. Again, this relates to the type and token frequencies: if the comparison was conducted without controlling for these factors, adult speakers might seem more productive than children due to the fact that they simply have more opportunities to use a variety of types, use them more often and with a bigger number of inflections.

The final issue relates to vocabulary. As mentioned earlier, usage-based approaches assume that children’s early language is lexically specific. Children become abstract as they are exposed to an increasing and vast sample of exemplars in their input of how e.g. a particular suffix can be combined with a majority of noun types. It is known that type and token frequencies of lexical items facilitate the transition from a fairly closed construction to an abstract one. When there is more evidence for a construction use with a variety of types in input, the concrete exemplars of that construction are more likely to be abstracted across each other compared to a construction with fewer types (Bybee, 1995; Bybee & Johnson, 1997). Older research on inflectional morphology productivity used to collapse error rates across all lexical items and suffixes together, which hid the lexically-specific nature of children’s early utterances. This way children’s knowledge was easily overestimated. To tackle this, a meaningful test of productivity will be sensitive to the differences between children’s knowledge of different lexical item + suffix combinations. Additionally, appropriate controls can show whether children’s
speech samples only seem lexically specific due to Zipfian distribution or whether there really are underlying differences between child and adult productivity with inflectional morphology.

1.4. **Most recent measures of productivity**

Aguado-Orea & Pine (2015) partly filled the void of carefully controlled naturalistic productivity analyses by investigating verb inflection in child Spanish. Spontaneous speech of two children aged 2;0-2;6 and their parents was video-recorded to compare child and adult productivity in verb inflection. This was done by matching the adult and child samples carefully to control for potential confounding variables like sample size and vocabulary. The average number of different inflections, i.e. person+verb combinations, was calculated for each verb for both samples which were compared against each other. They found that even though the adults’ verb inflections were somewhat lexically restricted, the children were significantly less productive. Additionally, despite low overall error rates, systematic patterns of errors were observed. As predicted by the constructivist approach, the children showed significantly higher error rates with low frequency verbs compared to those with higher frequencies.

Following the methodology of Aguado-Orea & Pine (2015), Krajewski et al. (2012) carried out a carefully controlled analysis of Polish noun morphology to examine case marking productivity between a young child (2;0-2;1) and her mother. They conducted three analyses to investigate this. First, they studied the frequency distributions of noun inflections in the child’s and the mother’s sample. Although all possible case-number combinations were present in both samples and they were
distributed fairly similarly, there were inflection errors in the child sample that deviate from the adult sample. Second, they evaluated how flexible the child’s noun inflection productivity was. The method was similar to the way in which verb inflection in child Spanish was examined: the number of inflections used per noun by the child was directly compared to the adult sample after carefully controlling for the variables outlined above. The results showed that even though the relative frequency distributions of noun and case endings used in the child and caregiver samples were very similar, the child’s noun inflection productivity was significantly more restricted compared to that of the mother. Finally, they examined how contextually flexible the child’s noun morphology was. Whereas the number of inflections per noun was utilised as a measure of productivity in the second analysis, the number of nouns per inflection was used to measure productivity in the third analysis.

Another way of investigating the nature of children’s productivity has been to examine English-speaking children’s determiner use. Pine, Freudenthal, Krajewski & Gobet (2013) compared strictly controlled samples of two- and three-year-old English-speaking children’s language and their caregivers’ child-directed speech to examine the extent of overlap in their determiner use. This was achieved by controlling for the vocabulary, lexical items and sample size. The authors’ findings demonstrate the importance of these controls well. It was found that the overlap measures were sensitive to sample size and particularly to the specific nouns. This shows the importance of controlling for vocabulary in addition to sample size when productivity is studied naturalistically. When vocabulary and
sample size were controlled, there were significant differences between the adults’ and children’s determiner use: the children’s uses were more restricted than their caregivers. These findings demonstrate the gradual nature of child language acquisition, as predicted by the usage-based approach.

Most recently, strictly controlled, naturalistic analyses on productivity showed how creativity and flexibility of subject and object noun slots in verb frames differ between child and adult speech (Theakston, Ibbotson, Freudenthal, Lieven & Tomasello, 2015). The aim of the study was to examine different measures of productivity of English-speaking children’s noun-verb-noun (NVN) frame. The measures examined flexibility and creativity in naturalistic corpora of children aged two to three years and their caregivers. The authors carefully matched the verb types, noun types and sample sizes between the child and adult samples. This is critical as the results cannot be accounted for by differences between the actual sample sizes or caregivers’ more advanced vocabulary. Children were shown to be less flexible with subjects in both developmental stages compared to child-directed speech, while they were considerably more proficient with objects. Their object use resembled the CDS samples at the later developmental stage.

The current study is a naturalistic productivity analysis comparing a young Finnish child’s (1;7-1;8) and her mother’s production of noun inflection. Morphology provides a useful test case for determining the extent of early productivity. It enables the investigation of contrastive use of a quantitatively controlled pool of noun+case combinations. When these combinations are
compared against adult language use (Krajewski, Lieven & Theakston, 2012), the extent of child productivity can be measured carefully.

There are a number of benefits to studying noun morphology acquisition in Finnish. Finnish has considerably more complex noun morphology compared to English and, in fact, to most languages of the world (Iggesen, 2013). The Finnish case marking system consists of 14 to 15 cases, depending on how they are counted. It is easier to study productivity with inflections in a morphologically rich language compared to a language with very limited morphology as lack of productivity is more readily detectable. In addition to avoiding sampling bias that would skew the data to potentially over/under-estimate productivity for the reasons outlined earlier, appropriate sampling is crucial for exposing items with low frequencies and for detecting early usage. Sparse data, e.g. a recording scheme of one hour a month, is not sufficient to investigate differences between child and adult language as low frequency items and multiple exemplars of the same items might not be documented. Distributions of naturalistic speech samples tend to consist of a few lexical items with very high frequencies while the majority of items have a very low frequency. Even adult speakers do not exhibit a full range of possible inflections in naturalistic samples. The implications are that sampling of naturalistic data plays a vital role in capturing items with low frequencies and the earliest possible uses of items, e.g. different inflected forms (Pine et al., 2013).

1.5.Earlier naturalistic studies on child Finnish

There have been previous attempts to study morphology acquisition in child Finnish naturalistically but with equivocal results. For example, Bowerman (1973)
conducted a longitudinal study on two Finnish children: Seppo (a boy) from 22.5 months of age for two years and Rina (a girl) from 24 months of age for two years and eight months, each recorded for 30 minutes per week (Bowerman, 1973). Despite starting the data collection relatively late, the data did not contain evidence of any nominal morphology other than rote-learned forms during the earliest stages of data collection. Both children almost invariantly used the singular nominative forms of nouns. The only exceptions were the partitive forms of ‘snow’ (lumi-NOM; lu+nta-PART) and ‘water’ (vesi-NOM; ve+ttä-PART) used by Seppo (Bowerman, 1973). These nouns are rarely used in the nominative, as the contexts the words usually occur in call for the partitive inflection, e.g. ‘I’m drinking water’ [Minä – I; juon – drink; ve+ttä – water + PART], ‘there is water in the glass’ [Lasissa – in the glass; on – is; ve+ttä – water+PART], ‘there is snow outside’ [Ulkona – outside; on – is; lu+nta – snow+PART], ‘it is snowing’ [Sataa – it is raining; lu+nta – snow+PART] (literally, ‘it is raining some snow’). It seems likely that Seppo rote-learned the partitive forms of ‘water’ and ‘snow’ without comprehending the semantic differences between their nominative and the partitive forms.

The most extensive naturalistic study of inflectional morphology in child Finnish is by Toivainen (1980). He investigated the early acquisition of inflectional suffixes in 25 one- to three-year-old children by recording them in their home environment for an average of fifteen minutes on a weekly basis. The median age of acquisition for each documented suffix was calculated. It was observed that Finnish children start using inflections from the age of 15 months, followed by the first case marker suffix at the median age of 1;8 (Toivainen, 1980). Finnish-speaking
children’s earliest inflections were in the partitive case with nouns that have a vowel stem e.g. leipä+ä-PART ([some] bread-PART) as opposed to nouns with a consonant stem e.g. ananas+ta-PART ([some] pineapple-PART). However, the study did not address or measure the extent of productive use, and these early partitive forms are mostly likely retrieved from memory.

Even though these studies on child Finnish provided some insights into the possible patterns of morphology acquisition, they were based on sparse data which allowed only tentative conclusions to be drawn. There is a need for richer and denser naturalistic corpora that capture low frequency items that have previously been challenging to examine.

Table 2.1. The Finnish noun case marking system.

<table>
<thead>
<tr>
<th>Case</th>
<th>Suffix</th>
<th>Meaning</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominative</td>
<td>Null-marked</td>
<td>The bare form</td>
<td>Auto (car)</td>
</tr>
<tr>
<td>Partitive</td>
<td>-(tt)a ; -(tt)ä</td>
<td>Incomplete direct object</td>
<td>Autoa</td>
</tr>
<tr>
<td>Accusative</td>
<td>Null-marked or -n</td>
<td>Complete direct object</td>
<td>Auto; Auton</td>
</tr>
<tr>
<td>Genitive</td>
<td>-n</td>
<td>The possessive</td>
<td>Auton</td>
</tr>
</tbody>
</table>

1 Traditionally there are 15 cases but more recently linguists have argued that the genitive and the accusative should be considered one case instead of two separate ones as their endings somewhat overlap. This thesis treats them as two individual cases in our study as they mark for different argument structure roles: the genitive marks for possession whereas the accusative marks for direct objects.
<table>
<thead>
<tr>
<th>Case</th>
<th>Form</th>
<th>Description</th>
<th>Autonym</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Inessive</strong></td>
<td>-ssa / -ssä</td>
<td>Internal locative; same as the preposition ‘in’</td>
<td>Autossä</td>
</tr>
<tr>
<td><strong>Elative</strong></td>
<td>-sta / -stä</td>
<td>Movement from a location; same as the preposition ‘from’</td>
<td>Autosta</td>
</tr>
<tr>
<td><strong>Illative</strong></td>
<td>-an, -en, -on etc. (depending on the stem)</td>
<td>Movement into a location; same as the preposition ‘into’</td>
<td>Autoon</td>
</tr>
<tr>
<td><strong>Adessive</strong></td>
<td>-lla / -llä</td>
<td>External location: same as the prepositions ‘in’ &amp; ‘at’</td>
<td>Autolla</td>
</tr>
<tr>
<td><strong>Ablative</strong></td>
<td>-lta / -ltä</td>
<td>Movement from one external location to another: same as the preposition ‘from’</td>
<td>Autolta</td>
</tr>
</tbody>
</table>
### 1.6. The aim of the study

The aim of our analysis is to examine a young Finnish child’s, Mailis’s (1;7-1;8), knowledge of noun inflection and compare her productivity to that of her mother. There were three reasons for conducting this study. First, in the study of morphology, a disproportionate amount of the literature has been devoted to English verbal morphology which, by virtue of its limited range of forms, does not provide a good test bed for degrees of productivity in children’s language (being more suited to the study of the nature and prevalence of children’s errors). Thus investigating morphology acquisition is not ideal in English due to the restricted...
nature of its inflectional system. There has been some improvement in this regard due to the dense Spanish and Polish corpora discussed above. Investigating noun productivity in Finnish will therefore make a valuable contribution to the field. Second, the Finnish case marking system is more transparent than, for example, the Polish system due to the lack of gender marking in Finnish. In languages with both gender and case marking it is challenging to tease apart their individual effects on morphology acquisition. Third, earlier naturalistic studies (Bowerman, 1973; Laalo, 1994; Niemi & Niemi, 1987; Toivainen, 1980) suffered from a sampling problem that skewed the findings to support the idea of a more instantaneous and error-free acquisition process. Children’s early use of inflections could reflect properties of their caregivers’ language instead of a child’s underlying knowledge. Lastly, young children simply know fewer inflections due to not having learnt them yet compared to adult speakers of the language they are acquiring, which restricts them from making the necessary form-meaning mappings sooner. Our study is in line with the kinds of controls implemented in the studies on Spanish (Aguado-Orea & Pine, 2015) and Polish (Krajewski et al., 2012), and therefore we utilise the same criteria in our analysis of child productivity with Finnish noun morphology, comparing data from a mother and a child. When these criteria are applied, the results cannot be accounted for by other factors such as restricted lexical knowledge.
2. Method

2.1. Participants

The data were collected from a monolingual Finnish family with a 1;7-1;8 month old child (Mailis) in Joensuu, East Finland. The parents are highly educated and represent a typical middle class family. They are originally from the town of Raahe in West Finland and speak a dialect typical to that region. Mailis was influenced by both West and East Finnish dialects during the data collection period and showed characteristics of both dialects, mostly vocabulary. It is important to note that the dialects do not differ in noun case marking.

2.2. Procedure

The recordings started when Mailis was 1;7.23 and ended at the age of 1;8.18. According to previous research (Laalo, 1994; Niemi & Niemi, 1987; Toivainen, 1980), Finnish children will have started uttering some case forms by this age but it is unclear whether their usage of case marking is adult-like or reflects more limited productivity or rote-learning. The mother was provided with an audio recorder and asked to record interactive play situations at home between her and Mailis for a total of 18 hours. The same set of toys was used in approximately half of the recordings in an attempt to promote repeated use of vocabulary items but the reason was not disclosed to the mother to keep her speech as natural as possible. The data were recorded in sessions of one hour, five hours a week for a period of four weeks. Each hour was recorded on a separate day. The dense recording scheme was chosen to address some of the problems that previous naturalistic
studies on child Finnish encountered. These were based on very limited data which reduces the reliability of productivity analyses.

2.3. Transcription and coding

The recordings were transcribed in CHAT format (MacWhinney, 2000) with a purpose-specific coding scheme. All nouns were coded for their case, number and case marking errors, if they occurred. Imitations, repetitions and partly intelligible utterances were excluded from the analysis. All nouns were extracted from the mother and the child samples, along with their frequencies of occurrence. Frequencies were extracted separately for singular and plural inflections.

The analysis applied in this study was based on those used in investigations into the acquisition of verb inflections in Spanish (Aguado-Orea & Pine, 2015) and Polish noun inflections (Krajewski, Lieven & Theakston, 2012). To examine the relative productivity of use of noun inflections, we compared data samples collected from the child and her mother. However, two important controls were applied to the data to match the speakers on vocabulary and sample size to ensure a fair comparison.

The first step to control sampling bias was to match the speakers on vocabulary. Adults have a larger vocabulary than young children, and consequently we are almost bound to find a larger number of noun types in the mother’s data. Different nouns appear in different contexts and thus with different noun inflections, so differences in the usage of noun inflections between the mother and child could simply reflect differences in the nouns found in each sample. To control
for this, lists of all the nouns produced by each speaker were created and only the noun types shared by both samples were included in the analysis. Types that did not occur more than once were eliminated to ensure the speakers had an opportunity to use them with more than one inflection. Similarly, as with noun vocabulary, if adults have a larger range of inflections at their disposal, they have the opportunity to use more different inflections with their nouns. Comparing the number of inflections used by the child against an adult sample is particularly important in studying the extent of productivity. In previous studies, the criterion for productivity was set at the use of a particular inflection with three different lexical items. But this measure is, of course, arbitrary and does not exclude the possibility that those forms could have been rote-learnt when not combined with novel nouns. For this reason, it is vital to compare the child and adult samples to examine the extent of the child’s knowledge in our study. To control for this, only same inflections attested in both samples were included. Singular and plural forms of any case were treated as separate inflections. For example, the mother had used both singular and plural forms of the genitive whereas the child used only the singular genitive inflection. Thus, the plural genitive forms were excluded from the adult sample. It is also worth noting that there are three different accusative endings: a marked singular form, an unmarked singular form and a plural form. The marked singular form is used in contexts with an agent and a patient (“I-NOM want to put a hat-ACC on”) whereas the unmarked singular form is used in contexts without an agent such as imperatives (“Put the hat-ACC on”).
The second step to control for sampling bias was to match the sample size. Adults speak more than children, so in any given recording we are likely to sample a larger number of noun tokens, and thus opportunities to display productive use of noun inflections, in the mother’s speech than for the child. A larger number of tokens increases the average number of inflections per noun type in the adult sample, making it tempting to conclude that the mother is more productive than the child. To control for this, the mother’s vocabulary-matched sample was matched to the child sample for the number of tokens included for each noun type across the child and mother. This was achieved by randomly reducing the number of noun tokens for each type down to the smaller number in either the child’s or caregiver’s speech on a noun-by-noun basis. The process of matching the number of tokens per type was repeated five times to control for the possibility that any individual randomly reduced sample may be unrepresentative of the mother or child’s overall use of inflections with a given noun. Five random reductions were considered sufficient as the variation between the five reduced samples was minimal.

3. Results

Altogether 10,067 noun tokens were extracted from the corpus: 3,414 tokens from the child sample and 6,653 tokens from the adult sample. Four hundred-forty-eight noun types were used by both speakers but this reduced to 293 types used more than once by both speakers. After the vocabulary reduction process, the child and adult sample sizes consisted of 3,246 and 6,038 tokens, respectively. The adult sample was then randomly reduced to match the size of the child sample in terms
of overall noun tokens, and then the number of tokens per noun type, precisely, as
described above.

**Table 2.2. Descriptive statistics for the controlled child (Mailis) and mother samples.**

<table>
<thead>
<tr>
<th></th>
<th>Mailis</th>
<th>Mother</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean number of nouns</strong></td>
<td>255.8</td>
<td>255.8</td>
</tr>
<tr>
<td><strong>Mean number of inflections per noun (± SD)</strong></td>
<td>2.21 (±1.55)</td>
<td>2.85 (±2.12)</td>
</tr>
<tr>
<td><strong>Range of inflections</strong></td>
<td>1-10</td>
<td>1-12</td>
</tr>
</tbody>
</table>

**Figure 2.1. Type frequencies distributed across numbers of inflections for the child**

and mother samples (averaged across five samples)
Figure 2.1. shows how many noun types occur on average with different numbers of inflections in the mother sample before and after controls, and in child sample after controls. The distribution is heavily skewed to the left, showing that both the mother and the child used the majority of noun types with only one or two inflections. In fact, there are relatively few types that were used with more than four inflections, which makes the distributions fairly similar. The figure also shows that the data are skewed to the left with a similar overall distribution before and after the controls in the mother sample. Despite the overall similarity, the figure demonstrates one key difference before and after the application of the controls: the mother’s inflection use is skewed to the left even more before the controls, seeming slightly more lexically specific, than after the controls were applied. Before the controls, the mother used a slightly larger number of noun types (109) with one inflection only than compared to the number of types (104) used with two inflections. The pattern is reversed when the controls were applied: the mother used more noun types (79.6) with two inflections than with one inflection (77 types).

Because the data were skewed, a Wilcoxon rank sum test was utilised to compare the mean numbers of inflections per noun between the mother and child. This revealed a significant difference in noun inflection productivity between the two samples (Z = -7.007, p < .001), with the mother producing a wider range of different inflections with her noun types than observed in the child’s speech. This suggests that the child’s knowledge of noun inflection is more restricted than that of her mother. Lastly, we also ran a correlation between the number of inflections
per noun between the two samples to find out whether it is mostly the same nouns that are used in a greater number of inflections or with a smaller number of inflections by both speakers. A moderate positive correlation was found between the number of inflections per noun between the speakers, \( r(152) = .51, p < .001. \)

4. Discussion

The aim of the study was to examine whether a young Finnish child is as productive with noun morphology as her caregiver. The findings were evaluated in the light of what productivity really means and how they fit with the usage-based approaches. Our productivity analysis comparing the child’s and the mother’s speech samples showed that even when the number of nouns, sample size and number of inflections are controlled for, the mother was significantly more productive with noun inflections than her child aged 1;7-1;8. The finding reflects the gradual and item-based nature of first language acquisition, showing that the child did not use each noun with as wide a range of inflections as the mother even for the same inflections and the same nouns. It is important to note that the difference cannot be accounted for by a sampling bias, differences in lexical knowledge or by other factors that were controlled for in the study.

