INCONSISTENT INPUT AND AMOUNT OF EXPOSURE:

CHILD LANGUAGE ACQUISITION OF FERING, A DIALECT OF NORTH FRISIAN

A Dissertation in
German with Applied Linguistics Option and Language Science

by
Alison Eisel Hendricks

Doctor of Philosophy

December 2014
The dissertation of Alison Eisel Hendricks was reviewed and approved* by the following:

Carrie N. Jackson  
Associate Professor of German and Linguistics  
Dissertation Advisor  
Co-Chair of Committee

Karen Miller  
Assistant Professor of Spanish Linguistics  
Dissertation Advisor  
Co-Chair of Committee

B. Richard Page  
Associate Professor of German and Linguistics  
Head of the Department of Department

Michael Putnam  
Associate Professor of German and Linguistics

*Signatures are on file in the Graduate School
ABSTRACT

This dissertation presents six comprehension and production experiments designed to test how the amount and type of language input influence child language acquisition in a naturalistic minority language. This project focuses on the acquisition of Fering, a dialect of North Frisian spoken on the island of Föhr in northern Germany. Acquisition of a consistently produced feature, plural marking, and an inconsistently produced feature, grammatical gender, was compared in groups of bilinguals with different amounts of language input. Elementary school students (ages 6-10), high school students (ages 16-19) and adults on Föhr completed a series of production and comprehension tasks examining their acquisition of Fering morphology. The results indicate that both the amount of language input and the type of input (consistent vs. inconsistent) influence the acquisition of morphology. Specifically, when children are exposed to inconsistent input, their production patterns depend on how much linguistic input they have access to. Children with high levels of input were able to faithfully acquire adult-like patterns of inconsistency, while children with roughly equal exposure to each language at home tended to regularize inconsistent input. Those who had a minimal amount of Frisian input tended to produce grammatical gender randomly. For consistently produced plural markings, elementary school students with larger quantities of input were more accurate in producing and comprehending plural markings. The results of this dissertation suggest that both the amount and type of language input strongly influence how children acquire their languages.
**TABLE OF CONTENTS**