The finding is interesting because research has provided evidence for a link between morphological complexity of a language and speed of inflection acquisition (Xanthos et al., 2011). While the Finnish child included in Xanthos et al. study was not the most morphologically advanced child out of the other children
speaking other languages who were included in the study, the results showed her 
development of noun and verb inflection was moderately advanced between the 
ages of 1;7-3;0 compared to children acquiring e.g. German and Dutch. Even though 
Mailis used a seemingly impressive number of inflections (10) for a child younger 
than two, the fact that they were used in a significantly more restricted way 
compared to the mother implies that some of the inflections must have been rote-
learned or only partly productive. We would not want to argue that all of the child’s 
noun+case combinations were rote-learned but rather that it is plausible to think of 
her noun inflection representations as partly productive. This is consistent with 
constructivist/usage-based approaches which predict that the language acquisition 
process is lexically-specific and gradual. These approaches also argue that children’s 
seemingly versatile inflectional use is somewhat deceptive since it creates an 
ilusion of a more flexible, adult-like productivity than is actually the case. Our 
results show that a child does not generalise case endings to as many different 
nouns as evidenced in the input in the earliest observable stages of morphology 
development.

Our findings are similar to those reported by Krajewski et al. (2012) and 
Aguado-Orea & Pine (2015). Krajewski et al. (2012) found that a two-year-old 
Polish child was restricted in her use of noun case marking both lexically and 
contextually. Aguado-Orea and Pine’s (2015) investigation into two Spanish-
speaking children’s productivity, starting from 2;0, with verb morphology showed 
how these children’s verb inflection was partly productive at best even though the 
adults’ speech seemed somewhat lexically restricted too. One important difference
in our study compared to those in Polish and Spanish is the age of the child. While those two studies start when the children are precisely 2;0, Mailis is only 1;7 at the beginning of our study. Finnish children have been reported to use case marked nouns, rote-learned or not, before the age two (Bowerman, 1973; Laalo, 1994; Toivainen, 1980). By examining productivity of a child even younger than in the previous studies, we have pushed the boundaries of what the ‘earliest observable stages’ of morphology acquisition are. This is needed to examine early child language more carefully to reveal the relevant underlying mechanisms that may help us to distinguish between theoretical accounts.

Another finding which provides support for the constructivist/usage-based approach was that even though there was a significant difference in productivity between the two samples, the child’s and the mother’s frequency distributions resembled each other. The correlation between the mean number of inflections per noun in the two samples shows that both speakers tend to use the same nouns more or less flexibly. This finding supports the idea that child language closely reflects the properties of child-directed speech, up to the point of being more productive with particular lexical items while showing less productive use of other items. As the usage-based/constructivist approaches emphasize the role that input frequencies play in the acquisition process (Ellis, 2002), it is hardly surprising that the child’s distribution frequencies reflect those of the mother.

Overall, we have shown that dense naturalistic data can be informative with regards to the extent of children’s productivity as long as variables such as sample size are controlled for appropriately. Naturalistic studies often fail to control
relevant variables, partly due to the kinds of data collection and analysis methodologies adopted by earlier child language research. An implication of our study for the language acquisition field is that naturalistic studies can provide us meaningful, quantitative and controlled data – to the extent that they were sufficient to measure the quantitative difference in noun inflection productivity between the two samples. This analysis adds to the ever-growing body of evolving naturalistic research on productivity also by demonstrating the effectiveness of this methodology with a morphologically rich language. These kinds of corpora are important for development of computational models of morphology acquisition as they could be used for testing theoretical accounts, designing future studies and developing evidence-based predictions for them. Particularly with data from somewhat later points in development, it would be possible to predict the developmental trajectory of a child’s morphology acquisition, something which is currently lacking in the field. For instance, it is yet not known when precisely children reach adult-like levels of inflectional productivity. Utilising dense naturalistic corpora in modelling would allow further refinement of the usage-based approach as the question of when children make syntactic generalisations in adult-like ways is one of the most central in the child language field.

A limitation of the constructivist and usage-based approaches is that they lack a detailed account of the acquisition process of a complex inflectional system. For example, past research has shown the importance of frequency and of analogies between constructions in language development (Gentner & Namy, 2006; Tomasello, 2000a; 2003), but there are no explicit explanations of how they operate.
Particularly the literature on frequency effects overwhelmingly shows that items with high frequencies tend to be learned before low frequency items (Berko, 1958; Bybee & Thompson, 1997; Ellis, 2002). While frequency effects are clearly important in language acquisition, they must be integrated in the current models more effectively to show how they interact with other aspects of learning, such as working memory and attention. In order to acquire language, it is essential to attend to the right property of language to learn it to begin with - e.g. to attend to morphology to learn it - and for working memory to encode instances of inflected words to store a sufficient amount of frozen phrases before abstraction can happen. The roles of implicit and explicit learning are also unclear (Ambridge, Kidd, Rowland & Theakston, 2015; Ellis, 2002). Future research should therefore focus on developing accounts of complex morphology acquisition that can combine a variety of effects and learning mechanisms reliably that can also predict developmental changes.


Suomalaisen Kirjallisuuden Seura.


Rationale for Study 2

Study 1 presented findings regarding noun morphology productivity of a young Finnish child. The study compared the average number of inflections used per noun in the child and the caregiver samples to examine the flexibility of both speakers’ case marking. The results showed that the child was more restricted in her inflection use than her caregiver by using fewer inflections per noun on average compared to the adult control. This finding adds to the ever-growing evidence for the gradual nature of language acquisition and how deceptively productive young children can seem until examined against a control sample. These findings support the usage-based approach that predicts child language to be

Study 2 was inspired by the first study. Study 1 investigated the child’s correctly produced inflections. The next study will examine the role of various input frequencies and what they might reveal about the child’s accuracy and error patterns with case marking. The role of input frequency is well established in the child language literature, but the impact of different types of frequencies is still unclear. This inspired us to examine the impact of various input measures on a young Finnish child’s accuracy with case marking in obligatory contexts and the likelihood of overextending case endings to incorrect contexts. We explored the effects of absolute noun and case token frequencies, absolute noun and case type frequencies, as well as absolute and relative (%) frequencies of noun+case ending strings to examine competition between these strings. Altogether seven measures were developed to examine these questions more closely.
Chapter 3: Naturalistic study 2: Case marking errors

Abstract

Previous literature of child language acquisition shows that input frequency has a significant role in productivity. Studies have shown that children are sensitive to different aspects of frequency in different stages of development. Our analysis focusses on investigating what aspects of input frequency a young child is sensitive to in Finnish noun inflection. We used a dense naturalistic corpus of mother-child interaction, collected between the ages of 1;7-1;8. The study examines seven measures of input frequency correlated with the child’s accuracy with case marking. The results showed that the child was more accurate with case marking with nouns that have high input token frequency than with nouns with low input token frequency, and that high token and type frequencies were associated with the likelihood of the child generalising a case ending to incorrect contexts. The findings partly support previous work regarding the role of frequency in language acquisition, but have implications for the further development of child language theories and which types of frequency young children are sensitive to in input.
1. Introduction

The question of how children acquire the knowledge to combine stems and appropriate inflections in obligatory contexts is widely debated. The study of children’s error patterns can provide insight into their linguistic representations. In this context it is important to examine the relationship between the input from a child’s language environment and in what ways this may relate to the child’s output. Examining child-directed speech can reveal the roots of systematic errors in child language. There is cumulative evidence showing that input frequency has a significant role in how children acquire inflectional morphology (Ellis, 2002; Ambridge & Lieven, 2011; Ambridge, Kidd, Rowland & Theakston, 2015). However different aspects of frequency have been shown to be important in different studies and for different aspects of morphological learning.

Type and token frequency has been shown to have different effects (Bybee & Thompson, 2000). While the two do interact, type frequency in particular, seems critical to morphological productivity. Type frequency is the measure of how many individual items can be used in a morphophonological construction (e.g. number of different nouns in the nominative case), and it is linked to how easily the construction is generalised to other appropriate types. Other things being equal, the higher the type frequency, the easier it should be to generalise the construction to other appropriate types (i.e. to use newly learned nouns in the nominative even if they have not been heard in the nominative). Morphological constructions with high type frequencies are used more regularly which strengthens the schema of that particular construction and promotes productivity (Bybee & Thompson, 2000;
An implication of knowing that type and token frequency facilitate productivity is that a child should be more accurate with highly frequent noun+suffix combinations compared to noun+suffix combinations with low frequency. The principle should similarly apply to case markers: children should make fewer errors with highly frequent case endings compared to markers with other frequencies. Bybee’s (1985; 1995) network model of inflectional morphology acquisition had shown that a shared affix between multiple different lexical items in input will link these lexical items together in a child’s lexicon. This will consequently further the child’s correct use of that affix.

In fact, input frequency of individual inflections has been shown to affect children’s accuracy with inflections. A critical question regarding the acquisition of a language with complex morphology is the interaction between type and token frequencies. Constructions (e.g. [NOUN]+ a particular case marker) with low type frequency are generalised to new lexical items more readily when the construction has been used with sufficiently phonologically distinctive lexical items in the past (Dąbrowska & Szczerbinski, 2006).

On the other hand, token frequency has an important role in the acquisition of irregular inflected items, as demonstrated by Bybee’s (1995) network model. High token frequency facilitates the acquisition of e.g. irregular verbs, as they otherwise would be generalised according to the regular –ed inflection paradigm. High token frequency of irregular inflections supports the construction of strong lexical representations of the items that are inflected with an irregular pattern. High token frequency is also argued to promote lexical strength. Items with high token
frequency are acquired independently of other lexical items, whereas items with low token frequency are acquired in relation to previously learned items (Bybee, 1985; Bybee, 1995). Due to the rote-learning dimension of high token frequency items, this kind of frequency is considered not to contribute towards accurate, productive use of inflectional morphology. Differences in productivity have been found between inflections with similar type frequency but different token frequencies: high token frequency has been reported to in fact hinder morphological productivity (Francis & Kučera, 1982).

Competition between available inflections is another important factor to consider in the acquisition of inflection morphology. The more competition there is, e.g. multiple possible inflections associated with a stem, the greater the learning challenge. Taraban & Roark (1996) investigated competition between French feminine and masculine adjectives in non-French speaking adults. The participants learnt the same set of nouns in either a high competition condition or in a low competition condition. The feminine adjectives were the same in both conditions whereas the masculine adjectives were different, which also manipulated the rate of competition between the two classes. Unsurprisingly, participants in the high competition condition learnt the feminine nouns more slowly than the participants in the low competition condition. Cue reliability, which refers to how consistent a cue is (e.g. a morphological marker), was observed to be the most significant predictor of the participants’ performance. A comparison of Russian and German case marking systems also highlighted the importance of cue reliability in second language learning. Native English speakers, who were learning Russian, mastered
the case system more readily compared to native English speakers who studied German case inflections. This is assumed to reflect the fact that although there are more cases in Russian than in German, the system is more reliable (Kempe & MacWhinney, 1998). More recently, Krajewski, Theakston, Lieven and Tomasello (2011) investigated case switching in young Polish speaking children. In the experiment, two- to three-year-old children were taught novel nouns in one case and required to produce these nouns in a target case that was different from the source form. The authors predicted that children would make fewer case switching errors with high frequency target forms than with low frequency target forms. However, they did not find a frequency effect for the source form or the target form as such. Children produced the same target forms at different rates depending on the source form and the direction of case switching. Similarity between the source form and the target form was found to affect children’s productivity with novel nouns. For example, children were better at producing the nominative target when the source form was the instrumental case in the masculine gender compared to a condition with a genitive target case, while the genitive (target) was produced more proficiently from the feminine form of the instrumental case (source). This effect was attributed to an interaction of similarity between the target and source inflections, and which inflection the items were originally presented in. These findings show that whereas some case endings are considered easier or more challenging than others per se, we should also focus on how children switch from one case to another.
Evidence for children’s sensitivity to lexical item+inflection combinations comes from experiments studying why children sometimes produce a non-finite verb form (the infinitive, e.g. *eat) in a finite context (*He eat) while also providing non-finite forms correctly in other obligatory contexts. Theakston, Lieven & Tomasello (2003) argued that this phenomenon could be accounted for by input frequency patterns of verbs in questions; when children hear third singular pronoun subjects combined with non-finite verbs ("Where does it go?") and with finite verbs ("It goes here"), children may treat third singular marking as optional. Two- and three-year-old English-speaking children were introduced to six familiar and three novel verbs in either a declarative context or a question context or both. In an elicitation paradigm the children had to produce the verbs in finite contexts. The mean proportional use of third singular marking was calculated for each condition. The authors found differences in children’s productions between familiar and novel verbs. Children produced finite forms (third singular marking) and non-finite, unmarked forms equally across the familiar verbs, independent of which context the children heard them in the experiment. Input effects of the experimental condition were reported with novel verbs, as predicted by the authors. Children produced finite verb forms significantly more often with the novel verbs that were only presented in a context with finite marking compared to verbs that were presented in non-finite and mixed contexts. These results demonstrated the input effects on the combination of verb+person marking. Supporting findings regarding lexical item+inflection frequencies have been reported also by Finneran & Leonard (2010). The authors examined typically developing English-speaking children’s acquisition of the 3SG (third singular) present tense verbal -s marker. Sixteen
children aged 30-36 months learnt novel verbs in either third singular contexts (‘The tiger heens’) or contexts with unmarked (non-finite) verbs (‘Will the tiger heen?’). In the experiment tasks, these verbs were elicited in both finite and non-finite contexts to examine the optional infinitive error rates. Again, the results demonstrated the importance of linguistic input on children’s accuracy with morphology and language in general. The verbs that were presented only in a 3SG context were subsequently more likely to be used with a third singular inflection in finite and non-finite contexts in the experimental task.

These kinds of findings have been applied in modelling to demonstrate the extensiveness of input effects on child language to account for patterns in acquisition. An example of such a computational model is The Model of Syntax Acquisition in Children (MOSAIC) (Freudenthal, Pine, & Gobet, 2006; Freudenthal, Pine, Aguado-Orea, & Gobet, 2007; Freudenthal, Pine, & Gobet, 2010). It is a model of a language learner without pre-existing knowledge of grammar or access to semantic information. The model is used to examine optional infinitive errors, an error of omitting finite verbal markers in obligatory contexts (e.g. *He eat). Primacy and recency effects are introduced into the model’s learning to simulate human learning as closely as possible. In fact, the errors are partly explained in terms of these two memory effects. The authors argue, and have shown, that optional infinitive errors are a result of truncation of compound-finite utterances (e.g. He can eat ice cream). When children hear compound-finite structures in input, they store the beginning and the end of an utterance due to the primacy and recency effects, omitting the auxiliary (e.g. can). When children retrieve these partly stored
utterances, they make optional infinitive errors (*He eat ice cream) as failure of storing all grammatically necessary components of the utterance. The model has been shown to simulate language acquisition patterns fairly accurately cross-linguistically in English, German, Dutch, French and Spanish. Particularly relevant to our study, the model has also demonstrated lexical effects in the acquisition process. Using MOSAIC, the authors found different optional infinitive error rates across different lexical items. This effect persisted across the five languages examined by the model. For example, the verbs ‘sleep’ and ‘sit’ in English were associated with a very high error rate (100%) while the error rates were low for ‘want’ (33%) and ‘hurt’ (25%). The error rates positively correlated with how often a particular lexical item (verb) occurred in compound-finite structures in child-directed speech of the target language (Freudenthal et al., 2010).

Researchers have also explored the idea of children’s errors being a result of defaulting to the form with highest frequency. Räsänen, Ambridge & Pine (2014) investigated Finnish-speaking children’s input-driven verb inflection errors to examine whether they could be accounted for by the most frequent form. Various person and tense inflection combinations were elicited with familiar verbs from two-, to four-year-old children. The findings showed that the children made more errors with low frequency verb forms, particularly the third plural. It was also found that most errors comprised using a high-frequency form (e.g. 3sg) instead of a low-frequency target form, but the errors were affected by lexical item frequencies as well: children performed better with high frequency verbs compared to low frequency items. Additionally, children’s accuracy was affected by phonological
properties of the target. Fewer errors were elicited with low-frequency targets with a high number of phonologically similar other verbs (i.e. phonological neighbourhood density) compared to low-frequency targets that had fewer phonologically similar verbs. Mutual phonological properties between a target and other verbs were argued to make the targeted inflectional paradigm more accessible, and thus, improve children’s performance. These findings yet again demonstrate the importance of lexical item frequency, frequencies of individual suffixes and phonology. Overall, the study implies that while it does seem like children default to the most frequent form, third singular, it is still competing against all other possible inflections. Frequencies of individual verb+inflection combinations, together with phonological effects, affect how accurately children inflect verbs in different person/number contexts.

Partly in line with the study by Räsänen et al. (2014), a series of studies on English-speaking children’s noun and verb inflection showed that stems and inflected forms compete against each other in production (Matthews & Theakston, 2006). Children aged between three and nine years were tested on their knowledge of noun plural inflections and verbal past tense forms. It was found that in an elicitation task of noun plurals with regular (e.g. ‘dog’), irregular (e.g. ‘man’) and sibilant (a stem ending with an s, e.g. ‘dress’) types, children between the ages of four and eight showed different frequency effects. The eight-year-olds were more proficient with high frequency nouns than with low frequency nouns, while five-, six-, and seven-year-olds showed an interaction effect between noun frequency and phonology. When the authors administered an equivalent elicitation experiment of
past tense verb forms, they found that seven-year-olds showed a lexically specific frequency effect but only with regular verbs ending in either /-d/ or /-t/. An additional effect of relative frequency was found that demonstrated the importance of competition between different inflected forms. The participants completed a task that entailed inflecting irregular plural nouns from two groups: nouns that were more frequent in singular than in plural (man, mouse, knife) and nouns that were more frequent in plural than in singular, or approximately equally frequent in both forms (child, tooth, foot). Five-year-olds were more likely to inflect those nouns that were at least as frequent in plural as in singular compared to the items that were relatively more frequent in the singular form. The authors also found varied token frequency effects: five-, six-, and seven-year-olds inflected sibilant nouns more often when they had a high token frequency; five-year-olds were better at inflecting high frequency regular nouns than those of low frequency and a similar effect was found in eight-year-olds’ performance with irregular nouns. Finally, children showed weakest performance with sibilant nouns. The results suggested an interaction of stem phonology, noun frequency and regularity. As mentioned earlier, inflected forms with high frequency are more likely to be used correctly in obligatory contexts compared to forms with low(er) frequencies, and even prevent errors. High frequency nouns often appear in one inflected form or in a small group of inflections, which promotes fairly high accuracy with of a limited set of inflected forms of that lexical item. On the other hand, high frequency items can also cause error. A child’s accuracy with high frequency lexical items is usually restricted to those inflections the lexical item most frequently appears with. Errors can occur as a result of highly frequent lemma + suffix combinations when the child
attempts to produce a low frequency inflection with a high frequency lexical item. These kinds of observations are reported by Aguado-Orea & Pine (2015), who show that Spanish children are accurate with 3rd person singular verbal marking in obligatory contexts due this form being the most frequent one in input. These Spanish-speakers also incorrectly extended the 3rd singular form to 3rd plural contexts as the former was relatively considerably more frequent than 3rd plural verbal forms. Thus, it is important to note that while high frequency forms promote accurate inflectional morphology use in obligatory contexts, children are more likely to use these forms in incorrect, low frequency contexts, too. Another possibility is children may find it easier to segment high frequency nouns that occur with a variety of cases from those suffixes compared to low frequency nouns. This could promote accurate use of these nouns, which in turn might encourage correct use of low frequency inflections with these nouns.

Finally, one of the biggest issues in older research into children’s grammatical errors is that the focus has been on the overall error rate of a child’s speech sample, which is often found to be below 5% (Hoekstra & Hyams, 1998; Wexler, 1998). The problem is that an overall error rate is not very informative of a child’s errors as it misses the variety of error rates between different inflectional contexts. Children can, and do, show different error rates in different contexts (Aguado-orea & Pine, 2015; Räsänen et al., 2015), a phenomenon disguised by a low overall error rate. Another issue in investigating only an overall error rate is that it does not distinguish lexical items by frequency. Maratsos (2000) points out
that items with high token frequency bias the overall error rate by, again, disguising relatively high error rates on low frequency items.

1.1. The aim and the research questions of the study

The aim of the present study is to examine the possible relationships between different types of input frequencies and the accuracy of a young child’s case marking. We asked the following research questions using the measures as defined below.

1) Does the absolute input frequency of individual nouns correlate with the child’s accuracy in using those same nouns? (Input frequency = Number of tokens per noun type (combining all case endings); % Child’s accuracy per noun type = Total number of tokens per noun type (combining all case endings) produced correctly / Total number of tokens per noun type (correct + errors)).

2) Does the absolute input token frequency of specific case endings correlate with the child’s accuracy in using those same case endings in obligatory contexts? (Input frequency = Number of tokens per case ending (combined across all nouns); % Child’s accuracy per case ending = In obligatory contexts for a given case ending, Total number of tokens (combined across all nouns) produced correctly / Total number of tokens (correct + errors)).

3) Does the absolute input type frequency of specific case endings correlate with the child’s accuracy in using those same case endings in obligatory contexts? (Input frequency = Total number of different nouns used with a specific case ending, from the selected input sample; % Child’s accuracy per case ending = Same as in 2).
4) Does the absolute token input frequency of specific case endings from the entire corpus correlate with the child’s absolute frequency of incorrect case uses, per case ending (e.g. use of the nominative in non-nominative contexts)? *(Input frequency = Same as in measure 2 but from the whole corpus; Child’s frequency of incorrect case uses = Absolute number of times a case ending is used in an incorrect context instead of the target case ending)*

The input frequency used in this measure differs from the previous three measures by including token frequencies of specific case endings from the entire input corpus instead of restricting the frequencies only to those nouns the child made errors with. We wanted to explore the relationship between these unrestricted case frequencies and the child’s incorrect cases usage, as by doing this certain skews from the input sample would be eliminated. When the case frequencies of only those nouns the child made errors with are examined, we are getting a fairly limited picture of the child’s case input. The fact that the child did make errors with these nouns suggests there is something about them that makes them more error prone in the child’s speech. The input frequencies were extracted from the entire input corpus also for measure 5.

5) Does the absolute type frequency of specific case endings from the entire corpus correlate with the child’s absolute frequency of incorrect case uses, per case ending (e.g. use of the nominative in non-nominative contexts)? *(Input frequency = Total number of different nouns used with a specific case ending, from the whole corpus; Child’s frequency of incorrect case uses = Same as in measure 4)*
6) Does the absolute input frequency of specific noun+case combinations correlate with the child’s accuracy in using those same noun+case combinations? (Input frequency = Number of tokens of a given noun in a given case; % Child’s accuracy = \( \frac{\text{Number of tokens of noun+case combination produced correctly}}{\text{Total number of noun+case tokens (correct + errors)}} \)).

7) Does the relative input frequency of particular nouns in particular cases, relative to the overall use of those nouns, correlate with the child’s accuracy in using those same noun+case combinations? (Input frequency = Number of tokens of a given noun+case combination / total number of tokens of a given noun (combining all case endings); % Child’s accuracy = as in measure 6)

Measure 1 examines the effects of the variety of different forms of the same noun on the child’s accuracy with noun case marking. We predicted the child would be more accurate (i.e. make fewer errors) with higher frequency nouns in the input than with nouns of lower frequency. Measure 2 tests the relationship between the child’s accuracy with each case and overall token frequency of different cases in the input. It was predicted that the child would be more accurate with cases with high overall token frequencies in the input compared to cases with lower input token frequencies. Measure 3 is otherwise identical to measure 2 except that the independent variable input is type frequency instead of token frequency. Again, case endings with high type frequency (which appeared combined with large numbers of different nouns) were predicted to be associated with higher accuracy.
compared to cases with low type frequencies (appearing with fewer different nouns).

Measures 4 and 5 examine the relationship between the token and (noun) type frequencies of the different case endings in the input and the likelihood that the child would overextend use of these case endings into incorrect contexts. The input frequencies for both measures were extracted from the entire input corpus. We predicted that the child would be more accurate with cases with high overall type and token frequencies in the input compared to cases with lower input type frequencies. As high type frequency, i.e. a particular case ending occurs with a high number of different nouns, is known to promote generalisation, we would expect a correlation between type frequency and overgeneralisation would be particularly high.

Measures 6 and 7 both examine the effects of the frequency of noun+case combinations in the input on the child’s accuracy with these combinations. Measure 6 used the absolute frequencies of these combinations whereas the frequency of use of particular case endings with a given noun, relative to the use of that noun with all other available endings formed the basis for measure 7. Including these two measures allowed us to find out whether it is the absolute frequency of individual noun+case combinations, or the relative competition between cases per noun that the child is sensitive to. If absolute frequency is more crucial than relative frequency, the child would be expected to perform equally well with two (or more) noun + case combinations that have similar absolute frequencies but different relative frequencies. On the other hand, if relative frequency is more important
than absolute frequency, the child would be predicted to perform more accurately with noun + case combinations that are relatively more frequent compared to competing strings in input. Most likely both kinds of frequencies interact, contributing to highest accuracy with case endings with high absolute token frequency, which at the same time make up the highest relative frequency of total uses of a particular noun. These combination variables also rule out a potential confounding variable: some cases are more likely to be combined with particular noun types than with some others due to semantics. For example, food and drink related nouns are often inflected in the partitive, a direct object case that denotes for ‘some’ or ‘incomplete’ whereas venues, buildings and locations are often combined with locative cases such as the inessive case, which is equivalent to the English preposition ‘in’.

2. Methods

2.1. Participants and data coding

The participants were the same as in study 1: one mother-child dyad. Case marking errors were coded simultaneously with morphological coding of all nouns in the corpus. To identify the presence of errors, the case context for each noun use was established with reference to the wider linguistic context. Uses of the target case as defined by the context were coded as correct (e.g. NOM form in NOM context); uses of a case ending that was not compatible with the case context were coded as errors (e.g. GEN form in NOM context). The code line was created by the first author as there was no usable morphological glosser developed for Finnish at the time of the study. The tier consisted of the noun lemma, number, and case. If there
was a case marking error, the correct target form was also included in the dependent tier. The corpus also contains some ambiguous and incorrect noun forms that do not have an identifiable target form. These forms were excluded from the analysis, as identifiable target forms were necessary.

2.2. Procedure

First, all of the child’s errors were extracted from the corpus to identify the noun types with which errors were made. Errors that were repetitions were excluded. All uses of these nouns (correct and errors) were then extracted from both the child’s speech and the input data. The frequencies of the different case contexts with each noun were calculated, along with the number of correct uses and errors in each context. Data were then combined across case endings to determine the overall frequency of use of individual nouns (measure 1), and across case endings to determine the overall token frequency (measure 2) and the overall type frequency (measure 3) of individual case endings. Measures 2 and 3 were repeated with one difference only: while the input frequencies were first extracted from the selected input sample, i.e. from the sample of nouns the child made errors with, for measures 2 and 3, the input frequencies were extracted from the entire input corpus for measures 4 and 5. The child’s overall relative correct use of individual nouns and individual cases was then calculated (measure 6), as well as her relative correct use of particular noun+case combinations in obligatory contexts (measure 7).
3. Results

*Descriptive statistics:* A calculation of the overall error rate for the child corpus (= token frequency of all of the child’s case marking errors / number of noun tokens in the child corpus before controls, multiplied by 100) showed the error rate to be 4.10%. While the overall error rate is low, the error rates across different inflectional contexts are varied (Table 3.1.). Error rates for each case context demonstrate that the nominative has the lowest error rate (0.40%) while the rate for the genitive case is as high as 56.25%. Rates for each case use in incorrect context instead of the target case also show considerable variability. The elative case has the smallest rate of being provided in incorrect contexts (0.03%) and the ablative has the highest rate of 15.79%. *(Correction no. 6)*

**Table 3.1.** Error rates for each case and rates of different cases used in incorrect contexts (%).

<table>
<thead>
<tr>
<th>Case</th>
<th>Error rates for each case context (%)</th>
<th>Rates for each case used in incorrect context (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOM</td>
<td>0.40</td>
<td>4.98</td>
</tr>
<tr>
<td>PART</td>
<td>5.10</td>
<td>1.22</td>
</tr>
<tr>
<td>ACC</td>
<td>19.77</td>
<td>2.33</td>
</tr>
<tr>
<td>GEN</td>
<td>56.25</td>
<td>3.75</td>
</tr>
<tr>
<td>ADE</td>
<td>6.87</td>
<td>3.05</td>
</tr>
<tr>
<td>ALLA</td>
<td>7.97</td>
<td>1.45</td>
</tr>
<tr>
<td>ABLA</td>
<td>31.58</td>
<td>15.79</td>
</tr>
</tbody>
</table>
Correlation analyses: Sixty-five noun types and ten cases were included in the analysis. Altogether 4,590 noun tokens were included: 2,997 tokens from the mother’s sample and 1,593 tokens from the child sample. Spearmans’ rho correlations were carried out to determine the relationship between the seven input frequency variables and the child’s accuracy of use of particular nouns, case markers, and their combinations.

The first analysis (measure 1) showed a positive, significant correlation between the overall input frequencies for individual nouns and the child’s accuracy of use of those nouns ($r_s(65) = .717, p = .001$, 2-tailed) (see Figure 3.1.).
Measures 2 and 3 were conducted with only those nouns the child made errors with. The second analysis (measure 2) revealed a non-significant relation between the input token frequency of particular case endings and the child’s relative accuracy of use of those endings ($r_s (10) = .365, p = .300$, 2-tailed) (see Figure 3.2.). The relation between the input type frequency of case endings and the child’s relative accuracy (measure 3) was also non-significant ($r_s (10) = .171, p = .636$, 2-tailed) (see Figure 3.3.).

Figure 3.1. Absolute noun input frequencies correlated with the child’s accuracy per noun type (measure 1)
Figure 3.2. Absolute case input token frequencies correlated with the child’s accuracy per case (measure 2).

![Graph showing the correlation between child's accuracy and mother's absolute token frequency per case.](image)

Figure 3.3. Absolute input type frequencies correlated with the child’s accuracy per case (measure 3).

![Graph showing the correlation between child's accuracy and mother's absolute type frequency per case.](image)
Following these two type and token frequency correlations, measure 4 examined the relation between absolute input token frequencies for each case ending and the likelihood of extending use into incorrect contexts (see Table 3.2). While the input frequencies were extracted from a restricted noun sample for measures 2 and 3, the input frequencies for measures 4 was extracted from the whole corpus. Spearman’s rho revealed a highly significant correlation between the token frequencies and incorrect case use ($r_s (10) = .740, p = .014, 2$-tailed). Measure 5 (see Table 3.1.) correlated input type frequency with the child’s incorrect case use, which showed a marginally significant relation ($r_s (10) = .624, p = .054, 2$-tailed). Equivalently to measure 4, the input type frequencies were extracted from the entire input corpus for measure 5.

**Table 3.2.** Input type and token frequencies correlated with the child’s number of incorrect case uses (measures 4 and 5).

<table>
<thead>
<tr>
<th>Case</th>
<th>MOT absolute token frequencies</th>
<th>MOT absolute type frequencies</th>
<th>CHI frequency of incorrect case uses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominative</td>
<td>3345</td>
<td>720</td>
<td>111</td>
</tr>
<tr>
<td>Partitive</td>
<td>1496</td>
<td>520</td>
<td>6</td>
</tr>
<tr>
<td>Accusative</td>
<td>921</td>
<td>255</td>
<td>2</td>
</tr>
<tr>
<td>Genitive</td>
<td>488</td>
<td>139</td>
<td>3</td>
</tr>
<tr>
<td>Adessive</td>
<td>449</td>
<td>177</td>
<td>4</td>
</tr>
<tr>
<td>Allative</td>
<td>416</td>
<td>113</td>
<td>2</td>
</tr>
<tr>
<td>Ablative</td>
<td>44</td>
<td>27</td>
<td>3</td>
</tr>
<tr>
<td>Inessive</td>
<td>351</td>
<td>152</td>
<td>3</td>
</tr>
<tr>
<td>Illative</td>
<td>556</td>
<td>163</td>
<td>1</td>
</tr>
<tr>
<td>Elative</td>
<td>165</td>
<td>91</td>
<td>1</td>
</tr>
</tbody>
</table>

Ten separate correlations were run, one for each case ending, to determine whether the absolute frequency of particular noun+case endings in the input
correlated with the child’s accuracy of use of those same noun+case endings (measure 6; see table 3.2). Eight of the correlations were non-significant whereas two were significant: the genitive ($r_s (63) = .929$, $p = .001$, 2-tailed) and the illative ($r_s (11) = .721$, $p = .01$, 2-tailed). A similar analysis was conducted to determine whether the relative proportional frequency of individual noun+case combinations as a function of that noun’s overall frequency of use in the input correlated with the child’s accuracy of use of those noun+case combinations (measure 7; see table 3.3.). This time only the allative was significant ($r_s (16) = .507$, $p = .01$, 2-tailed), while the nine other correlations were non-significant.

**Table 3.3.** Absolute input frequencies of noun+case combinations correlated with the child’s accuracy with these combinations (measure 6). Relative (%) input frequencies of noun+case combinations correlated with the child’s accuracy with these combinations (measure 7).

<table>
<thead>
<tr>
<th>Case</th>
<th>N (type)</th>
<th>Correlation (measure 6)</th>
<th>Correlation (measure 7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominative</td>
<td>52</td>
<td>-.030</td>
<td>-.174</td>
</tr>
<tr>
<td>Partitive</td>
<td>38</td>
<td>.145</td>
<td>-.276</td>
</tr>
<tr>
<td>Accusative</td>
<td>21</td>
<td>.129</td>
<td>.140</td>
</tr>
<tr>
<td>Genitive</td>
<td>63</td>
<td>.929**</td>
<td>.177</td>
</tr>
<tr>
<td>Adessive</td>
<td>21</td>
<td>-.364</td>
<td>-.245</td>
</tr>
<tr>
<td>Allative</td>
<td>16</td>
<td>-.132</td>
<td>.507*</td>
</tr>
<tr>
<td>Ablative</td>
<td>10</td>
<td>.282</td>
<td>.238</td>
</tr>
<tr>
<td>Inessive</td>
<td>10</td>
<td>.039</td>
<td>-.068</td>
</tr>
<tr>
<td>Illative</td>
<td>11</td>
<td>.712*</td>
<td>.311</td>
</tr>
<tr>
<td>Elative</td>
<td>8</td>
<td>.260</td>
<td>.020</td>
</tr>
</tbody>
</table>

It is possible that neither absolute nor relative frequencies had a systematic relationship with the child’s output because there was not enough competition or variation between cases per noun type. We investigated this by excluding nouns
that the mother used in one case only for at least 70% of the time and re-ran the correlations. Again, noun input frequency was significantly correlated with the child’s accuracy with case marking per noun type, \( r_s (35) = .552, \ p = .001, \) 2-tailed) (measure 1). When we examined whether the absolute frequency of particular noun+case endings in the input correlated with the child’s accuracy of use of those noun+case endings with this subset of noun types, of the ten case endings only the illative was significant \( r_s (9) = .773, \ p = .01, \) 2-tailed). Comparable correlations based on the relative input frequency of particular noun+case combinations showed that this time the adessive \( r_s (21) = .912, \ p = .001, \) 2-tailed) and the illative \( r_s (9) = .773, \ p = .01, \) 2-tailed) were significant while other correlations were not. This was conducted to examine whether a particular noun was used almost exclusively with one case.

4. Discussion

The aim of the study was to investigate what kinds of input frequencies are linked to a young child’s accuracy with case marking. We examined the effects of absolute token and type frequencies of nouns and cases, absolute noun+case combination frequencies and relative (%) noun+case frequencies. The results supported previous findings on lexical specificity by showing relative lexical item input frequency correlated with the child’s inflection accuracy with each noun (Freudenthal, Pine & Gobet, 2010; Theakston, Lieven & Tomasello, 2003). The other two statistically significant variables were input type and token frequencies on the child’s number of incorrect case uses. This suggests that a case ending with high input token and type frequencies does not necessarily translate into correct, productive use of that
ending in the early stages of acquisition. High input frequencies may promote the probability of a suffix to be generalised into incorrect contexts when children are unsure of the appropriate target form. Similar findings have been reported previously. High input frequency has been found to cause errors in naturalistic (Aguado-Orea & Pine, 2015; Cameron-Faulkner & Kidd, 2007) and experimental studies (Dąbrowska, 2008; Rowland & Theakston, 2009). There were two more correlations, the absolute noun+case combination frequencies of the genitive and the illative, that were also significant. When we removed nouns that occurred in one case for at least 70% of the time, the overall trend of the findings remained the same as with all of the nouns the child made errors with. This shows that the results do not merely reflect the child learning to use each noun with a single high frequency inflection. However, overall, no relation was found between the number of tokens of a given noun in a given case in input and the child’s correct use of the same noun+case combinations in obligatory contexts. The correlations were non-significant for both absolute and relative input frequencies in these measures. The case input frequency variable was also non-significant.

As mentioned earlier, Krajewski et al. (2011) examined Polish children’s case switching and found the direction of case switching to affect children’s performance. The authors suggest that there may be conceptual differences between case endings which could make some of them more error prone than others. For example, the Finnish partitive, which denotes incomplete direct objects and actions, is the second most frequent case yet the child’s accuracy with it (a relatively high 83% correct overall) was not related to input frequency. It may be
an example of a conceptually challenging case as ‘incompleteness’ is not always a visibly salient attribute. The illative case, which is equivalent to the English preposition ‘into’ (e.g. *into the house*), could be an example of a conceptually easier case. It refers to concrete locations systematically which makes it a conceptually less ambiguous case than e.g. the partitive. It has been found that conceptual and grammatical properties do interact in noun and verb learning (Bassano, 2000). However, the child’s overall accuracy with the illative case was 79%, which is fairly similar to that of the partitive, showing the child was fairly accurate with both cases overall despite their potential conceptual differences. The partitive is a case of frequency, with a token frequency of 1,496 and appearing with 520 different nouns in the entire input corpus. The illative has considerably lower, but not low, token (556) and type (163) frequencies in the entire input corpus. One possibility is that both cases have passed a frequency threshold by having sufficiently high token and type frequencies that supported the child’s accuracy with these cases well. Alternatively, the child is accurate with a (potentially) conceptually challenging case like the partitive due its high input frequencies, but at the same time, a relatively high accuracy can be achieved with a conceptually easy case, the illative, with lower token and type frequencies.