List of Figures ........................................................................................................ vii
List of Tables ................................................................................................ .......... ix
Acknowledgements ................................................................................................. xii
Chapter 1. Introduction ............................................................................................ 1
Chapter 2. Previous Literature on Acquisition ...................................................... 6
  2.1 Introduction ....................................................................................................... 6
    2.1.1 Models of Acquisition ............................................................................... 7
  2.2 Acquisition of morphology .............................................................................. 11
    2.2.1 Grammatical gender ................................................................................. 12
    2.2.2 Plural marking .......................................................................................... 18
  2.3 Type of Input .................................................................................................... 22
    2.3.1 Variable Input .......................................................................................... 23
    2.3.2 Inconsistent Input ..................................................................................... 25
    2.3.3 Discussion and Summary .......................................................................... 31
  2.4 Quantity of Input ............................................................................................. 32
    2.4.1 Balanced bilingual acquisition ................................................................ 33
    2.4.2 Unbalanced bilingual acquisition: Minority languages ......................... 36
    2.4.3 Discussion and Summary .......................................................................... 43
  2.5 Discussion and Summary ................................................................................ 43
Chapter 3. Description of Fering ........................................................................... 46
  3.1 Introduction ....................................................................................................... 46
  3.2 Bilingualism on Föhr ...................................................................................... 48
    3.2.1 Languages spoken on Föhr ...................................................................... 48
    3.2.2 Status of Fering ...................................................................................... 49
    3.2.3 Child Bilingualism on Föhr ...................................................................... 50
  3.3 Fering grammatical gender .............................................................................. 51
    3.3.1 Fering grammatical gender system .......................................................... 52
    3.3.2 Inconsistency in Fering Grammatical Gender ............................................ 58
    3.3.3 Adult Survey ........................................................................................... 64
  3.4 Fering plural marking ....................................................................................... 71
    3.4.1 Previous Literature .................................................................................. 71
    3.4.2 Corpus Study ........................................................................................... 73
    3.4.3 Adult Survey ........................................................................................... 74
  3.5 Summary and Discussion ................................................................................ 78
Chapter 4. Results: Grammatical Gender ............................................................... 80
  4.1 Introduction ....................................................................................................... 80
    4.1.1 Fering Grammatical Gender .................................................................... 82
    4.1.2 Participants .............................................................................................. 83
    4.1.3 General Materials .................................................................................... 85
    4.1.4 General Procedure ................................................................................... 86
  4.2 Definite article elicitation task ......................................................................... 88
    4.2.1 Background ............................................................................................. 88
    4.2.2 Methods and design ................................................................................ 89
    4.2.3 Results .................................................................................................... 92
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2.4</td>
<td>Discussion</td>
<td>121</td>
</tr>
<tr>
<td>4.2.5</td>
<td>Summary</td>
<td>123</td>
</tr>
<tr>
<td>4.3</td>
<td>Pronoun Production Task</td>
<td>124</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Background</td>
<td>124</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Methods and design</td>
<td>125</td>
</tr>
<tr>
<td>4.3.3</td>
<td>Results</td>
<td>126</td>
</tr>
<tr>
<td>4.3.4</td>
<td>Discussion and Summary</td>
<td>134</td>
</tr>
<tr>
<td>4.4</td>
<td>Pronoun Comprehension Task</td>
<td>136</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Background</td>
<td>136</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Methods and design</td>
<td>137</td>
</tr>
<tr>
<td>4.4.3</td>
<td>Results</td>
<td>140</td>
</tr>
<tr>
<td>4.4.4</td>
<td>Discussion</td>
<td>146</td>
</tr>
<tr>
<td>4.4.5</td>
<td>Summary</td>
<td>147</td>
</tr>
<tr>
<td>4.5</td>
<td>Summary and Discussion</td>
<td>147</td>
</tr>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>151</td>
</tr>
<tr>
<td>5.1.1</td>
<td>Plural Marking in Fering</td>
<td>153</td>
</tr>
<tr>
<td>5.1.2</td>
<td>General Participants</td>
<td>154</td>
</tr>
<tr>
<td>5.1.3</td>
<td>General Materials</td>
<td>155</td>
</tr>
<tr>
<td>5.1.4</td>
<td>General Procedure</td>
<td>157</td>
</tr>
<tr>
<td>5.2</td>
<td>Plural elicitation task</td>
<td>159</td>
</tr>
<tr>
<td>5.2.1</td>
<td>Background</td>
<td>159</td>
</tr>
<tr>
<td>5.2.2</td>
<td>Participants</td>
<td>160</td>
</tr>
<tr>
<td>5.2.3</td>
<td>Methods and design</td>
<td>160</td>
</tr>
<tr>
<td>5.2.4</td>
<td>Results</td>
<td>162</td>
</tr>
<tr>
<td>5.2.5</td>
<td>Discussion</td>
<td>182</td>
</tr>
<tr>
<td>5.2.6</td>
<td>Summary</td>
<td>185</td>
</tr>
<tr>
<td>5.3</td>
<td>Acceptability Task</td>
<td>186</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Background</td>
<td>186</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Participants</td>
<td>187</td>
</tr>
<tr>
<td>5.3.3</td>
<td>Methods and design</td>
<td>187</td>
</tr>
<tr>
<td>5.3.4</td>
<td>Results</td>
<td>189</td>
</tr>
<tr>
<td>5.3.5</td>
<td>Discussion</td>
<td>200</td>
</tr>
<tr>
<td>5.3.6</td>
<td>Summary</td>
<td>202</td>
</tr>
<tr>
<td>5.4</td>
<td>Summary and Discussion</td>
<td>203</td>
</tr>
<tr>
<td>6.1</td>
<td>Introduction</td>
<td>207</td>
</tr>
<tr>
<td>6.1.1</td>
<td>Background</td>
<td>208</td>
</tr>
<tr>
<td>6.1.2</td>
<td>Participants</td>
<td>209</td>
</tr>
<tr>
<td>6.1.3</td>
<td>Materials and Procedure</td>
<td>211</td>
</tr>
<tr>
<td>6.2</td>
<td>Perceived Fluency</td>
<td>212</td>
</tr>
<tr>
<td>6.2.1</td>
<td>Background</td>
<td>212</td>
</tr>
<tr>
<td>6.2.2</td>
<td>Analysis</td>
<td>213</td>
</tr>
<tr>
<td>6.2.3</td>
<td>Results</td>
<td>214</td>
</tr>
<tr>
<td>6.2.4</td>
<td>Discussion and Summary</td>
<td>217</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 2-1. Wug-test (Berko, 1958) ................................................................. 19

Figure 4-1. Stimulus for the Gender Production Task ........................................ 90