Matthews, Lieven, Theakston & Tomasello (2009) found an unexpected frequency effect in a study investigating 4-6-year-olds’ co-referencing errors with sentences like “Mama Bear is washing her”. Children are inclined to comprehend that ‘Mama Bear’ and ‘her’ co-refer to the same entity even though the sentence lacks a reflexive pronoun. In the experiment, children made truth value judgements...
of such sentences with verbs of high frequency and low frequency to examine verb frequency effects on the children’s performance. While the usage-based approach predicts that children should have made more accurate judgements with high frequency verbs, the results showed a contrary effect. In fact, children were significantly better at making truth value judgements with low frequency verbs compared to verbs with high frequency. This was explained by the so called kindergarten path effect – an effect which is caused by children making quick judgements of more familiar verbs before carefully listening to the end of these sentences, leading to a higher error rate with them as opposed to less familiar verbs. The finding showed that even if there is a vast body of evidence for children being more proficient with lexical items with high frequencies, including our results, frequency effects interact with other aspects of language processing and general cognitive development. The authors reached a similar conclusion when they examined frequency effects of the different pronominal accusative forms. While the third plural accusative form, them, is the most frequent case marked pronoun in English child-directed speech out of the forms included in this study, children were least accurate with it in the experiment. Matthews and colleagues speculated this might be related to other aspects of cognitive development like understanding number instead of being exclusively linked to language acquisition and frequency effects. This reasoning is in line with Krajewski et al.’s (2011) suggestion that some cases of noun inflection may be conceptually more challenging, which may interfere with pure frequency effects.
Based on the possibility that children’s productions are affected by similarity between a recently heard case and the child’s target case, it is possible that a case in the child’s input activates a phonologically or semantically similar case ending in the child’s phonological and/or syntactic network, like a priming effect. Studies on morphological priming have shown that inflections, especially regular ones, are potent primers for adults and older children. Clahsen & Fleischauer (2013) examined 8- and 9-year-old German children’s morphological processing with a multi-modal priming task. The aim of the experiment was to study whether there are differences between priming regular and irregular participles in verb inflection. They found that children of both ages and adults were readily primed by the regular inflection whereas only 9-year-olds and adults were additionally primed by the irregular participle. It is also known that magnitude of a priming effect is related to the degree of overlap in phonology between the prime and its target (Gonnerman, Seidenberg & Andersen, 2007). Such findings suggest that phonologically close case endings, e.g. –lla and –lle or –sta and –ssa in Finnish, could prime each other and make them more accessible for the child to produce in a naturalistic setting. While phonological effects were not examined in our study, it is likely that that phonological similarity of inflections does interact with other frequency effects.

There is some evidence suggesting that phonology and morphology do interact with each other. The relationship between the phonology of stem endings and suffixes has been studied in both typically and atypically developing children. For example, Polite (2011) examined phonological factors in morphology production in three-year-old typically developing English-speaking children as a
control group for children with specific language impairment. He analysed spontaneous speech data of typically developing children recorded during standardised test battery. These children were included for their inconsistent use of plural marking. Polite investigated whether the length of the noun stem and stem-final phoneme type affect children’s provision of the plural marker. The former was not found to have an effect whereas stem phonology was significant. Three stem types were identified: vowel, consonant and consonant clusters. It was found that stems ending with a vowel were the easiest, followed by consonant stems. Consonant clusters were the most challenging stem group, suggesting that the phonetic context of inflection affects productivity. There is evidence also from Spanish showing that the phonetic context of stems has an effect on children’s productivity with suffixes. When three-year-old Spanish speakers are required to produce plural forms of real words and pseudo-words, they were more proficient with stems ending in a vowel than in a consonant (Arias-Trejo, Abreu-Mendoza & Aguado-Servín, 2014). Furthermore, it has been previously shown that children are not equally productive with all consonant stems. In an elicitation experiment utilising Berko’s wug test method, Solomon (1972) found that even children as old as six- and seven are more proficient at inflecting stems ending with /p/ and /k/ than with /f/.

Previous studies have also provided evidence for children being more likely to omit inflections when the stem sounds like it is already inflected. For example, Köpcke (2002) re-analysed data on English-speaking children’s plural elicitations and also examined German-speaking children. More specifically, he analysed the
effect of allomorphs and near-allomorphs of the English and German plurals on plural production. It was found six- and seven-year-old English-speaking children and two-year-old German-speaking children did not provide the plural marker as frequently with noun stems ending with phonemes similar to the relevant plural suffix compared to noun stems that did not end with a plural allomorph. While research is lacking in the area of possible impact of allomorphy in child Finnish, a study of adult language processing suggests that low levels of allomorphy in Finnish improve processing of derivational morphology due to high salience of affixes (Järvikivi, Bertram, & Niemi, 2006).

Thus evidence from morphophonological experiments suggests that stem phonology is a critical factor in children’s productivity with noun inflection. Finnish morphophonology is extremely complex and would require very detailed analyses and a whole separate study. This is an important issue for future research and should be included in studies on inflectional morphology acquisition. Perhaps there is an interaction effect between the phonology of individual suffixes and stems which would make some noun-case combinations more challenging than others, e.g. combinations that create a consonant cluster by attaching a consonant-initial suffix to a consonant stem.

An issue with the current study that must be addressed is that while most of the correlations of three variables, relative case frequency, relative noun+case combination frequencies and relative noun+case combination frequencies, were non-significant, there were a few correlations that were statistically significant. For example, the relative frequencies of the noun+genitive and the noun+illative
combinations correlated significantly with the child’s accuracy. These two significant results could be a product of a relatively small sample size which has been shown to skew naturalistic findings. Rowland & Fletcher (2006) examined whether a small naturalistic sample size affects the observed lexical effects. They reported that error rates with infrequent items and lexical specificity were equally likely to be under- and overestimated in small samples, suggesting caution to be practiced when drawing conclusions on children’s errors based on small data sets.

Even though our original corpus, from which the child errors were extracted, is not small, our error data of the 65 noun types is not extensive either. A larger sample of errors might have allowed us to observe other trends in the data in terms of relations between the frequencies of particular noun+case combinations in the input and their accuracy of use in the child’s speech (in addition to the significant relations observed in the current data for the genitive, illative, and allative cases).

A feature of the Finnish language to consider is transparency. There is no gender to interact with the case marking system and overlap between suffixes is minimal: only the singular forms of the genitive and the accusative overlap as well as do the plural nominative and the plural accusative. This is different from children acquiring languages like German and Polish as their task is to not only learn the meanings of case and number inflections but also to figure out the role of grammatical gender. Experimental evidence is suggesting that there are cognitive benefits to a more transparent inflection system. When adult participants and 9- and 10-year-olds are required to learn an artificial language with two subclasses of nouns, each of which has an identifiable common marker, most participants learn
to distinguish between the two noun groups. No participant learned the difference between the two classes in a control condition with ambiguously marked nouns (Brookes, Braine, Catalano & Brody, 1993). Even though this study demonstrated the effect of transparency in adult learners and older children, it showed that transparency is a critical component to how easily a cognitive system like language can be learned.

In summary this study has found that absolute noun token frequency is an important factor in Finnish children’s accuracy with case-marking and that when they do make errors they default to the most frequent form, nominative. This illustrated that hearing a particular noun or case does not necessarily mean the child is accurate with them, but merely that it is more likely for the child to overextend lexical items and inflected forms of high frequency. Type frequency was also found to have a marginal impact on how likely the child extended case usage to incorrect contexts. These results illustrate that while type and token frequency have a vital role in how accurate children are with inflection, frequency measures alone cannot account for error patterns. According to other research, factors also affecting children’s case marking are stem phonology, transparency, possibly priming and non-linguistic, cognitive development. The field should also start exploring types of frequencies that have not been studied yet, such as transition frequencies between cases. Previous research suggests that morphophonology could be a significant contributor in inflectional morphology acquisition, by for example making the acquisition of certain stem + suffix combinations easier than
others. It is an issue that needs to be investigated more closely in future research on inflectional morphology acquisition.

References


Cameron-Faulkner, T., & Kidd. E. (2007). I’m are what I’m are: The acquisition of first-person present BE. *Cognitive Linguistics, 18*(1), 1-22.


Rationale for Study 3

Studies 1 and 2 examined one child’s productivity with noun inflection errors. The studies were naturalistic case studies of mother-child interaction. To complement our naturalistic studies, we designed an experiment to measure comprehension instead of production, and to investigate how two grammatical cues work together to support comprehension at three ages. The results from studies 1 and 2 have so far shown that even though a young Finnish child ages 1;7-1;8 may superficially seem highly productive with noun case marking, this is still significantly more restricted than adult productivity. In the following experiment, we attempted to obtain data from as young children as possible as well as investigating children at two later ages. An experiment with three age groups, two-, four-, and six-year-olds, will allow us to have a better idea of the actual developmental trajectory of how two grammatical cues, noun case marking and word order, are interpreted. As in our naturalistic studies, we used adult language as a control by testing a group of adults with the same experimental task to compare children’s performance against.

In a preliminary analysis, we conducted a naturalistic input analysis of case marking and word order to gain a better understanding of the interaction between these two cues in child-directed Finnish, and to make informed predictions. This idea is based on the Cue Competition Model developed by MacWhinney & Bates (1989), who argued that children’s grammatical acquisition reflects the strengths of the different grammatical cues in input (see Chapter 1: General Introduction). The model has been used to explain cross-linguistic variation in grammar acquisition.
Chapter 4: How do Finnish children comprehend word order and case marking cues in transitive sentences with novel verbs?

Abstract

A comprehension experiment utilising a pointing task was conducted to investigate 2-, 4-, and 6-year-old Finnish children’s understanding of word order and case marking in transitive sentences. Word order and case marking were manipulated in six conditions to examine how children weigh these cues and whether this changes developmentally. Adults tested on the same task used case marking to interpret sentences when it was unambiguous. However, they strongly preferred first noun as agent interpretations when case marking was ambiguous. Two-year-olds correctly identified agents and patients in conditions with unambiguous case marking and with first noun as agent. They performed at chance when case marking was ambiguous, and when the first noun was case marked as the patient. Four-year-olds were able to use word order with ambiguous case marking, however, like the two-year-olds, they were at chance in conditions with case marked, patient-initial sentence types. Six-year-olds performed similarly to adults in all but one condition: patient-agent verb (OSV) with unambiguous case marking. The results suggest the youngest children were sensitive to cues in coalition and that children gradually develop the ability to utilise cues in isolation. Even though six-year-olds performed most accurately out of the three child groups, they were still not performing like adults. As suggested by previous research, sentences with a coalition of cues might provide a framework for children to tease apart the function of individual cues.
1. Introduction

One of the challenges a child will tackle during early language development is to acquire accurate knowledge of the syntax of various sentence constructions. The transitive construction is a language structure that communicates who did what to whom, typically with a combination of cues such as word order, case marking, prosody or animacy. Cross-linguistic research on how cues are acquired has shown that the relative weights of grammatical cues vary between languages. For example, an English-speaking child relies heavily on word order while a German child relies on a coalition of word order and case marking (Chan, Meints, Lieven & Tomasello, 2010; Dittmar, Abbot-Smith, Lieven & Tomasello, 2008). Investigation of the inter-relation between these cues is key to understanding how children can begin to develop an awareness of language as a manipulable system. Even though child language researchers have devoted a lot of attention to studying English-learning children, English is an unusual language in its relatively impoverished inflection and rigid syntactic word order in the transitive construction. In this study we attempt to clarify the roles of case marking and word order in sentence comprehension in Finnish-speaking children, and how comprehension of these cues changes developmentally. Finnish is an interesting language for studying the acquisition of grammatical cues due to its relatively transparent case marking system and alternative word orders.

The degree to which speakers are more sensitive to one cue over another is measured by the Cue competition model (MacWhinney et al., 1985). This is an account of, and a method for, predicting the relative importance as well as the
acquisition order of grammatical cues in different languages. One of the most fundamental assumptions of the model is that children will initially rely more on cues that are weighted heavily in the language they are acquiring. The measure of a cue weight, also called cue validity, consists of two subparts: cue availability, which equates to the raw frequency of a cue, and cue reliability, which equates to the frequency with which the cue reliably signals a particular meaning (i.e. is not homophonous with one or more meanings). For example in English, the reliability of a first noun being the agent instead of the patient is high but not 100% accurate due to the possibility of topicalisation. For example, patients take the subject position in the English passive construction: The vase was broken by John. Interpreting the first noun as the agent would in this case be incorrect.

The effect of cue validity on children’s comprehension is often studied by comparing and contrasting how children acquiring languages with differently weighted cues assign agent and patient roles. For example, Bates et al. (1984) studied sentence interpretation in two- to five-year-old English- and Italian-speaking children. The participants enacted 54 grammatical or semi-grammatical transitive sentences with familiar verbs in three word orders with animate and inanimate agents and patients. They found that even the youngest Italian-speaking children were sensitive to animacy as a cue to sentence comprehension, probably due to the ambiguity and irregularity of word order in Italian, whereas the two-year-old English-speaking children relied more on word order than on animacy. The model does not apply to child language only: adult speakers of different languages have also been shown to rely on different cues in comprehension depending on cue
validity. MacWhinney, Bates and Kliegl (1984) studied cue competition in adult speakers of English, German and Italian. The participants were presented with transitive sentences which were variously manipulated for word order, stress, agreement and animacy. As predicted, English-speakers relied most on word order whereas German-speakers weighted agreement and animacy the most with Italian-speakers weighting agreement as the most important cue. By understanding how the various cues are weighted in the adult end-state, and how this differs from children’s performance over development, language acquisition researchers are better able to understand how the distributional characteristics of any given language might influence the developmental trajectory, providing insights into the underlying universal cognitive mechanisms involved in language acquisition.

Some evidence in support of a developmental shift in cue weights can be found in English-speaking children. An English-speaking child will rely on word order to identify the agent of a causal action from as early as 2;3 when asked to point to one of two scenes with reversed agent and patient roles that matches a corresponding audio test sentence (Noble, Rowland & Pine, 2011). In general, word order cues for English speakers are so strong that they are likely to interpret the first noun as agent even with improbable inanimate agent and animate patient pairings (Bates et al., 1984; MacWhinney, Bates & Kliegl, 1984; Slobin & Bever, 1982, although see Chan et al., 2009). However, even young children learning a word order dominant language are sensitive to case marking. Ibbotson, Theakston, Lieven and Tomasello (2011) examined the role of case-marked pronouns in two- and three-year-old children’s comprehension of active and passive constructions.
Children were presented with either one or two grammatical or ungrammatical case marked pronouns in transitive sentence contexts, and were asked to point to a corresponding cartoon picture. The results showed that the younger children successfully comprehended grammatical sentences with two case marked pronouns better than those with only one case-marked form, but failed to use either case marking or word order to comprehend ungrammatical sentences. The older children comprehended grammatical sentences with either one or two case marked pronouns, but showed differences in their comprehension of ungrammatical sentences as a function of the number of available case marked forms. These results suggest that early in development case marking and word order cues are used in coalition to support performance, indicating that even for these English-learning children, word order does not yet over-ride the information conveyed by highly frequent pronominal forms.

This idea of the importance of a coalition of cues was considered by Slobin & Bever (1982) in the context of canonical sentence representations. A canonical sentence schema is taken to be the most frequent instantiation of a transitive sentence with two noun phrases, consisting of the most typical features of complete clauses in a given language. They conducted a cross-linguistic experiment to investigate whether children operate with this schema in comprehension, and especially, whether word order is part of that schema when it is not grammatically necessary. English-, Italian-, Serbo-Croatian- and Turkish-speaking children aged 2;0-4;4 took part. The tested languages vary in terms of word order flexibility and morphology. The participants listened to transitive sentences with animate nouns
and various word orders and acted them out with toys. They found that Turkish children were relatively accurate with grammatical sentences from the youngest age group compared to the children speaking other languages. In the group of 2;0-2;4-year-old Turkish children, seventy-nine percent performed the experimental tasks correctly whereas the scores for other language groups varied between 58-66%. Even though there was no effect of word order on Turkish children’s accuracy with grammatical sentences, word order manipulations had an impact on parsing ungrammatical sentences. In those cases, children’ interpretations reflected these word orders’ input frequencies. The authors consider the cross-linguistic differences to have occurred due to Turkish children having to acquire one transparent patient-marking morpheme, which is cognitively less challenging than, for example, English speakers’ need to learn word order patterns when both the agent and the patient are nouns. The study suggested that children are most sensitive to the most valid syntactic cue, whether it is word order, case marking or something else and that these canonical schemas are central to children’s sentence comprehension in later stages of development by perhaps allowing children to generalise their knowledge of syntax to other sentence types (Slobin & Bever, 1982).

More recently, Dittmar et al. (2008) suggested that children build comprehension of the transitive structure from sentences with a coalition of cues which form the basis for later acquisition of lower frequency sentence frames. They investigated this by examining cue competition in German-speaking children aged 2;7, 4;10 and 7;3. The children listened to transitive sentences with three novel verbs and one familiar verb, and were asked to point to a matching scene from a
choice of two options. The sentence types were manipulated in terms of word order and case marking: the stimuli included prototypical transitive sentences, i.e. the most frequent word order (SVO) and supporting case marking, a word order condition with ambiguous case marking, and a conflict condition with a patient-first word order but with appropriate case marking. The results showed the youngest age group was above chance in the cues-in-coalition condition – which the authors call the prototype condition - but not in the others. Four-year-olds were able to utilise word order but not case marking as a cue whereas only seven-year-olds treated case marking as a more valid cue than word order. These findings suggest that children start parsing from prototypes with multiple cues in coalition and are only able to use cues separately in sentence comprehension much later.

These findings are consistent with earlier work suggesting that the precise properties of the input that children are sensitive to change over the course of development. Sokolov (1988) examined how accurately cue validity predicted sentence interpretation in Hebrew-speaking children. He found that cue availabilities of word order, a direct object case marker (the accusative), reflexive noun phrases (equivalent to the English –self constructions) and subject-verb agreement were considerably more important for four-, five-, and seven-year-olds compared to reliability whereas reliability was the stronger cue for nine-year-olds and adults. The developmental differences arise from sensitivity to the nominal accusative marker, which was a surprisingly challenging cue to the three youngest groups, possibly due to its similarity to a verbal inflection. Also a low frequency reflexive noun phrase cue was not as readily utilised by the younger ages compared
to nine-year-olds and adults. This suggests that younger children are more sensitive to raw frequency (e.g. of word order patterns) but that as the linguistic system develops as a function of accruing distributional evidence, how reliable a cue is (e.g. a case marker) starts to matter more.

The work of Dittmar et al. (2008) was continued by Chan et al. (2009) in a study of how Cantonese-, German-, and English-speaking children weigh grammatical cues in active transitive sentences with novel verbs. Whereas Dittmar et al. (2008) contrasted word order and case marking, Chan et al. (2009) manipulated word order and animacy by presenting sentences with animate agents and inanimate patients (AI), animate agents and patients only (AA) and inanimate agents and patients (IA). All the sentences were in noun-verb-noun word order. This allowed the authors to compare the extent to which children weigh word order and animacy as individual cues, in contrast to when they are used in coalition in a sentence and, finally, when they conflict (inanimate noun first - IA). Cross-linguistic differences in the results supported the idea that the regularities of a child’s first language strongly affect syntactic acquisition patterns – in this case, how informative word order is in assigning agent and patient roles. For example, when word order was the only cue, both English-speaking and German-speaking children chose the first noun as agent significantly above chance at the age of 2;6. Cantonese-speakers performed above chance later than their English-, and German-speaking peers at 3;6. All language groups chose first noun as agent equally often at 4;6. When animacy was conflicted with word order with an inanimate first noun and animate second noun, all groups were at chance at 2;6,
but three-, and four-year-olds preferred word order over animacy in all languages. Finally, the cues-in-coalition condition (or the prototype condition), was the easiest sentence type for children of all ages to interpret. When the first noun is animate and the second noun is inanimate, all groups are significantly above chance from 2;6. Overall, these results are in line with the study by Dittmar et al. (2008) on German children’s comprehension of case marking and word order: sentence types with a coalition of cues are comprehended first without an understanding of the contribution of individual cues. Children acquire a more sophisticated appreciation of isolated cues later in development as a result of more input that enables computation of accurate cue validities.

So far, we have mainly considered work based on languages with relatively fixed word orders. Children acquiring languages with even greater reliance on case marking also need to learn how to weigh up the cues present in their language. Early work demonstrated that from the age of 2;3-2;6, children acquiring primarily case marking languages (e.g. Hungarian, Turkish) are sensitive to inflections denoting agents and patients with familiar verbs in coalition with other cues such as verb morphology, stress and animacy (MacWhinney, Pléh & Bates 1985; Slobin & Bever, 1982; Sokolov, 1988). More recently, in similar vein to the aforementioned studies investigating cue competition, a study of Turkish children’s verb learning has shown that Turkish-speaking children utilise both syntactic and morphological cues to assign meaning (Göksun, Küntay & Naigles, 2008). Two-, three-, four-, and five-year-old children and a control group of adults enacted transitive and intransitive sentences manipulated in terms of the number of arguments (either
one or two noun phrases), presence of the accusative case marker and presence of a causative verb marker. The enactments were rated on the level of causativity. The authors found that enactments were more causative when sentences contained two noun phrases than one, and with the accusative case marker, providing further evidence that children rely on a combination of cues when initially assigning causativity. Three-, four-, and five-year-olds demonstrated consistent effects of frame (number of arguments), but two-year-olds showed only an effect for intransitive verbs in the frame with two arguments. The authors admit that two-year-olds might have been biased to act out intransitive sentences causatively for the reason that the participants were asked to enact more transitive than intransitive verbs, and by default, the youngest age group could have been affected by the bias in design the most. Additionally, or alternatively, the results demonstrated that two-year-olds are yet to fully form the frame-meaning mappings of either causative and non-causative events or one- and two-argument frames or both.

Studying a language like Turkish has shed some light on cue comprehension in child and adult speakers of a morphologically rich language. Turkish resembles Finnish grammar in many ways with case marking, alternative word orders and the lack of determiners and grammatical gender. Finnish is a semi-null subject language, which means that some pronouns as subjects can be dropped (first and second singular and plural) while some cannot (third plural and singular). Omitting nouns as agents, and pronouns and nouns as patients, is not permitted. From the perspective of noun case marking, studying Finnish adds another layer of
complexity with 14-15 noun cases (depending on how they are counted) while Turkish has six and German four. A larger number of cases implies more competition between the available nominal forms, although it is unclear how this might impact on children’s ability to identify and utilise the various cues to sentence meaning. Generally Finnish case marking is quite transparent as there is no case and gender interaction, most case endings are unique to one particular case, and case marking is always present. However, the singular form of the accusative overlaps with the singular form of the genitive and the plural accusative form overlaps with the plural nominative form with all nouns leading to occasional ambiguity, for instance if agents and patients of a transitive sentence are in plural nominative and plural accusative, respectively. Word order, however, is relatively free in Finnish transitive sentences as case marking denotes agents and patients. Even though the most frequent word order in adult directed speech is SVO, it may be scrambled to highlight the object (e.g. A book-ACC I’m reading) (Hakulinen & Karlsson, 1977) or to distinguish between given (object) and new information (subject) (Kaiser & Trueswell, 2004). Because these alternative word orders are allowed for pragmatic functions in Finnish, investigating the interaction of case marking and word order cues in Finnish-speaking children provides insight into how children approach the problem of teasing cues apart.

As in many case marking languages, the subject of a transitive sentence, the agent, is in the nominative case which is not inflected. The nominative form is considered the basic form of a noun, the same form as found in dictionaries. The direct object i.e. the patient, is either in the partitive or the accusative case
depending on the verb and telicity of the action. Direct objects of unfinished or 
continuous actions are marked with the partitive (e.g. *I’m reading a book-PART*) 
whereas direct objects of completed actions are marked by the accusative case (e.g. 
*I read a book-ACC*). However, even though the equivalent English examples are in 
the progressive present tense to demonstrate the partitive and in the past tense to 
demonstrate the accusative, both case markers can be used in the same tenses too. 
For example, case marking can indicate a potential action outcome: *I’m going to 
read a book-PART* implies the book may not be finished whereas *I’m going to read a 
book-ACC* suggests the book will be read from start to finish.

In the current study, we first conducted a naturalistic corpus analysis of 
word order and case marking in transitive sentences in Finnish child-directed speech (CDS), in order to have 
a more informed idea of the types of transitive sentences Finnish children hear 
(study 1) and to derive predictions about their likely order of acquisition in relation 
to cue availability and reliability. We then investigated how Finnish children 
understand case marking and word order in causative transitive sentences at the 
ages of two, four and six years by manipulating word order and either unambiguous 
or ambiguous case marking (study 2).

2. Study 1: Word order and case marking in Finnish child-directed speech

We examined the frequencies of the various transitive word orders and case 
marking in Finnish child-directed speech (CDS). This analysis allowed us to make 
more accurate predictions for study 2 based on the cue competition model. As
discussed above, cue validity should determine the order of acquisition of case marking and word order cues and how children’s knowledge of them changes developmentally.

2.1. Method

Ten hours from an unpublished longitudinal corpus (Kirjavainen-Max Planck Finnish corpus\textsuperscript{2}) were utilised. The corpus consists of more than 200 hours of audio-recordings of a monolingual Finnish family. The corpus follows one child’s development from the age of 1;7 till 4;0 but the transcripts included in this analysis were from the time the child was 2;0-2;1 years old. Input in this age was particularly interesting to us as two-year-olds are the youngest age group in our experiment. Transitive sentences, including sentences with two nouns, two pronouns or one of each (N= 216), in child-directed speech were manually extracted from the transcripts as the corpus was not tagged for word classes that would have allowed CLAN searches for appropriate utterances. Next, the extracted utterances were coded for word order and case marking. Word order was coded in terms of whether the sentence was subject-initial or object-initial, as e.g. SVO or OSV. Also the sentence frame (e.g. NNV, NVN) and the cases of subjects and objects were coded. Case marking was coded as unambiguous when subjects and objects were distinguished by suffixes and ambiguous when subjects and objects could not be teased apart based on case marking alone due to overlapping forms (see above). When case marking was ambiguous, subjects and objects were identified based on animacy and context.

\textsuperscript{2} For details contact Minna Kirjavainen (leimk@ogu.ac.jp) or Elena Lieven (elena.lieven@manchester.ac.uk)
We wanted to know the relative input frequencies of the various transitive frames to determine how often children are exposed to agent-initial and patient-initial frames. We were also interested in how often case marking is ambiguous and unambiguous. We calculated cue validities separately for Argument+Verb and Verb+Argument frames as in Abbot-Smith & Serratrice (2013). All of the utterances consisted of two arguments but this way of calculating the cue validities was merely a way of capturing the two argument frames. We opted to calculate cue validities separately for each type of combination, because it is more informative of the relationship between each of the two arguments and the verb. If we had calculated cue validity for ‘first noun as agent’ cue only, we would have missed cue validities for arguments that come after the verb. For example, when we calculated the cue validities for a NNV frame, the frequency and the thematic role of noun right before verb (i.e. the second noun) was counted. For example, if the word order was OSV, we counted it as an argument+verb frame for the agent (e.g. Cereal Piia ate). Sentences in the SOV word order were counted as an argument+verb frames for the patient cue (e.g. Piia cereal ate). Cue availability equated to the raw frequency of the particular cue out of all transitive sentences included in the analysis and multiplied by a hundred to convert to percentages. For example, cue availability for the argument+verb frame consisted of counting the number of utterances that have a noun before the verb, regardless of whether it was an agent or a patient (e.g. Mommy brings the ball; Piia cereal ate). Cue reliability was determined by dividing the raw frequency of sentences where the cue in question denoted the correct argument structure role by the raw frequency of sentences that contained this cue. For example, when the reliability was calculated for the agent cue in the
argument+verb frame, **Mommy brings the ball** was counted towards the cue whereas **Piaa cereal ate** was not. Lastly, **cue validity** was calculated by multiplying cue availability and cue reliability. We also calculated the cue validity for the ‘first noun as agent’ cue to capture the validity of using the first noun as a cue to agency across all frames as it is a likely cue for children to use because of its saliency and primacy effect. All of the values were multiplied by a hundred to convert to percentages.

### 2.2. Results and discussion

Altogether 216 transitive sentences in the input were examined. Unambiguous case marking was found in 177 sentences (82%) of which 85 took a partitive direct object and 92 sentences were marked with an accusative direct object. Ambiguous case marking was found in 39 sentences (18%). Three word orders were attested: SVO (N = 205; 95%), OSV (N = 3; 1%), and SOV (N = 8; 4%). There were no examples of OVS. Thus, an overwhelming 213 sentences had a subject-first word order (99%).

Figure 4.1 shows the cue availability, reliability and validity values (%) for the agent cue across these transitive sentences. The argument+verb frame was the cue with the highest validity (96%), showing that agents are considerably more likely to be pre-verbal than post-verbal. In fact, there were no instances of post-verbal agents. Case marking was also a highly valid cue to agency (82%) but due to some instances of overlapping nominative and accusative plural forms, not as valid as the pre-verbal frame cue. Figure 4.1 also shows that the validity of the ‘first noun as agent’ cue (79%) is only slightly lower than that of case marking. The calculations confirmed what the descriptive statistics also showed: despite the fact that there
are many possible grammatically correct transitive word orders, the actual number of word orders attested in naturalistic child-directed speech is very limited and largely restricted to SVO.

**Figure 4.1.** Cue availability, reliability and validity values (%) for agent noun phrase (NP) frames in Finnish child-directed transitive sentences.³

Cue competition calculations for patient noun frames were in agreement with the word order frequencies (Figure 4.2.). They showed that Argument+Verb interpretation to patient has a very low cue validity (.04%) whereas the Verb+Argument frame is highly valid (90%). We calculated cue validities separately for the accusative and the partitive cases to allow a more direct comparison with our experiment results. The calculations show that the accusative case is less

³ N.B. Reliability and validity for Verb + Argument are 0% due to no instances of post-verbal agent, and thus do not to appear in the figure.
reliable than the partitive (70% vs 100%, respectively). The former is still a slightly more valid cue than the latter due to higher frequency, with validities of 43% and 39%, respectively. Overall, it is unlikely for Finnish children to hear patient-initial transitive sentences.

**Figure 4.2.** Cue availability, reliability and validity values (%) for patient noun phrase (NP) frames in Finnish child-directed transitive sentences. The values were calculated separately for the accusative and the partitive cases.

The results have provided insight into the input of transitive sentences that Finnish children hear. Based on these data, there is very little variation in transitive word order in Finnish CDS. This is surprising since case marking allows grammatically correct, alternative word orders but children do not seem to be exposed to them.
The results are interpreted as children hearing the majority of transitive sentences in the SVO word order with a small proportion of two Noun+Noun+Verb (NNV) orders: SOV and OSV. However, even though patient-initial sentences are used to distinguish between given and new information in adult language, our data suggests young Finnish children may not be exposed to sentence types with patient as the first noun in CDS, at least at two years of age. Even when they do hear patient-initial sentences, they are almost exclusively SOV.

Even though the Finnish case marking system is relatively transparent compared to, for example, German noun endings, the most frequent transitive word order in child-directed Finnish is SVO. Our analysis demonstrates six aspects of transitive sentences in Finnish child-directed speech: 1) the first argument is very likely to be an agent 2) the second argument is very likely to be a patient 3) the agent is very likely to be in a pre-verbal position 4) the patient is very likely to be in a post-verbal position 5) case marking is fairly unambiguous and a valid cue. Thus, the predictions for our comprehension experiment, based on previous literature and our input analysis, are as follows: 1) SVO frame with unambiguous case marking will be comprehended more easily compared to the other sentence types, even by the youngest age group. The youngest age group is predicted to perform significantly above chance only in conditions in which both word order and case marking are available: SVO+ and SOV+. 2) Four-, and six-year-olds will show an agent-initial bias in sentence types with ambiguous case marking. The youngest age group is predicted to perform at chance when only word order is available. 3) Cue conflict sentence types, i.e. patient-initial unambiguously case marked frames, will
be the most challenging conditions. Two-year-olds are unlikely to interpret them correctly whereas it is possible that four-year-olds might be beginning to rely on case marking due to transparency and reliability of the case endings. The oldest children will be starting to develop an awareness of case marking but will not yet appreciate in an adult-like way how it interacts with word order.

3. Study 2: Comprehension experiment

3.1. Method

3.1.1. Participants

Four age groups participated in the study: sixteen 2;6-year-olds (range: 2;2-2;11, eight girls and eight boys), sixteen 4;6-year-olds (range: 4;0-4;11, eight girls and eight boys), sixteen 6;9-year-olds (range: 6;3-7;2, six girls and ten boys) and eight adults as a baseline. Four more children were tested but excluded either because of lack of concentration (1) or side bias in pointing (3). The data were collected in nurseries in Joensuu, Finland. All participants were monolingual Finnish speakers with no reported language problems.

3.1.2. Materials

For this experiment, we decided to use transitive sentences in the past tense with patients marked with the accusative case. Finnish case marking is quite transparent without much overlap between the suffixes. Conditions with ambiguous case marking were needed to test children’s comprehension with word order as the only
cue. There is some overlap between case markers in the plural: plural accusative and plural nominative are identical. As a result, we used plural forms of agents and patients in the ‘word order only’ sentence types. All of our test sentences were in the past tense because the accusative case denotes completed actions. Patients would have been marked with the partitive case if the present tense was used and thus, no ambiguous case marking would have been possible.

Three novel verbs and actions were devised for the experiment. All of them were causative, reversible transitive verbs involving a nominative active agent and an accusative patient as the direct object. *Miikkaus* (‘meeking’) involved a patient standing on a hair roller while the agent pushed the hair roller three times trying to knock the agent off. The patient fell off after the third push. *Tammaus* (‘tamming’) was an action where the patient stood on a sponge while the agent was jumping on the patient’s back, pushing the patient deeper into the sponge. After three jumps, the agent jumped off and landed on the floor next to the sponge. Lastly, *viiffaus* (‘weefing’) began with the agent standing on the floor while the patient was on a ladder. The agent jumped on the same level with the patient, and tapped the patient three times until it fell off the ladder. After the patient fell off onto the floor, the agent returned to the same spot as in the beginning of the action. The videos of these novel actions lasted six seconds with a voice-over of the test sentence which the children heard twice during each simultaneous video pair. The audio stimuli were recorded by the experimenter as neutrally as possible with conscious effort to avoid emphasizing any argument. Finnish prosody is very consistent: stress is always on the first syllable and sentence intonation falls off.
All the animal names were common bi-syllabic animals that even the youngest children were likely to know. The animal pairings were based on properties of the toys. Animals that were fairly similar in size but different in colour were paired to avoid bias in which animal a child might visually focus on more but different colours made the animals easier to tell apart. Each pair appeared in two conditions: once in part 1 and part 2 of the experiment.

3.1.3. Design

A within-participants design of six conditions was utilised (see Table 4.1). Condition 1 measured children’s comprehension of the subject-object-verb frame with unambiguous case marking, indicated by the plus sign in the condition name (SOV+). Condition 2 measured a noun-verb-noun condition with ambiguous case marking, indicated by the minus sign in the condition name (SVO-). Note that there is not one correct interpretation in ambiguously case marked conditions as word order does not mark agent-patient relationships syntactically in Finnish. Condition 3 included the sentence type object-subject-verb with unambiguous case marking (OSV+). Condition 4 was our sentence type with a coalition of cues with subject-verb-object order with unambiguous case marking (SVO+). Condition 5 was a control condition for conditions 1 and 3 with an ambiguously marked noun-noun-verb frame (SOV-). Lastly, condition 6 was an unambiguously marked object-verb-subject sentence type (OVS+). It is important to note that although the nouns of ambiguously case marked sentence types, i.e. condition 2 (SVO-) and condition 5 (SOV-), are labeled as subjects and objects, the agents and patients are not overtly marked in these conditions, but rather the sentences may be interpreted based on a first-noun-as-
agent strategy. There were three test sentences, one per novel verb, in each condition. To ensure the procedure was child friendly, the study was split between two consecutive days. The children completed conditions one, two and three on the first day and the remaining three conditions on day two. This way, the children heard nine test sentences on each day, which equals 18 test sentences altogether. The test items were reversible transitive sentences in that both agent and patient were animate. The children’s comprehension was tested with a pointing task. They saw two simultaneous videos of the novel actions with a voice-over that matched only one of the scenes. The children were then asked to point to a still picture of the matching scene. The order of the testing blocks and test sentences within each block was randomised for each participant. Also agency and sidedness were counterbalanced in the stimuli: the agents and patients as well as the sides of the target stimuli were reversed for half of the participants to prevent biases in pointing.

**Table 4.1.** Experiment conditions by condition number, name, meaning and day of testing.

<table>
<thead>
<tr>
<th>Condition number</th>
<th>Condition name</th>
<th>Details</th>
<th>Day of testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SOV+</td>
<td>Agent-Patient-Verb word order with unambiguous case marking</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>SVO-</td>
<td>Noun-Verb-Noun word order with ambiguous case marking</td>
<td>1</td>
</tr>
</tbody>
</table>
3.1.4. Procedure

Day 1 of the experiment consisted of three warm up tasks, familiarization of the three novel actions and nine test items which were evenly split between three conditions.

3.1.4.1. Part one

*Warm up.* The first task was aimed to warm up the child to singular and plural forms as both were used in the test items. The child was presented with a single toy animal, e.g. a dog, and two animal toys that were attached together that were the same kind as the single toy i.e. two dogs. The toys were placed apart from each other in front of the child. The child was asked to show ‘a dog’. If the child didn’t respond, the experimenter encouraged her by asking if the child could touch the dog. It is important to note that when the procedure is written up in English, the first prompt uses an indefinite article while the second prompt uses a definite
article, there is no difference in the original Finnish prompts. There are no
determiners in Finnish and the difference that they convey in English was not
present in the Finnish experiment script. Next, independently of the child’s
response, the dogs were replaced by rabbits. Similarly, the experimenter asked the
child to ‘show the rabbits’. This game was played altogether four times: twice with
dogs and twice with rabbits. On half of the occasions the child was asked to show a
single toy and the other times plural toys were requested.

The next warm up task involved familiarising the children with all the
animals used in the experimental stimuli and ensuring they recognized them. The
children were shown a series of photos on a laptop computer screen of the same
toy animals as in the video stimuli. There were two animals in each picture and the
experimenter requested the children to point to every animal in turn by asking
‘show me a [animal name]’. Children recognized all the animals with ease. The
third warm up task was conducted to ensure the children could point clearly to one
still picture out of two options on the same computer screen. The children saw four
picture pairs, one by one, which consisted of the same animals they had just been
familiarized with in the previous task. In each picture pair, the two animals were
doing a same, familiar intransitive action, e.g. eating or jumping. The children were
asked to point to only one of the pictures at a time. For example, the experimenter
requested children to point by saying ‘can you show me: a bear is eating’. If
children’s pointing was ambiguous, the experimenter guided them by showing how
to point and asking ‘can you point like this?’.
**Novel verb familiarisation.** Following the warm up tasks, the child was familiarised with the first novel action. The experimenter acted out the action with two toy animals that the child had already seen in the warm up pictures and said ‘Look, this is meeking. This is called meeking. Would you like to try meeking?’ (Katso, näin miikataan. Tämä on miikkaamista. Haluatko kokeilla miikkausta?). Children heard the novel verb inflected in three ways to ensure they treated it as an action instead of a label to an event. Similarly, each verb was modeled three times. The children were asked to try the act out to make the experiment more fun and interactive but it was not essential. The order of the actions was counterbalanced for each child.