Figure 4-2. Grammatical gender vocabulary accuracy: Home language comparison ..... 97

Figure 4-3. Grammatical gender production: Home language comparison .............. 105

Figure 4-4. Grammatical gender production: Age Group Comparison .................. 110

Figure 4-5. Scatterplots of masculine responses compared to Adult responses by item 119

Figure 4-6. Scatterplots of masculine responses compared to Adult responses by item for correct word response .................................................................................. 120

Figure 4-7. Stimulus for Pronoun Production Task ........................................... 126

Figure 4-8. Stimulus for Gender Pronoun Comprehension Task .......................... 139

Figure 5-1. Sample elicitated production item .................................................... 161

Figure 5-2. Plural vocabulary accuracy: Home language comparison .................. 169

Figure 5-3. Plural accuracy: Age group comparison ........................................... 178

Figure 5-4. Proportion of Plural Markings for Frisian Plural Responses ............... 182

Figure 5-5. Acceptability Scale ........................................................................ 188

Figure 5-6. Proportion of responses by age and home language ......................... 194

Figure 6-1. Pictures for Story-telling Task .......................................................... 211

Figure 6-2. Scale for native speaker ratings of story-telling task .......................... 214

Figure B-1. Example item for definite article production task ............................. 264

Figure B-2. Example item for gender pronoun production task ............................ 265

Figure B-3. Example item for grammatical gender comprehension task .............. 266

Figure C-1. Example item for plural marking production task ............................ 270
Figure C-2. Example item for plural marking acceptability task ……………………………271
LIST OF TABLES

Table 3-1. Grammatical Gender and Plural Marking in Fering ........................................ 53
Table 3-2. Pronoun Agreement in Contemporary Fering .................................................. 56
Table 3-3. Historical Gender System in Fering ................................................................. 56
Table 3-4. Categories of Items for Gender Assignment Task ........................................... 65
Table 3-5. Participant Responses by Grammatical Gender .............................................. 68
Table 3-6. Regular Plural Marking in Fering (Wiltz, 2011) ............................................. 72
Table 3-7. Irregular Plural Marking in Fering (Wiltz, 2011) ............................................. 72
Table 3-8. Categories of Items for Plural Assignment Task .............................................. 75
Table 4-1. Fering grammatical gender system ................................................................. 82
Table 4-2. Example stimuli for grammatical gender tasks ............................................... 85
Table 4-3. Vocabulary accuracy by Home Language and Category ................................ 94
Table 4-4. Sub-set of vocabulary accuracy by Home Language and Category ................. 95
Table 4-5. Vocabulary accuracy by Age Group and Category ........................................ 98
Table 4-6. Vocabulary accuracy by Age Group and Category for Matched Participants 99
Table 4-7. Proportion of Masculine gender responses for correct vocabulary responses by Home Language and Category ................................................................. 103
Table 4-8. Sub-set of gender production by Home Language and Category .................... 104
Table 4-9. Masculine gender responses by Age Group and Category ............................. 107
Table 4-10. Sub-set of gender responses by Age Group and Category ............................ 108
Table 4-11. Regularized Speakers by Home Language Group ........................................ 114
Table 4-12. Regularized Speakers by Age Group ......................................................... 115
Table 4-13. Form accuracy by Home Language and Category ........................................ 128
Table 4-14. Form accuracy by Age Group and Category ............................................. 129
Table 4-15. Pronoun production accuracy by Home Language and Category ............ 131
Table 4-16. Pronoun production accuracy by Age Group and Category .................. 132
Table 4-17. Pronoun production frequency of responses by Home Language .......... 133
Table 4-18. Pronoun production frequency of responses by Age Group ................. 134
Table 4-19. Pronoun comprehension accuracy by Category and Home Language ..... 142
Table 4-20. Pronoun comprehension accuracy by Category and Age Group .......... 142
Table 4-21. Pronoun comprehension improvement by Home Language ................ 143
Table 4-22. Pronoun comprehension improvement by Age Group ......................... 144
Table 4-23. Gender comprehension practice items by Home Language ................ 145
Table 4-24. Gender comprehension practice items by Age Group ......................... 145
Table 5-1. Regular Plural Marking in Fering (Wilts, 2011) ........................................ 153
Table 5-2. Level of adult consensus on plural production ........................................ 164
Table 5-3. Vocabulary accuracy by Home Language and Category ....................... 166
Table 5-4. Sub-set of vocabulary accuracy by Home Language and Category .......... 167
Table 5-5. Vocabulary accuracy by Age Group and Category ............................... 170
Table 5-6. Sub-set of vocabulary accuracy by Age Group and Category .................. 171
Table 5-7. Plural accuracy by Home Language and Category ................................ 173
Table 5-8. Subset of plural accuracy by Home Language and Category ................... 174
Table 5-9. Plural accuracy by Age Group and Category ........................................ 176
Table 5-10. Sub-set of plural accuracy by Age Group and Category ..................... 177
Table 5-11. Acceptability accuracy by Home Language and Category .................... 190
Table 5-12. Acceptability accuracy by Age Group and Category .......................... 191
Table 5-13. Average ratings for double negative filler items ........................................... 194
Table 5-14. Difference scores by Home Language and Category ................................. 197
Table 5-15. Difference scores by Age Group and Category ..................................... 199
Table 6-1. Native speaker ratings of story-telling task by Home Language ............. 215
Table 6-2. Utterance fluency by Home Language ......................................................... 221
Table 6-3. Utterance fluency by Age Group ................................................................. 222
Table 6-4. All Types, Tokens, and TTR by Home Language ........................................ 227
Table 6-5. Frisian Types, Tokens, and TTR by Home Language ............................... 228
Table 6-6. Types, Tokens, and TTR by Age Group ..................................................... 229
Table 6-7. Only Frisian Types, Tokens, and TTR by Age Group ............................... 230
Table B-1. Stimuli for grammatical gender tasks ......................................................... 262
Table A-1. Responses to adult survey for grammatical gender marking task ............ 257
Table A-2. Responses to adult survey for plural marking task .................................. 259
Table B-1. Stimuli for grammatical gender tasks ......................................................... 262
Table C-1. Stimuli for plural marking tasks ................................................................. 267
ACKNOWLEDGEMENTS