**Testing.** After learning only one novel action, the child watched three pairs of videos of the action, one pair at a time. During each video pairing, the child heard a sentence twice which matched only one of the videos. After the videos, two still pictures matching the videos appeared on the screen and child was asked to point to the scene that matched the audio (e.g. Can you show me: the wolf-NOM meeked the goose-ACC; Näytä mulle: susi miikkasi hanhen). Each video pairing of the same novel action corresponded to a different experimental condition, which in the first part of the experiment were unambiguous SOV order (SOV +), ambiguous SVO order (SVO -) and unambiguous OSV order (OSV +). The testing procedure was repeated with the other two novel actions. It is important to remember that when case marking is ambiguous, i.e. in condition 2 (SVO -), the sentence could plausibly be interpreted as agent-initial or patient-initial.
3.1.4.2. Part two

The second part of the experiment was conducted on the following day. The testing blocks were the same as in part one, i.e. they consisted of three test sentences per block, representing every condition included in part two: SVO with unambiguous case marking (SVO +), SOV with ambiguous case marking (SOV -) and OVS with unambiguous case marking (OVS +). Each testing block involved only one novel action at a time but the novel actions used in part two were the same as in part one. As in part one, the experimenter acted out the novel actions before every testing block but this time the children were not asked to repeat the act out. The testing procedure was otherwise identical to part one. All the children did the same conditions on days 1 and 2. The reason for doing this instead of counterbalancing the six conditions between the testing days was to ensure there was an equal spread of similar manipulations each day. On day 1, the children did conditions SOV+, SVO- and OSV+ which involved a varied range of manipulations: one sentence type with first noun as agent and unambiguous case marking, one sentence type with ambiguous case marking and one sentence type with first noun as patient and unambiguous case marking. On day 2, the children did conditions SVO+, SOV- and OVS+. Again, there was a sentence type with first noun as agent with unambiguous case marking, a sentence type with ambiguous case marking and a sentence type with first noun as patient and an unambiguous case marking. In this way, we aimed to minimise interaction between the sentence types. For example, if we had included two patient-initial sentence types for the same day, they might
have primed each other as well as comprehension of the third sentence type on that testing day.

4. Results

We calculated the mean percentages for ‘first noun as agent’ interpretations in each condition. We could not score the data as ‘correct’ or ‘incorrect’ since the word-order only conditions 2 (SVO-) and 5 (SOV-) with ambiguous case marking do not have a clear, correct interpretation. If the child did not point clearly, pointed to both scenes or did not point at all, the trial was excluded. Altogether three trials out of 288 were excluded in the youngest age group in condition 5 (SOV-) and condition 6 (OVS+). Furthermore, six trials were excluded in the four-year-olds between all conditions except condition 2 (SVO-). No trials were excluded in six-year-olds or adults.

We started by testing the results against chance using one-sample t-tests (50%). In the two agent-initial conditions with case marking (SOV+ and SVO+), all groups were above chance: t(15) = 3.020, p = .009 (two-year-olds), t(15) = 3.503, p = .003 (four-year-olds) and t(15) = 16.102, p < .001 (six-year-olds) for SOV+; t(15) = 7.263, p < .001 (two-year-olds), t(15) = 3.656, p = .002 (four-year-olds), and the six-year-olds showed a ceiling effect in SVO+. In the two ambiguous, word-order only conditions (i.e. two nouns without case marking (SVO- and SOV-), only the two-year-olds were at chance (SVO- t(15), p = .079; SOV- t(15) = 1.140, p = .27), whereas the others were above chance, (four-year-olds SVO- t(15) = 2.423, p = .029; SOV- t(15) = 4.521, p < .001; six-year-olds SVO- the statistic could not be computed due to a ceiling effect; SOV- t(15) = 7.889, p < .001). In the ‘conflict conditions’ (patient-
initial with case marking, OSV+ and OVS+), the two youngest groups were at chance for both (two-year-olds OSV+ $t(15) = 1.168, \ p = .261$; OVS+ $t(15) = 1.073, \ p = .300$; four-year-olds OSV+ $t(15) = -.532, \ p = .603$; OVS+ $t(15) = .193, \ p = .846$) and the six-year-olds were also at chance in OSV+$(t(15) = .259, \ p = .799)$ but were above chance in OVS+ $(t(15) = -.952, \ p < .001)$.

This suggests a clear developmental pattern in which agent-initial sentences with case marking are the first to be understood, that two-year-olds need the congruent case marking in order to interpret sentences with two nouns, while all other groups interpret these as agent-initial even without case marking. Finally, non-ambiguous, patient-first case marked sentences are the last to be understood by the child groups, though the six-year-olds can manage if the subject comes after the verb but not if it comes directly after the object and before the verb. Adults always showed either a ceiling effect or a floor effect without variety.

To get a more detailed insight into the results a two-way mixed factorial analysis of variance (ANOVA) was utilized to determine whether there were significant differences between the four age groups and six experimental conditions (sentence types) (Figure 3). The dependent variable was how often the first noun was chosen as the agent of the action in each condition and age group. The analysis revealed a main effect for sentence type, $F(5, 260) = 66.505, \ p < .001$, but not for age, $F(3, 52) = 2.287, \ p = .089$. The results also showed an interaction effect for sentence type and age, $F(15, 260) = 9.620, \ p < .001$ (Figure 4.4).

*Post hoc analyses.* Post hoc tests with the Bonferroni correction were conducted to examine the underlying causes of the interaction between sentence
type and age. Individual one-way ANOVA tests for each sentence type with an adjusted p-value of $p = .004$ showed the age groups performed significantly differently from each other in every condition. In condition 1, an agent-initial sentence type with unambiguous case marking (SOV+), six-year-olds and adults chose the ‘first noun as agent’ scene significantly more often than two-year-olds, $F(3, 52) = 4.976, p = .004$. In condition 2 (SVO-), two-, and four-year-olds performed differently from six-year-olds and adults by, again, not choosing the ‘first noun as agent’ scene as frequently as the two older groups, $F(3, 52) = 12.322, p = .004$. This finding was not predicted: we hypothesised that all age groups would show an agent-initial bias due to the overwhelming majority of transitive sentences having an SVO frame, according to our naturalistic input analysis.

In condition 3 with first noun as patient and unambiguous case marking (OSV+), all child age groups were significantly different from adults but no child group performed differently from one another, $F(3, 52) = 4.904, p = .004$. This time, the difference arose from the children interpreting the first noun as agent significantly more often than the adults did. In fact, the adults showed a floor effect as no adult participant chose the ‘first noun as agent’ scene in this sentence type with accusative case marking on the first noun. We predicted that patient-initial (‘conflict’) sentence types would be the most challenging conditions but also that six-year-olds would be beginning to rely on case marking. The prediction was met as condition 3 (OSV+) indeed was the most challenging condition for all child groups but on the other hand six-year-olds were able to utilize case marking in condition 6 (OVS+).
In condition 4, which was the most frequent sentence type (SVO+), only four-, and six-year-olds and four-year-olds and adults were significantly different, F(3, 52) = 3.889, p = .004. The difference arose from the four-year-olds scoring the lowest percentage (79.2%) in choosing the ‘first noun as agent’ scene compared to 87.5% for the two-year-olds and ceiling performance for six-year-olds and adults. This condition was predicted to be overall the easiest which is supported by the results, shown by mean percentages of first-noun-as-agent interpretations per age group. Also, two-year-olds were most accurate in this condition compared to their performance in any other condition.

In condition 5 (SOV-), six-year-olds and adults interpreted a first noun with ambiguous case marking as agent significantly more often than two-year-olds, F(3, 52) = 6.913, p = .004. We predicted that all age groups but two-year-olds would interpret the sentence type with the ‘first noun as agent’ strategy as no other cues were available. This prediction was completely met as two-year-olds showed no preference for one particular interpretation but the other two age groups preferred the ‘first noun as agent’ interpretation. In condition 6 (OVS+), two-, and four-year-olds differed significantly from six-year-olds and adults by interpreting the sentence based on word order instead of case marking, F(3, 52) = 9.832, p = .004). These results are in accordance with our hypothesis. We predicted that patient-initial sentence types would be the hardest and that the youngest children would perform at chance while six-year-olds would appreciate case marking as a more reliable cue than word order, which they did.
Figure 4.3. Mean proportion of ‘first noun as agent’ interpretations of each age group in every condition. Performance above chance with a 2-tailed t-test, adjusted $p = .002^*$. 
To summarise, our hypotheses were mostly supported but the results also included some surprising findings. First, we predicted that the SVO frame with unambiguous case marking (SVO+, coalition of cues condition) would be the easiest sentence type to interpret for all age groups. This hypothesis was supported, shown by the highest overall percentages of ‘first noun as agent’ responses averaged across children. In addition, all age groups performed above chance in this condition. Second, we predicted that all age groups would show an agent-initial bias when case marking is ambiguous. This hypothesis was supported in part. In condition 2 (SVO-) with a noun-verb-noun frame and ambiguous case marking and condition 5 (SOV-) with a noun-noun-verb frame and ambiguous case marking, two-year-olds were the only group that performed at chance while the other ages
preferred the ‘first noun as agent’ interpretation. Last, we also predicted that cue conflict sentence types with first noun as patient and unambiguous case marking would be the most challenging conditions as they are rare in CDS. More specifically, we predicted that six-year-olds might be beginning to rely on case marking as it is more transparent in Finnish than in e.g. German. German-speaking children treat case marking as a reliable cue from the age of seven (Dittmar et al., 2008). Younger children were not predicted to prefer one particular strategy. Again, this hypothesis was partly supported. All of the child groups were at chance in condition 3 (OSV+) with patient as the first noun and unambiguous case marking. It was confirmed that this sentence type would be the most challenging one. However, on the last sentence type, condition 6 (OVS+) with patient as the first noun and unambiguous case marking, six-year-olds performed significantly above chance while the other two child groups showed no preference for case marking or word order.

5. Discussion

We investigated the development of Finnish children’s comprehension of two grammatical cues, case marking and word order in order to assign agent-patient roles in transitive sentence contexts. First, we examined these two cues in transitive sentences in child-directed speech to gain insight into the kind of input Finnish children might be hearing. The results showed that despite the grammatical correctness of alternative word orders, SVO was the most dominant word order with as high a frequency as 95% of transitive sentences in CDS. Next, a pointing experiment was conducted to study children’s comprehension of these cues. Only novel verbs were used as this allowed us to control for the children’s experience
with those particular verbs. It is well established that children perform better with more frequent lexical items than low frequency or novel items (Ambridge, Kidd, Rowland & Theakston, 2015; Matthews, Lieven, Theakston & Tomasello, 2005; Theakston & Rowland, 2009) thus, novel verbs enabled us to avoid some potentially confounding factors. Three child age groups (2;6, 4;6, 6;9) and an adult baseline were tested with six sentence types, in three different word orders, crossed with ambiguous and unambiguous case marking.

The experimental results revealed interesting differences between the age groups and sentence types.

The coalition as cues sentence-type: First, all age groups performed above chance in condition 4 (SVO+), which was the ‘coalition of cues’ sentence type. We predicted that even the youngest age group would perform above chance with this sentence type as we found SVO to be the most frequent transitive word order in Finnish child-directed speech. The findings of the children’s performance in this condition supported the prediction. However, post hoc tests revealed a significant difference between four-, and six-year-olds and four-year-olds and adults whereas there was no statistical difference between two-year-olds, six-year-olds and adults. This finding is surprising, because we did not predict that four-year-olds would perform worse than the youngest age group. One possibility is that the four-year-olds’ performance reflects the gradual adjustment in the implicit weightings of the cues within the linguistic system, leading to a temporary worsening in performance. Further replication of this result is needed to determine its robustness.
**Subject-initial case marking:** Similarly to the SVO+ condition (condition 4), all age groups performed above chance in the SOV+ condition, (condition 1). Here, two-, and six-year-olds as well as two-year-olds and adults differed whereas there was no difference in how two-year-olds and four-year-olds performed in this condition. This suggests that even though most two- and four-year-olds were able to utilise one or both cues, their comprehension is developmentally different from the six-year-olds. However, although the oldest child group did not perform as consistently as adults, the results are showing that they have almost reached adult-like understanding of the cues when they hear a case-marked subject-initial sentence. So in this sentence-type, two adjacent nouns do not provide major difficulties providing the first noun is case marked as subject. This finding is consistent with our CDS cue values that revealed the ‘first noun as agent’ cue to be a strong cue with a validity of 79%.

While children’s performance with SVO+ was relatively homogeneous, condition 5 (SOV-) illustrates a clearer developmental pattern. Only two-year-olds were at chance when there were no additional cues to word order, followed by other age groups in order of age. Four-year-olds, six-year-olds and adults all performed significantly above chance, which showed that when other cues are either ambiguous or absent, they interpret transitive sentences based on the first noun as agent strategy. Again, this is in line with our input analysis.

**Conditions with ambiguous case marking:** There were two conditions where all but two-year-olds performed above chance: condition 2 (SVO-) and condition 5 (SOV-). In both of these, case marking was ambiguous. In the SVO- condition, even
though two-year-olds were at chance and four-year-olds performed above chance, their performances were not statistically different from each other. Both two- and four-year-olds performed differently from six-year-olds and adults who performed identically. The latter two always preferred the first noun as agent interpretation over 'second noun as agent’. These results are intriguing for two reasons. First, the SVO- sentence type was a contrast for the SVO+ sentence type discussed above. Comparing the two allowed us to find out whether young children follow a fixed word order strategy (e.g. first noun as agent) to parse transitive sentences or if they utilise case marking. Similarly, the SOV- condition was a contrast to the SOV+ and OSV+ sentence types, all containing two adjacent nouns. The finding that even the youngest age group was above chance in a condition with unambiguous case marking but at chance with ambiguous case marking suggests that even children as young as 2;6 have some sensitivity to case marking as a cue to agent-patient relations. Second, when six-year-olds and adults were presented with ambiguously case marked sentences, they reliably made first noun as agent interpretations. This reflects our finding in study 1: SVO is the most frequent transitive word order, taking up to 95% of all transitive word orders. Six-year-olds and adults show a strong preference of this word order when there are no other cues, presumably as an effect of frequency.

Our result of two-year-olds’ sensitivity to case marking is interesting when compared to the findings of Dittmar et al. (2008). Similarly to our study, they found that two-year-olds performed most accurately in the prototype condition with unambiguous case marking and the most frequent word order. However, German
two-year-olds were at chance when either cue was absent, implying that neither case marking nor word order alone are heavily weighted yet by the youngest age group. Dittmar et al. (2008) account for this finding by concluding prototype sentences with more than one cue are special ‘landmarks’ or anchors in a young child’s language representations that allow them to be built on and expanded to other sentence types. Our findings on child Finnish support this approach: two-year-olds can use the two cues in coalition but are at chance in sentence types with either ambiguous case marking or patient-initial word order with unambiguous case marking. The developmental differences seen between the age groups show how four- and six-year-olds are in the process of isolating the two cues.

If we compare Conditions 2 (SVO-) and 5 (SOV-), they show a slightly clearer developmental pattern compared to the SOV+ and SVO+ sentence types. While two-year-olds were at chance, four-year-olds were above chance despite not performing significantly better compared to the youngest age group. Six-year-olds outperform the other two child age groups significantly. Four-year-olds seem to have acquired more knowledge of word order frequencies and are more likely to interpret the first noun phrase as the agent than two-year-olds. However, they are in the developmental middle ground since their performance does not resemble that of adults and six-year-olds. This might be because even though SVO was observed to be a highly frequent transitive frame in Finnish CDS, the canonical transitive is a construction with a relatively low frequency across all constructions in CDS, at least for English (Cameron-Faulkner, Lieven & Tomasello, 2003) Thus, two- and four-year-olds might not have that much experience with straightforward SVO
transitive sentences, especially ones with two full noun phrases rather than at least one pronoun, which is why they are struggling to use word order as the only cue in comprehension.

Object first plus case-marking: In contrast, condition 3 (OSV+) and condition 6 (OVS+) paint a slightly different picture. All child groups were at chance with the former and none of the child groups performed statistically differently from each other. This was also the case for OVS+ with the exception that the six-year-olds were above chance. Study 1 showed that object-initial sentences are very rare and despite unambiguous case marking, children of all three ages found these sentence types challenging. Another factor that could make condition 3 (OSV+) even harder than condition 6 (OVS+) is that it contains two adjacent nouns whereas the noun phrases are separated by a verb in the latter. Even though all age groups were above chance in two agent-initial conditions with unambiguous case marking (condition 1 (SOV+) and condition 4 (SVO+)), children were significantly more competent with the SVO+ sentence type. More accurate comprehension of this sentence type largely stems from high frequency in CDS as found in study 1. However, input frequency does not explain six-year-olds’ better performance in condition 6 (OVS+) compared to condition 3 (OSV+) as we did not find any examples of OVS. The only object-initial word order found was in fact OSV. This supports the idea that two adjacent nouns in a transitive frame are more challenging to process, possibly due to forgetting the second noun when it is in the middle of a sentence. When the noun phrases are the first and last items in a sentence, they may be retrieved as a result of the primacy effect and the recency effect.
In condition 6 (OVS+), two- and four-year-olds chose the first noun as agent interpretation significantly more often than six-year-olds and adults. The two younger age groups did not differ statistically and neither did six-year-olds and adults. As discussed earlier, children are more sensitive to coalition of cues as opposed to one particular cue. Their performance is improved by the presence of both cues as those sentence types (SOV+, SVO+) probably sound most natural to them. The OSV+ sentence type might be particularly challenging because of its combined effects of low input frequency and lower saliency of the accusative case ending that denotes the patient. A case ending might be more salient if it is followed by a verb instead of a noun. Children could also utilize a strategy with OVS+ that is unavailable with OSV+: even if a case marker of a patient is more challenging to detect when it is followed by another word, a noun or a verb, children might find it easier to detect the lack of case marking in the last lexical item of the utterance, the nominative case of the agent. By hearing an unmarked noun phrase at the end of a sentence frame with two nouns and a verb might lead children to conclude the first noun phrase must have been the patient.

Other possibilities: One point to consider is children’s familiarity with plural forms as agents and patients. There were 53 plural patients out of 216 transitive sentences in CDS in study 1, whereas there were no plural agents. Thus, the input data suggests that transitive sentences with both plural agents and plural patients, like condition 2 (SVO-) and condition 5 (SOV-), are extremely rare. This could have affected the two-year-olds’ performance as they did not show an agent-initial preference with conditions 2 and 5. There is also evidence suggesting that parsing
sentences with two unmarked lexical noun phrases can create additional processing demands, even in adults (Bever, 1974), and eliciting sentences containing two lexical noun phrases can be challenging with children up the age of eight (Gropen, Pinker, Hollander, Goldberg & Wilson, 1989). The fact that all of our test sentences contained two noun phrases might have created additional cognitive load. It is of course possible that the precise pattern of results may change if children were tested on sentences containing differentiated noun phrases (e.g. pronouns, proper names, Rowland & Noble, 2011), especially if the patterning mapped on to the kinds of sentences children are regularly exposed to in terms of their form and information structural properties (Theakston, 2012).

Another possibility is that the particular challenges of these sentence types were highlighted by lack of context. Kaiser & Trueswell (2004) examined the extent of discourse context on comprehension of non-canonical word orders in Finnish adults. They found appropriate contexts to reduce the difficulty of infrequent constructions in a self-paced reading task and in an auditory-visual eye-tracking experiment. Grünloh, Lieven & Tomasello (2011) also found that a discourse context for object-first word order with the appropriate intonational pattern improved five-year-old German-speakers’ comprehension. Adults could interpret such sentences based on case marking alone, while five-year-olds utilized case marking and intonation only in coalition. These results showed that five-year-olds can interpret object-first word orders correctly under some circumstances, whereas according to findings by Dittmar et al. (2008) children under the age of seven systematically interpreted object-first sentences incorrectly. In our study, the
participants were familiarised with the novel verbs but no additional context beyond training was introduced which may have made comprehension more challenging.