There are many people I would like to thank for helping me to develop and carry out the research presented in this dissertation. First, I would like to thank my advisors, Carrie Jackson and Karen Miller. Without their help I could not have completed this project. Particular thanks to Carrie, who has challenged and encouraged me throughout my graduate studies at Penn State. From when she called me to let me know that I was accepted into the program, to our recent weekly skype meetings, Carrie has been a constant source of practical advice and insightful discussions. The addition of Karen as a co-advisor was a tremendous help, and I am indebted to her for her expertise in child language acquisition. I am grateful for the extensive time and care Carrie and Karen invested in this project, and I am deeply thankful for their extensive knowledge and mentorship throughout my graduate studies. Thank you also to my committee members, to Richard Page for his insight into historical linguistics and to Mike Putnam for his guidance on problems related to Low German.

I would like to recognize the many individuals and groups that helped this project along on Föhr and in Kiel. In particular, Dr. Volkert Faltings, Uta Marienfeld, and Reinhard Jannen at the Ferring Stiftung on Föhr. Personal thanks to Uta who always made sure I had a place to stay in the cozy apartment above the Ferring Stiftung when I was visiting Föhr. Thank you also to Drs. Jarich Hoekstra and Alastair Walker from the University of Kiel for their feedback and support on this project. This dissertation would also not have been possible without the help of several research assistants, including Elke Nickels, Göntje Christiansen and Göntje Erichsen. Thank you also to Joslyn Johnson for
the playful pictures she drew for this project. Special thanks to Elke for traveling with me all across Föhr many times over while we collected data.

Importantly, I am grateful for the support from the teachers and principals of the Grundschule Föhr-West in Süderende, Grundschule Föhr-Ost in Midlum, and the Eilun Feer Skuul in Wyk on Föhr. Thank you to Erk Roeloffs whose high school Frisian classes are a hopeful reminder of the next generation’s excitement about Frisian on Föhr. Particular thanks to Enken Tholund who not only helped me in organizing data collection in her classes, but also taught my Fering course in the winter evenings. This research would not have been possible without the participation and enthusiasm of many community members and families on Föhr. I was grateful to be welcomed into your homes and to hear your stories. Such hospitality was truly touching.

To carry out this research, I had to live on Föhr and in Kiel for a substantial time, and pay research assistants to help with data collection and analysis. This would have been impossible without the financial support from the following grants: Fulbright Fellowship to Germany, NSF Doctoral Improvement Grant #1251989, and an Early Career Development Award from the Center for Global Studies at Penn State.