It could also be argued that children are not initially acquiring a prototypical transitive structure. Perhaps the performance of the youngest age group could simply be explained by two-year-olds being more likely to perform accurately when they can use of the two cues (case marking or word order), to comprehend agent-patient relationships. Even if two-year-olds utilised only one cue when two were available, they were still more likely to perform correctly. It is a possibility that only the older age groups showed evidence for using two cues in coalition to comprehend transitive sentences.

A relatively recent approach to examining the nature of young children’s syntactic knowledge is to view representations of constructions as gradient. It is an alternative to thinking of language acquisition and development of other cognitive domains in terms of rigid developmental stages. Viewing representations of constructions as graded, plastic entities could provide us with a more fine grained insight into the discrepancies in performance between ages and tasks. Evidence for this is provided by a cross-linguistic study comparing English-speaking and German-speaking two- and four-year-olds (Abbot-Smith, Lieven & Tomasello, 2008). The experiment utilised a so called weird linking elicitation task to examine how children map semantic roles to syntax in transitive sentences. Children listened to transitives with novel and familiar action verbs. The sentences were grammatical but mismatched the scene children saw: with a sentence like ‘the lion is pushing the
“tiger”, the tiger was actually the agent in the scene. The authors observed a distinct difference between the English-speaking and German-speaking two-year-olds. A majority (66%) of the English-speakers avoided using novel verbs whereas German children provided novel verbs significantly more frequently. When the authors re-analysed the subgroups of responses that provided the appropriate target verb, the proportion of grammatically correct responses were the same for both English- and German-speaking children. Thus, there might be something about the prototypical, agent-initial English transitive construction with SVO word order that makes it harder for most two-year-old English-speaking children compared to German children of same age. One cognitive factor affecting the results might have been that perhaps German children identified two cues pointing to the agent (case marking and word order) while English speakers could rely on one cue, word order, only. This cues-as-coalition approach could help us understand the results of our experiment too: while it is not clear that Finnish two-year-olds or even four-year-olds understand the contribution of individual cues in agent-patient relations, their performance is markedly enhanced by the presence of both cues. Our findings illustrate that two-year-olds need more than just one cue to accurately interpret transitive sentences. These cues seem to be unambiguous case marking and first-noun-as-agent as the only conditions two-year-olds were above chance were condition 1 (SOV+) and condition 4 (SVO+).

More evidence for gradedness of construction representations is provided by a study that compared 2-, and 3-year-old English-speakers’ comprehension of the SVO word order in an act-out task as well as in an intermodal preferential looking
task (Chan, Meints, Lieven & Tomasello, 2009). The purpose of using two tasks to measure the same thing, comprehension of the English transitive, in the same experiment was to find out whether children would show differences in their performance between the tasks. Both familiar and novel verbs were included. The study revealed that the youngest children (2;0) were at chance in both tasks with novel verbs but showed comprehension with familiar verbs in preferential looking. The other two age groups, 2;9 and 3;5, were above chance with familiar and novel verbs in act-out and preferential looking, but children aged 2;9 showed better understanding of familiar than novel verbs in act-out, the more challenging task. These results illustrate how representations of constructions are graded in different ages. In our study, children from all three age groups can decipher agents and patients under some conditions but not in all of them. This shows gradedness of construction representations between age groups, as there are clear developmental differences between them, in addition to differences between sentence types within each age group too.

6. Conclusion

Overall, our research has shown that, even though Finnish-speaking children as young as 2;6 can utilise case marking in coalition with the most frequent word, SVO, only six-year-olds were able to weigh word order and case marking in adult-like ways in most conditions. However, the results also suggest that even in six-year-olds, comprehension is still developing as they performed at chance in a conflict condition (OSV with unambiguous case marking) whereas adults consistently adopted a first noun as agent strategy. Thus, even though none of the age groups
was completely oblivious to grammatical cues, children’s understanding of individual cues seems to mature gradually. Our findings are in agreement with previous findings suggesting that prototypical sentences with a coalition of cues, i.e. SVO with unambiguous case marking in Finnish, are a cornerstone of the language acquisition process (Chan et al., 2009; Dittmar et al., 2008; Slobin & Bever, 1982). Sentences with a coalition of cues might provide a framework for young children for teasing apart functions of individual cues.
References


Chapter 5: General discussion

Two naturalistic studies and one experiment were designed to investigate aspects of the acquisition of noun case marking by Finnish-speaking children. A dense corpus consisting of 18 hours of naturalistic mother-child interaction was audio-recorded at the age of 1;7-1;8 for the purpose of the first two studies. In study 1, the extent of the child’s productivity was compared against the mother’s use of noun inflections. Productivity was measured as the number of different inflections used with each noun. The study was designed particularly to examine whether a young child shows adult-like productivity with case marking in the earliest stages of grammar development or whether the acquisition process is gradual. The results showed that there was a significant difference in the productivity of the two samples: the child used fewer inflections per noun than the mother did. These results support the more gradual, usage-based, account of language acquisition.

It has been established that frequency can account for some patterns in the acquisition process. Knowing this motivated another set of naturalistic analyses in study 2 that examined the kinds of frequencies that could be related to a child’s accuracy with noun case marking. The same corpus as in the previous study was utilised. We examined seven measures of input frequency in relation to a child’s accuracy with case marking. First, we examined whether the token frequencies of individual nouns correlated with the child’s case marking accuracy, defined as provision of the required inflection in obligatory contexts, with those same nouns. Second, we correlated input token frequencies of the different case markers with the child’s accuracy with those same case markers. Third, as type and token
frequencies have been shown to have different effects on productivity, we correlated input type frequencies of the different case endings with the child’s accuracy. These three measures were based on a sample of nouns that the child made errors with. The next two measures are based on the entire input corpus to eliminate potential verb-specific biases in the overall case marking frequencies. Thus, the fourth measure examined the relationship between the absolute input frequencies of case endings from the entire input corpus and the child’s likelihood of extending cases to incorrect contexts. The fifth measure repeated the previous measure but with type input frequencies. We also wanted to know whether children are more sensitive to the combinations of nouns and case endings than the independent frequency of each in isolation which led us to correlate the input frequencies of noun+case marking combinations with the child’s accuracy with these combinations. These six measures examined the effect of absolute frequency. Last, we correlated the relative (%) frequencies of particular noun+case combinations as a function of all uses in the input with the child’s accuracy rates for those same combinations to examine whether children are sensitive to the competition between noun+case marking combinations. We found that the child was more accurate with nouns that have high token input frequency compared to lower frequencies, which confirmed findings in previous literature. It was also found that absolute type frequencies of case endings marginally affected the likelihood of the child extending case usage to incorrect contexts. No other measure revealed a significant overall correlation even though some individual noun+case marking combinations reached significance for both absolute and relative frequency measures. The findings regarding noun+case combinations did not support the
predictions, as an effect was expected for all the different case endings but only some of them were significant with both absolute and relative frequency measures for the noun+case combinations. An issue that must be explored by future research is the interaction between type and token frequencies, and absolute and relative frequencies to allow more accurate predictions. Non-significant findings might be a result of the child’s age, too. It is a possibility that children’s sensitivity to different aspects of frequency change during development, and consequently, different frequency effects might be found as a function of age if the child’s accuracy was examined against the same input measures in a later developmental phase. It may also be that the semantic properties of some case endings makes them more challenging to learn at such an early stage. Overall, the study showed that when there is an overwhelmingly frequent form, the nominative in Finnish, children are more likely to be accurate with it in obligatory contexts, as well as more likely to extend it to incorrect contexts compared to forms with low frequency.

While the naturalistic studies focussed on a child’s speech production, the third study investigated how Finnish-speaking children comprehend case marking and word order in transitive sentences with novel verbs. The agents were in the nominative case, which in Finnish equates to the unmarked form of the noun, and the patients took accusative marking. In this study two-, four-, and six-year-old children and an adult control group saw two videos simultaneously of animal toys acting out a novel transitive action. During the simultaneous videos the participants heard a voice-over, test sentence that matched only one of the videos. The participants were asked to point to the scene that matched the test sentence. The
sentences were manipulated for word order and case marking, making the latter either ambiguous or unambiguous. We found developmental differences between the age groups in all but one condition. Two-year-olds relied most on the coalition of word order and case marking out of all age groups, four-year-olds could interpret sentences based on word order alone but struggled with unambiguously case marked, patient-initial sentences and six-year-olds had the most sophisticated appreciation of the two cues both in coalition and independently. However, they struggled with sentences with a noun-noun-verb word order when the first noun was an unambiguously marked as patient. The experiment demonstrated the importance of sentences with a coalition of cues in the language acquisition process.

1. Theoretical issues

Each of our studies addresses a different theoretical issue. While Study 1 compares nativist-generativist and constructivist-usage-based theories with different underlying assumptions, distinguishing between them is empirically challenging. The difficulties arise from the fact that the approaches can account for the same phenomena, but the assumed mechanisms are different. Another theoretical issue relates to the development of noun inflection. Our results showed that a young child does not combine as many inflections with a pool of nouns as her primary caregiver. As the data covered only four weeks of the child’s life, we could not meaningfully investigate the developmental trajectory of inflection acquisition. A key question for the theoretical approaches discussed is when children bridge that gap in productivity and use as many inflections with nouns, and vice versa, as adults. A detailed study of development across from about 18 months to three or four
years of age with dense enough data would make a hugely important contribution to this question. It is crucial for the usage-based approach to have a more detailed account of this key developmental stage in order to have an understanding of when children have received sufficient linguistic experience to start generalising morphology fully and reliably. Currently it is unclear how much input is enough to move from the stage of operating partly restricted construction frames, such as ‘He’s VERBing’ or Finnish noun + case marking combinations, to operating a fully abstract language system.

Study 2 raises issues of the role of frequency and the types of frequency young children are sensitive to in the input. Constructivist and usage-based approaches emphasise the role of frequency as an essential part of language acquisition and learning in general. While our results did demonstrate a positive correlation between the child’s case marking accuracy and noun input frequency, our other measures of input frequency did not reveal a more detailed pattern in case marking accuracy. It may be that there are other factors affecting the child’s accuracy that also interact with frequency effects, which possibly made it more difficult to examine frequency effects only. As further discussed in chapter 3, one such factor might be the phonological properties of the nouns and/or case endings. The child might struggle to pronounce particular noun+case combinations, particular noun stems, or particular case endings, which could result in the child omitting certain case endings regardless of their frequency. Another factor could be the transition between case endings. Krajewski, Theakston & Lieven (2011) showed that the degree of children’s productivity with inflectional morphology is affected
by the case ending children hear just prior to the point they use a case marked noun. The study revealed that children are more proficient at using some case markers after being exposed to certain different case markers from the child’s target case endings. Consequently, hearing some other case markers decreased children’s productivity with the target case endings. These are unique findings, and the effect is not well understood yet. It might be an effect that impacts languages with rich morphology more compared to languages with less inflectional morphology. A third factor could be the semantic properties of the case endings. While this is challenging to explore with naturalistic data, there may be conceptual differences between the case endings that make some endings more error-prone than others due to the meaning each ending carries. Generally speaking, it is a possibility that as no relative frequency effects were found for most cases in Finnish unlike in the previous literature, these others factors may interact at varying degrees depending on the properties of the child’s target language. Finnish, for instance, is a language with complex phonology entailing consonant gradation and vowel harmony, which could have a different impact on the various frequency effects found in literature on children acquiring languages with different phonology to Finnish.

Another theoretical issue relevant to the input frequency literature is which aspects of frequency are relevant, and whether these effects are stable across development or change as children’s knowledge and experience with language increases. One method to explore this is computational modelling. As mentioned earlier in chapters 2 and 3, children’s input has been utilised to account for their
optional infinitive errors in a model of syntax acquisition (Freudenthal, Pine, Aguado-Orea & Gobet, 2007; Freudenthal, Pine, Aguado-Orea & Gobet, 2010).

Study 3 highlights some questions regarding the Cue Competition Model. One of the strengths of the model is that it aims to make tangible predictions of the order of acquisition of grammatical cues. A problem that has been raised before (see Dittmar et al., 2008) is that the predictions based on cue validity calculations are not always accurate because it is not always clear which cues should be counted and how. For example, Dittmar et al. (2008) pointed out that the cue calculations might miss cue properties that are important, if not essential, for acquisition. Dittmar et al. (2008) conducted a series of experiments of German children’s comprehension of transitive sentences with various manipulations of word order and case marking cues across three age groups. They found four-year-olds’ performance inconsistent with input-based predictions, which were based on the Cue competition model. When the cue validity calculations followed the formula used by Kempe & MacWhinney (1998), the validity of word order was underestimated. It was proposed that perhaps children use word order as a cue to detect the position of a noun, e.g. the agent, and an inflected verb. If this is the case, word order as a cue would be available in fragments, too, which would increase the overall validity of the cue. On the other hand, it may be that children comprehend case marking in an item-based way that the model does not account for. It might be more informative to calculate cue validities separately for different inflections instead of case marking as a whole. This way, cue validities for the different case marking cues would be lower than for case marking with all inflections together,
which then would bridge the gap between the validities of case marking and word order. When the authors re-calculated the cue validities separately for individual markers instead for case markers as a group, the validity for case marking was considerably lower than for word order (21% and 87%, respectively). The model also does not consider the fact that children may be sensitive to different aspects of cue validity in different developmental stages (see Sokolov, 1988). Another theoretical issue is shared by study 1: even though the model makes predictions of children’s performance based on their input, it does not address at what age children begin to weigh different cues in adult-like ways. Some researchers (Rispoli, 1991; Tomasello, 2003) have voiced their concerns regarding a theoretical inconsistency between the model and the constructivist approach. As previously discussed, nativist-generativist approaches assume that even young children operate with linguistic categories like subject, verb and object, while such operations are assumed considerably later in development by constructivist and usage-based approaches. In order for the cue competition model to work, children must be assumed to have subject and object categories or something equivalent. This is not in line with the theoretical background the model originated from. On the other hand, sensitivity to the relationship between the first noun of a sentence and the agent of action does not equate to having a subject category, but could provide a means of interpreting the kinds of sentences used in cue competition-style experiments. As discussed earlier in Chapter 1 of the thesis, form-meaning mappings are not direct. Even though the first noun of e.g. a transitive sentence often is the agent in English and in Finnish (e.g. The lion chases the gazelle), there are times when the first noun takes a different role, such as the experiencer (e.g. The lion fears the gazelle).
Examining what cues speakers rely on in language processing is a multifaceted issue that is illustrated well by the complexity of these form-meaning mappings.

2. Methodological issues

Each study also highlights some methodological issues. A common issue in naturalistic research is transcription and coding. First, the data was collected with an audio-recorder only and therefore without video data. Transcribing audio-recorded data is more challenging than transcribing data with visual feedback – particularly when one of the participants is under two-years-old. Even though transcription is supposed to involve typing speech as it is, the transcriber does essentially make a decision every time an utterance is typed. This is particularly pertinent to transcribing a language with rich morphology. Stress is always placed on the first syllable in Finnish, meaning that speakers have a tendency to articulate word endings less clearly than beginnings of words. Case markers are at the very end of words, which made hearing them quite challenging at times. This issue was tackled by listening to unclear utterances multiple times, which often resolved the problem. If an inflection remained unclear after a number of times being listening to, the case marking of that noun was coded as ‘ambiguous’. This issue means that the available sample of inflections was smaller than the actual amount of inflections produced by the child during the recordings. Although this may have affected all inflections equally, one possibility is that certain inflectional endings are phonologically more difficult to hear than others, thus potentially underestimating the child’s knowledge of those forms.
Additionally, coding for errors can be subjective. For instance, a seemingly correct use of an inflection might not be the target the child meant to produce. In these cases, contextual factors were considered to make a decision of the target case. This is another reason why video recorded data is easier to transcribe and analyse compared to audio-recordings, as context is clearer in combination of visual information than without. Most of the time the context was well-established, but it is an aspect that researchers might want to be mindful of when naturalistic studies are designed.

In addition to the transcription issue, another potential problem with study 1 was the randomisation procedure of the two samples. To ensure that the results were not skewed due to a sampling issue, we matched the larger adult sample to the smaller child sample by randomly reducing the adult sample size. For example Krajewski et al., (2012) repeated the procedure 1000 times using automated computational methods which does ensure that the samples are randomised as effectively as possible. In the case of our study, the randomisation was done manually five times. This was deemed sufficient as the samples after each randomisation were consistent with each other in terms of the number of inflections per noun in the input sample. As there was little variance, it is unlikely to have had a significant impact on the findings. Overall, the findings of our productivity study closely resemble the findings by Krajewski et al. (2012). Both studies conclude that there is a gap between child and adult productivity with noun inflection, demonstrating the gradualness of the acquisition process.
In study 3, we needed to manipulate test sentences in terms of word order and the level of case marking ambiguity. This was challenging as there is not much overlap between the different case markers in Finnish. There is some syncretism in plural forms: the plural nominative and plural accusative cases are identical. As the test sentences were causative transitives, the agent took a nominative marking and the patient was marked with an accusative. When we manipulated case marking, we used the syncretic plural forms to make case marking ambiguous and the different singular forms when case marking was unambiguous. Ideally we would have liked to use singular forms in both ambiguous and unambiguous sentences to keep the stimuli as controlled as possible. It is important to note, though, that the participants never needed to distinguish between a singular and a plural form in a same sentence. In that sense, no additional cognitive load was introduced in sentences with plural noun forms.

Another methodological issue with the experiment relates to making accurate predictions of the developmental trajectory of cue comprehension. We conducted an input analysis of word order and case marking in the transitive sentences a two-year-old Finnish-speaking child hears. This was useful for having a general idea of, for instance, the ratio of different word orders but it was still challenging to make detailed predictions of the comprehension of four-, and six-year-old children based on it. It is known that child-directed speech changes as a child develops with e.g. more diverse vocabulary and total amount of speech (Rowe, Pan & Ayoub, 2005), but it is unclear whether this is an effect of adults reacting to the child’s improving language skills or whether the child’s language develops partly
as an effect of adults’ more complex child-directed speech when the child is older. The implication of this for our study was that we could not reliably make predictions of cue validities for children older than two as child-directed speech aimed at two-year-olds is very likely to be different than CDS for older children. Ideally, we would have needed naturalistic CDS data from all three ages included in the experiment to make more informed hypotheses of the four-, and six-year-olds’ performance.