Finally, I would like to thank my family for their support throughout my graduate studies. I especially appreciate my mama and papa who taught me the importance of education through modeling a deep commitment to learning. I am thankful for their encouragement and perspective which made this process more enjoyable. Thank you also to Brad, whom I made the mistake of teaching just enough linguistics to be annoying. Thank you for your constant support and annoyingly on point questions, even during the short, winter days on Föhr.
Chapter 1. Introduction

Language acquisition research examines how children learn their language based on the input they receive. Investigating differences in the type and amount of input children receive and the impact these factors have on acquisition is important to understanding how children acquire language. While a number of studies have investigated how the quantity of language input influences language acquisition (e.g., David & Li, 2008; Gathercole, 1997), fewer studies have investigated differences in the type of language input and its impact on language acquisition (Hudson Kam & Newport, 2005, 2009; Miller, 2007). In this dissertation, I examine how children acquire one type of input: inconsistent input. Specifically, I investigate how inconsistent input and the amount of input interact in language acquisition through an investigation of the acquisition of morphology in Fering, a minority language spoken on the island of Föhr in northern Germany.

Inconsistent input is produced probabilistically and unpredictably, and can refer to inconsistent alternations between two or more forms, or between the presence or absence of a single form. To contrast consistent input with inconsistent input, consider the case of consistent indefinite article selection in English. In English, a child hears the following forms: *an apple, a banana, a pear,* and *an onion.* The child learns that “an” always precedes words beginning with a vowel while “a” always precedes words that begin with consonants. This pattern is categorical and relatively easily acquired. In contrast, consider the case of inconsistent grammatical gender in Fering. The child might hear the following forms: *de reiluk ‘the.MASC curtain’, det bleed ‘the.COM newspaper’, de bleed ‘the.MASC*
newspaper’, *det reiluk* ‘the. Com curtain’. From this input, it is difficult to determine which gender is associated with each noun, as there are no linguistic or extralinguistic cues to gender assignment. This is the task facing children learning grammatical gender in Fering.

While there have been studies on how the amount of language input influences acquisition, much less is known about the effect of the type of input, and how these two properties of language interact. To date, there are very few studies on the acquisition of inconsistent input. Those studies that do exist are based on case studies (Singleton & Newport, 2004) or artificial languages with short exposure times (Austin, 2010; Hudson Kam & Newport, 2005, 2009). The results of these studies suggest that children tend to regularize morphology when they are exposed to inconsistent input.

These initial studies on the acquisition of inconsistent input leave several questions unanswered. How do children acquire inconsistent morphology when they have access to high levels of language input? Will children regularize inconsistent input when they have access to high levels of input, and will they be able faithfully reproduce the inconsistency in their input? If children are able to acquire inconsistent morphology, how much input is necessary? To date, there have been no larger scale investigations of inconsistent input in natural languages, which would help to address these questions. This is the gap that this dissertation seeks to fill.

By focusing on a minority language, the current project also has implications for minority language communities. While intergenerational transmission—passing a language from generation to generation—is essential for language vitality, little is known about how much language input is needed to successfully acquire a language or how to
best support children learning minority languages. The results from this project can help communities understand the challenges facing children learning a minority language. Further, I investigate whether the amount of input needed for acquisition differs across features. Language planners can use this information to ensure that language programs adequately meet the needs of children learning their minority language (Grenoble & Whaley, 2006).

Importantly, this dissertation adds to the body of literature on minority language acquisition in that it uses experimental tasks to investigate language acquisition. Much of the previous literature on minority language acquisition has used descriptive methods to examine general patterns of intergenerational transmission (Grenoble & Whaley, 2006; Hinton & Hale, 2001). Recently, Peter and Hirrtta-Edds (2008) called for more experimental studies that can evaluate the learning outcomes of minority language programs. In this project, I explore how experimental studies can complement descriptive studies in evaluating the outcomes of language programs.