3. Implications for future research

Our studies have shown that the language acquisition field still needs to work on developing more robust theories to account for various phenomena. First, the field currently lacks a detailed model of inflectional acquisition that can account for the acquisition of a complex inflection system like Finnish, Polish and Hungarian. As one part of a more robust model of child language, the precise role of frequency, and of different kinds of frequency effects, must be well defined and considered regardless of the theoretical approach to child language. While frequency is not suggested to be the most important mechanism in learning (Ambridge, Kidd, Rowland & Theakston, 2015), there is a vast body of research suggesting that it is a significant factor and its effects should not be underestimated. Due to the attention frequency effects have received, many generativist accounts are more mindful of frequency than in the earlier stages of child language research (e.g. Legate & Yang, 2007). The next step in the field is to account for the complexity of frequency effects, including type and token frequency and raw and relative frequency (Ambridge et al., 2015).
The field would also benefit from an account that could explain what other factors might be important at different ages of acquisition. For example, the role of working memory in language skills has been well established and is associated with e.g. reading comprehension (Daneman & Carpenter, 1980; Swanson, 1999) and word recall and phonological skills (Gathercole & Baddley, 1990). Thus, immaturity of a child’s working memory has a significant impact on their language processing. It has been found that young children struggle to prioritise information effectively in working memory; seven-year-olds, pre-teens and adults perform similarly in visual memory tests with small arrays of stimuli, whereas the youngest group struggled to pay attention to, and memorise, relevant information when cognitive load was increased (Cowan, Morey, AuBuchon, Zwilling & Gilchrist, 2010). While this study utilised visual stimuli only, modality effects have been found across different tasks assessing attention capacity. Children with Specific Language Impairment (SLI) struggle more with linguistic and non-linguistic auditory attention tasks than their typically developing peers whereas both groups perform equally with visual stimuli (Spaulding, Plante & Vance, 2008). It has also been found that typically developing children outperform children with SLI in tasks measuring linguistic resource allocation skills (Montgomery & Evans, 2009).

In fact, some of the findings of study 3 might be related to the working memory research by Cowan et al. (2010). Study 3 showed the importance of cue coalition in sentence comprehension in different stages of development. These kinds of studies (see Abbot-Smith & Serratrice, 2013; Chan, Lieven & Tomasello, 2009; Dittmar et al., 2008, MacWhinney, Bates & Kliegl, 1984; MacWhinney, Pléh &
Bates, 1985) aim to investigate when and how children tease apart individual grammatical cues and how they are weighted against each other. In addition to having less linguistic experience, younger children may struggle with multiple cues due to inaccurate attention to the most relevant cue(s) in a sentence. Pragmatic relevance of sentences is also important to accurate comprehension. The input analysis in Chapter 4 showed that even though alternative word orders are grammatical in Finnish, they are not utilised frequently in child-directed speech. This has implications for children’s comprehension: children find alternative word orders, as opposed to the most frequent one(s), considerably more challenging. These kinds of word orders, e.g. object-initial transitives in Finnish, are used for specific pragmatic purposes, and pose processing demands when used without a fitting pragmatic context (Lidzba, Konietzko, Schwilling, Krägeloh-Mann & Winkler, 2013; Smolík, 2015) Cowan et al.’s (2010) findings also imply that working memory begins to work in adult-like ways with small arrays of stimuli only around the age of seven. This is also the age that Dittmar and colleagues (2008) found children to weigh case marking and word order in adult-like ways. In our study 3, six-year-old participants performed well in most conditions, but their overall performance was not adult-like. Their behaviour could be partly explained by still maturing attention and working memory capacity. Even though this kind of information is important in making judgements of the precise level of comprehension in different developmental stages, cue competition studies have shown that understanding coalition of cues is also important in understanding syntax acquisition. Cross-linguistic studies have highlighted the universality of cue coalition in language development which shows that two or more cues pointing to a particular
interpretation is more effective than one single cue. This finding is intuitively logical, and is supported also by a number of other studies. For example, it has been found that using visual communication methods (e.g. pictures and symbols) in addition to verbal language supports language development of children with learning difficulties (Charlop-Christy, Carpenter, Le, LeBlanc & Kellet, 2002; Hibbing & Rankin-Erickson, 2003; Schreibman & Stahmer, 2014).

A recent study on children’s acquisition of time related words, e.g. ‘a minute’ and ‘an hour’, also demonstrates how young children struggle to prioritise seemingly conflicting information appropriately (Tillman & Barner, 2015). Children may infer the relative order of time words, such as that a minute is shorter than an hour and an hour is shorter than a day, from input but in reality, the true meaning of this type of terminology is acquired only around the age of seven or eight. In one task, three- to seven-year-old children were asked to make judgements of durations with questions like ‘Farmer Brown jumped for a minute. Captain Brown jumped for an hour. Who jumped more?’ The participants were able to make accurate judgements when they were asked to do this based on time words only by the age of four. However, when the task was amended by adding superficially conflicting numbers, e.g. ‘Farmer Brown jumped for three minutes. Captain Brown jumped for two hours. Who jumped more?’, the children found it considerably more challenging compared to the first scenario. Children showed more in-depth understanding of durations only at the age of six to seven, which is in accordance with the findings of children’s working memory. It seems like children under the age
of seven misjudge the importance of rank of cues in a variety of cognitive tasks, which may lead to misunderstandings in cue conflict contexts.

Finally, an important source of cue coalition for language comprehension and production is the combination of speech and gestures, such as pointing. Studies on the effects of parental gesturing with speech show that, generally speaking, gestures support children’s language comprehension. There are certain contexts, just like with grammatical cue coalition, that make it more likely for the child to utilise multimodal cues, such as simultaneity between gesture and speech (Gogate, Bahrick & Watson, 2000), supportive use of speech and gesture instead of contradictory use (Bates, Thall, Whitesell, Fenson & Oakes, 1989) and age – children aged two to three years utilise multimodal cues less effectively than five- to six-year-old children (McNeill, Alibali & Evans, 2000). These findings, again, show that cue coalition is a vital part of language acquisition and is not restricted merely to grammar development.

An interesting study idea would be to explore the role of cue coalition in language acquisition in other contexts than grammatical development more thoroughly. As it seems to play a more important role in language development than perhaps previously thought, investigating its precise contribution could improve our understanding of typical and atypical cognitive development. For example, coalition of prosody and language comprehension could provide interesting insight into how important non-verbal but vocal properties, such as intonation, are in supporting the acquisition of questions versus declarative sentences.
These data show that the overall developmental trajectory of the acquisition of Finnish noun morphology might go somewhat like this: Finnish children will already use an impressive number of case inflections by the very young age of 1;7, but not as many different inflections per noun as an adult. In the same developmental period, a Finnish-speaking child will make some errors, albeit not many different kinds. The vast majority of errors found in our data were errors of omission, i.e. the child provided the stem (the nominative form) without an appropriate case ending, which might be a function of defaulting to the single most frequent case across the entire sample. It does seem that a child’s use of case markers is still developing and while they may use a case marked noun in an obligatory context, use of inflected nouns is not completely reliable yet, let alone abstract. Finally, the third study paints a picture of the contribution of noun inflection in transitive sentences with multiple cues – which is what children have to do when they are acquiring their first language. The experiment suggests that children aged two, four and six are sensitive to a coalition of grammatical cues to support their comprehension of transitive sentences. Particularly the youngest children rely on cue coalition, which shows the importance of these kinds of sentences in the early stages of acquisition. Around the age of four, Finnish children will have a more sophisticated understanding of word order without the support of case marking, but continue to find conflicting cues challenging. By the age of six, Finnish speakers are weighing the two cues accurately with most word order and case marking combinations but do still struggle with case marked but adjacent nouns. Children will still continue to adjust the relative contribution of each
grammatical cue appropriately probably until the age of seven, as found by Dittmar et al. (2008).

The child language field lacks an account that can explain the inflectional morphology acquisition at the level of a single morpheme. Currently, the issue of morphology acquisition is approached at the level of how morphology emerges generally. A detailed account of single morpheme acquisition might provide more insight into the question of why children struggle more with some morphemes than others as it might be a function of the acquisition mechanism rather than a property of an inflection (e.g. concrete vs. abstract semantics of a morpheme). For example, a study comparing the acquisition of a very limited set of morphemes with different phonological and semantic properties, but roughly the same input frequencies, could shed some light on this question.

4. Conclusion

This thesis investigated the acquisition of morphology acquisition by focusing on Finnish children’s acquisition of noun inflections. This was achieved by conducting two naturalistic case studies on a Finnish child’s noun case marking productivity and noun case marking errors, both of which were compared against an adult control. A comprehension experiment on interaction of case marking and word order in transitive sentences was also conducted. The studies demonstrate the gradual nature of inflection acquisition, and language acquisition in general: superficially advanced, case marking productivity is in fact more restricted compared to adult productivity; this productivity is affected by noun type frequency; and even six-
year-olds do not perform completely in adult-like ways in a methodologically simple comprehension task.

This thesis is an addition to research on morphology acquisition in children acquiring a complex morphological system. The three studies reported in this thesis as well as other research in the field suggest that a strong account of language acquisition must include and define more clearly the role of frequency, different types of frequency, the gradual nature of language acquisition and the function of cue coalition. Examining Finnish allowed the exploration of a fairly transparent and consistent case system without interference from e.g. gender marking, definiteness, noun class or other noun morphology categories. An implication for the field is that it must continue to develop an account that can fully explain the morphology acquisition process generally as well as at a single morpheme level. All of the phenomena discussed in the thesis are interrelated – productivity, errors, combination of grammatical cues – and must be accounted for by a neat, cognitively economical theory.
References


APPENDICES

APPENDIX A. Recording guidelines for the parents for Studies 1 and 2.

APPENDIX B. Complete set of test sentences used in Study 3. The sentences are translated from Finnish to English.

APPENDIX C. Information sheet for parents for Studies 1 and 2.

APPENDIX D. Information sheet for parents for Study 3.
APPENDIX A. Recording guidelines for the parents for Studies 1 and 2.

Guidelines for the parents

1. The context of each recording should be a play situation at home consisting of interaction between the child and the parent, who has agreed to participate in the study.

2. Each recording should take approximately 60 minutes.

3. You are requested to record altogether five hours per week. Should you miss e.g. an hour one week, you are requested to make it up by recording six hours another week.

4. The recording situation should be as free from distraction as possible. For example, it is preferable to switch off other sources of sound like TV and the radio.

5. We are interested in spontaneous utterances of the child and the parent. Thus, we do not encourage reading during recordings.

6. To ensure a high quality of the recordings, please stay in the same room with the audio-recorder. If you and/or your child wish to leave the room during a recording, please take recording device with you.
7. To ensure we are able to collect a sufficient and comparable sample of child and maternal speech, it would be helpful if you could play with a particular set of toys e.g. for half of each recording.

8. Should the child be upset, tired or restless during a recording, it may be paused and continued at a more suitable time.

9. Please locate the recorder out of the child’s reach without compromising the quality of the audio (i.e. not too far).

10. To ensure your anonymity, the audio files must be encrypted before e-mailing them to the investigator. First, the audio files must be transferred from the audio-recorder to your computer. Next, you need to download free software for encryption called AxCrypt (http://sourceforge.net/projects/axcrypt/) if you do not have one already installed on your computer. Click ‘Download’ and follow the installation instructions. Once the software is appropriately installed, click a sound file with the right button of a computer mouse. A selection menu with ‘AxCrypt’ will appear. Click ‘Encrypt a copy’. Choose a password and an encrypted sound file will appear. Repeat the procedure to all sound files you are going to send to the investigator. The password will be given to the investigator over the telephone to ensure confidentiality of the data. Please also keep a copy of the audio files on your computer until the investigator notifies you that the recordings have been backed up to avoid any loss of data.
APPENDIX B. Complete set of test sentences used in Study 3. The sentences are translated from Finnish to English.

### DAY 1

<table>
<thead>
<tr>
<th>Meek</th>
<th>The goose-<em>NOM</em> the wolf-<em>ACC</em> meeked. (SOV +)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>The donkeys meeked the sheep. (SVO -)</td>
</tr>
<tr>
<td></td>
<td>The panda-<em>ACC</em> the dinosaur-<em>NOM</em> meeked. (OSV +)</td>
</tr>
<tr>
<td>Tam</td>
<td>The moose-<em>NOM</em> the bear-<em>ACC</em> tammed. (SOV +)</td>
</tr>
<tr>
<td></td>
<td>The pigs tammed the cocks. (SVO-)</td>
</tr>
<tr>
<td></td>
<td>The ram-<em>ACC</em> the mouse-<em>NOM</em> tammed. (OSV +)</td>
</tr>
<tr>
<td>Weef</td>
<td>The hedgehog-<em>NOM</em> the bird-<em>ACC</em> weefed. (SOV +)</td>
</tr>
<tr>
<td></td>
<td>The horses weefed the cows. (SVO -)</td>
</tr>
<tr>
<td></td>
<td>The zebra-<em>ACC</em> the duck-<em>NOM</em> weefed. (OSV +)</td>
</tr>
</tbody>
</table>

### DAY 2

<table>
<thead>
<tr>
<th>Meek</th>
<th>The goose-<em>NOM</em> meeked the wolf-<em>ACC</em>. (SVO +)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>The sheep the donkeys meeked. (SOV -)</td>
</tr>
<tr>
<td></td>
<td>The dinosaur-<em>ACC</em> meeked the panda-<em>NOM</em>. (OVS +)</td>
</tr>
<tr>
<td>Tam</td>
<td>The moose-<em>NOM</em> tammed the bear-<em>ACC</em>. (SVO +)</td>
</tr>
<tr>
<td></td>
<td>The cocks the pigs tammed. (SOV -)</td>
</tr>
<tr>
<td></td>
<td>The mouse-<em>ACC</em> tammed the ram-<em>NOM</em>. (OVS +)</td>
</tr>
<tr>
<td>Weef</td>
<td>The hedgehog-<em>NOM</em> weefed the bird-<em>ACC</em>. (SVO +)</td>
</tr>
<tr>
<td></td>
<td>The cows the horses weefed. (SOV -)</td>
</tr>
<tr>
<td></td>
<td>The duck-<em>ACC</em> weefed the zebra-<em>NOM</em>. (OVS +)</td>
</tr>
</tbody>
</table>
APPENDIX C. Information sheet for parents for Studies 1 and 2.

School of Psychological Sciences
Participant Information Sheet

Title of project: Productivity of a Finnish child's inflectional noun morphology: a naturalistic study

Introduction

Young children acquiring Finnish are learning how words are inflected and how inflection changes the meaning of words (e.g. a dog vs. dogs). Despite the fact that child language has been an interest of researchers for several decades, the precise acquisition mechanism is still not fully understood. Previous research suggests that the process of word inflection acquisition is affected by the frequency of different word forms the child hears in her parents’ speech. This study investigates the relationship between different noun forms in both the child’s and the parent’s speech.

What will my child and I be asked to do if we take part?

This study is a naturalistic study, which involves collecting data by audio-recordings of parent-child interaction. You will be provided with an MP3 audio-recorder to document you and your child’s interaction in your home. You will also be asked to send the audio files to the investigator via e-mail on a weekly basis. The duration of the study is two sets of six (6) weeks with a break of approximately four weeks between the time periods. You are required to do five (5) hours of recordings in sessions of one hour every week during the periods of data collection. Each one hour recording should be done on separate days so that no more than one hour will be recorded during one day.

Will our data be confidential?4

One aim of the study is to provide the child language research community with a publicly accessible dataset of child Finnish and child-directed Finnish in the CHILDES (The Child Language Data Exchange System) database

4 Nb - within the bounds of any requirements pertaining to our duty of care to children (HM Government, 2006)
This will allow researchers to address a wide range of research questions relating to the acquisition of Finnish. The recordings will be transcribed by the investigator, and only she and other researchers will have access to the recordings. Identifying information such as your address or date of birth which might be mentioned during the recordings will not be transcribed (but will be indicated with ‘www’ in the transcript). Any identifying information such as your or your child’s name will be changed in the transcriptions of the audio recordings and these pseudonyms will be used in research publications and/or published in CHILDES. Consent forms will be held, under responsibility of the principal investigator for 10 years following completion of the study for publishing, teaching and research presentation purposes and will be stored in a lockable file and then destroyed by a shredder.

Do we have to take part?
You or your child will not have to take part. If you decide to take part and then change your mind (whether before starting the study, half way through it or even after it has finished) you can withdraw without giving your reasons, and, if you wish, your data will be destroyed. However, once the data is donated to CHILDES it can no longer be destroyed as it may have been downloaded by someone else.

Benefits and risks
There are no direct benefits to you or to your child for participating in this study. However, both children and parents generally indicate their enjoyment of taking part in such studies, and are generally enthusiastic about taking part in research on child development. The parent will be compensated by a payment of £10 for each one hour recording. There are no direct risks involved in taking part in the study either. If you or your child is reluctant to take part, you have the right to withdraw at any point of the study.

If there are any issues regarding this research that you would prefer not to discuss with members of the research team, please contact the Research Practice and Governance Co-ordinator by either writing to 'The Research Practice and Governance Co-ordinator, Research Office, Christie Building, The University of Manchester, Oxford Road, Manchester M13 9PL', by emailing: Research-Governance@manchester.ac.uk, or by telephoning +44 (0)161 275 7583 or 275 8093.

Where can I obtain further information if I need it?
If you require any further details, please contact the principal investigator, Henna Lemetyinen, by email: henna.lemetyinen@postgrad.manchester.ac.uk. Alternatively, you may also contact the supervisor of the project, Dr Anna Theakston, by email (in English): anna.theakston@manchester.ac.uk.
APPENDIX D. Information sheet for parents for Study 3.

School of Psychological Sciences

Parent Information Sheet

Title of project: Understanding children’s comprehension of case marking and word order

Introduction

Young children acquiring Finnish are learning how different grammatical cues such as case marking and word order interact in a sentence. For example, in Finnish it is possible to change the order of the words in a sentence but have the sentence still mean the same thing by using grammatical devices such as case marking to indicate who does what to whom. In this study we are investigating how children learn to understand different kinds of sentences.

What will my child be asked to do if s/he takes part?

If your child takes part s/he will be invited to play a game with the investigator. The game uses simple pointing and act-out methodologies to assess what children understand about sentences. Your child will sit down with the investigator by a table. The investigator will show the child a few simple actions with toy animals, asking the child to imitate the actions with the toys. Next, the child will be shown two short videos displayed side by side on a computer screen, whilst hearing a sentence matching one of the videos on the screen. The sentences will contain new, made-up words (e.g. The panda meeked the sheep), which the child will have been familiarised with in the act-out. Finally, the child will be asked to point to one of two still pictures taken from the previous video clips, according to which one matches the sentence they heard. The procedure will be repeated with different sentences and videos. After the task, the child’s knowledge of how nouns change to indicate things like plurality, e.g. dog – dogs will be tested with a standardised test. The investigator will spend approximately 45 minutes with your child.

Will my child’s data be confidential?

Yes, only the researchers will have access to the data. If your child takes part s/he will be video-recorded for data coding purposes. Any identifying information such as the name of your child will be deleted from the video-recording immediately. The recordings will only
be identified by your child’s date of birth. Consent forms and data will be held under responsibility of School of Psychological Sciences for 5 years following completion of the study for publishing purposes and will be stored in a lockable file and then destroyed by a shredder.

**Does my child have to take part?**
You do not have to allow your child to take part in the study. If you decide to allow your child to take part and then change your mind (whether before the child starts the study, half way through it or even after it has finished) you can withdraw without giving your reasons, and, if you wish, your child’s data will be destroyed. Your child will be asked whether s/he is happy to play a game with the researcher. If your child does not want to participate, s/he will not be included, even though you have given your consent for your child to participate.

**Benefits and risks**
There are no direct benefits to your child participating in this study. However, both children and parents generally indicate their enjoyment for taking part in such studies, and are generally enthusiastic about participating in research on child development. Similarly, there are no direct risks to taking part in the study either. If your child demonstrates any form of reluctance s/he will not be required to take part. Children who show any signs of distress will also not be required to take part.

**Where can I obtain further information if I need it?**
If you require any further details, please contact the principal investigator, Henna Lemetyinen, by email: henna.lemetyinen@postgrad.manchester.ac.uk.

If there are any issues regarding this research that you would prefer not to discuss with members of the research team, please contact the Research Practice and Governance Coordinator by either writing to (in English) ’The Research Practice and Governance Coordinator, Research Office, Christie Building, The University of Manchester, Oxford Road, Manchester M13 9PL, UK, by emailing: Research-Governance@manchester.ac.uk, or by telephoning +44 161 275 7583 or 275 8093.