Beyond minority language contexts, studying how inconsistent input influences acquisition can also be applied to cases where children receive potentially inconsistent input from parents who are themselves second language learners. This includes parents who learned a minority language later in life, as in cases of language revitalization. Successful language revitalization results in new second language learners who, hopefully, pass their new language on to their children (see Robert, 2009, for an example in Wales). This dissertation will provide much needed additional information about how children who receive inconsistent input acquire language and about the amount of input needed to successfully acquire their language.
This dissertation addresses the following broad research questions:

1. How do children acquire grammatical morphology that is produced inconsistently in the input?

2. Within the context of inconsistent input, what role does the amount of exposure play?

Specifically, I examine the acquisition of morphology in Fering as a test case for investigating the role of the quantity of input in the acquisition of inconsistent morphology. Adult native speakers produce grammatical gender inconsistently, while plural markings are produced consistently. That is, adults largely agree on the plural markings for nouns, but there is no general consensus in Fering as to the gender of any given noun. This results in inconsistent input for gender markings but consistent input for plural markings. Further, children on Föhr have different amounts of exposure to Frisian, resulting in natural sub-groups of bilinguals. To examine the role of quantity of language input, I compare the results of grammatical gender, plural marking, and general proficiency tasks across speakers with different amounts of language input. Together, Fering provides an exceptional test case for how the quantity of input and the type of input interact in language acquisition. To this end, this project addresses the following specific research questions:

1. Do Fering children exposed to inconsistent input for gender marking produce gender marking inconsistently in their own speech or do they initially regularize gender marking?

2. How does the acquisition of gender marking (produced inconsistently) compare to the acquisition of plural marking (produced consistently)?
3. Do children exposed to Fering both in school and at home show more adult like patterns of inconsistency in their production of grammatical gender than children only exposed to Fering at school?

This dissertation is organized as follows: Chapter 2 discusses the previous research on child language acquisition, specifically focusing on the acquisition of grammatical gender and plural markings. I also discuss the relevant previous literature on how the quantity of input and input type each influence language acquisition. Chapter 3 provides a description of Fering grammatical gender and presents the results of an adult online survey that documents the use of morphology in contemporary Fering. Chapter 4 presents the results of three experiments on the acquisition of Fering grammatical gender: a definite article production task, a gender-marked pronoun production task, and a gender-marked pronoun comprehension task. Chapter 5 presents the results of two experiments on the acquisition of Fering plural marking: a plural marking production task and a plural marking acceptability task. Chapter 6 presents the results of a story-telling task in order to examine Fering acquisition more generally. Chapter 7 summarizes and discusses the results of chapters 4, 5, and 6, and returns to the theoretical and applied importance of this dissertation. In chapter 7 I discuss the questions that remain unanswered and discuss the future of research in child language acquisition.
Chapter 2. Previous Literature on Acquisition

2.1 Introduction

It is not surprising that the linguistic input children hear influences how they acquire language. Children who hear English, acquire English and children who hear German, acquire German. Behind this simple idea, lies a more complex story. The exact relationship between language input and language acquisition is still unknown. Research has shown that both the type (Hudson Kam & Newport, 2005; 2009, Miller 2013) and the quantity of language input (David & Li, 2008; Gathercole, 1997; Gathercole & Thomas, 2009; Thomas & Gathercole, 2007) influence the speed of acquisition. However, few studies to date have investigated how the type of input and the quantity of input influence acquisition in the same study.

Research on acquisition of different types of input has recently received attention. Most of this research focuses on the acquisition of variable input, that is, variation that is sociolinguistically constrained. This dissertation focuses on a different type of input – inconsistent input. Inconsistent input is unreliable input which is neither linguistically nor sociolinguistically constrained. Recent studies by Hudson Kam and Newport (2005, 2009) suggest that adults and children deal with inconsistent input differently. While adults replicate statistical patterns in inconsistent input, children tend to regularize inconsistent input. However, these studies fail to account for how the quantity of language input influences acquisition.

Research on bilingual acquisition, particularly minority language acquisition, documents the importance of quantity of input in language acquisition. Bilingual children
split their time between two or more languages, meaning that they receive a portion of their input from each language. As such, comparisons of bilingual children with different ratios between their languages allow for testing the role of quantity of input. While some research has studied bilingual children to investigate the amount of input, this research fails to account for how the type of input influences acquisition. This dissertation provides a more complete investigation of language acquisition by simultaneously exploring the quantity of input and the type of input.

This chapter is organized as follows: I begin by summarizing two models of language acquisition in Section 2.1.1. In Section 2.2, I outline the expected timeline of acquisition of grammatical gender and plural markings and discuss several factors which can alter the timeline of acquisition. In Section 2.3, I present research on how the type of language input influences morphological acquisition, while in Section 2.4 I present studies about how the amount of language input influences language acquisition. While there has been research both on how the type of language input and the amount of language input influence acquisition, this dissertation investigates the intersection of input type and input quantity. Section 2.5 discusses and summarizes the chapter.

2.1.1 Models of Acquisition

Although the interpretation of the results of this dissertation does not rely on any specific model of acquisition, I will briefly summarize two models of child language acquisition here. Much of research on child language acquisition follows one of two approaches: generative approaches (Legate & Yang, 200; Yang, 2004; Yang, 1999) and usage-based approaches (Lieven & Tomsello, 2008; Tomesello, 2003; see also Bybee,
2007; Goldberg, 2013). In this section I will describe one model of language acquisition from each of these approaches and discuss the predictions that they make for the acquisition of Fering.

Within the generativist framework, the Variational Model put forth by Yang (1999, 2002) is one of the few language acquisition models that takes language variation into account (see also Roeper, 1999). According to this model, children have access to a universal set of features through Universal Grammar (UG), and use the linguistic input available to them to determine which of these features their language encodes. The Variational Model proposes that children construct possible grammars from UG, and that each of these grammars has a probability associated with it. This probability represents the chance that the grammar that the child has constructed matches the adult target grammar. As the child encounters linguistic input, she must decide whether it is compatible with the possible grammars. When the input is consistent with the grammar she has constructed, the grammar is rewarded, and the probability that that grammar represents the adult grammar increases. Conversely, when a new sentence is inconsistent with the grammar she has constructed, then the grammar is punished, and the probability that it represents the adult grammar decreases. Acquisition takes place when the child converges on a grammar, and the probability approaches 1. The amount of input needed to acquire a feature in this account will differ based on the reliability of the input. When children receive unambiguous input, they will require less input to converge on the correct grammar, as the input would favor only one grammar. When children receive unreliable input, such as variable or inconsistent input, the child may require more input.
to converge on the correct grammar, as the input will support several competing grammars.

Focusing on the current project, the Variational Model makes two predictions. First, it predicts that when a child determines that a feature is present in her language, it should be applied to all items equally. Consequently, children should regularize their input from the beginning of acquisition. Perez-Lerouz, Pirvulesu and Roberge (2009) call this a “structural default” meaning that prior to determining how the parameters of her language are set, the child will assume a default setting, until there is sufficient evidence in favor of or against the parameter. When there is sufficient evidence, a parameter will be set, and all words should be equally affected.

Perez-Lerouz, Pirvulesu and Roberge (2009) further suggest that bilinguals may show “default retention”, meaning that because they do not receive as much input in their languages, they may maintain the default parameter setting longer than monolingual children. This can be seen through an extended period of regularization before the correct parameter setting appears.

Applying the Variational Model to the current project makes a further prediction, namely, that inconsistent input will require more time and input to acquire. That is, because inconsistent input provides evidence for conflicting grammars, children will need more time and consequently more input in order to acquire inconsistent morphology. In the case of Fering, this means that children will acquire consistently produced plural markings before inconsistently produced grammatical gender. Further, children with more language input will be able to acquire inconsistent morphology more quickly than those with less Frisian input.
While the Variational Model can be useful in framing this project, the application is limited in that it is not concerned with how children determine which marking is associated with a given noun. Specifically, the Variational Model is concerned with how children determine whether their language marks grammatical gender, and less with whether, for example, *fask* ‘fish’ is associated with masculine or common gender. In this way, the Variational Model does not make strong predictions about how children will acquire a lexically conditioned feature, such as grammatical gender in Fering.

This dissertation can also be situated within usage-based models (Tomasello, 2003). Within usage-based approaches, language acquisition is considered the result of domain-general pattern finding, rather than selecting parameters from Universal Grammar. Usage-based models argue that children use domain general skills of pattern finding to first learn items on a word-by-word basis, and then derive more general patterns. Children are powerful statistical learners and are sensitive to frequency patterns in their input. Because these theories believe that language emerges as a means of communication and to serve social functions, they argue that grammar is secondary and is constructed after initially learning in a piece-meal fashion.

There are several models of acquisition within the usage-based framework. Research within the usage-based framework typically uses frequency effects to support their claims. That is, if language is learned first word by word, rather than via general parameter setting, this predicts that learners should acquire more frequent forms more quickly than less frequent forms. Some have studied token frequency, or how often individual words occur in the input (Tomasello, 2003). Others have studied the frequency of type of forms or constructions, which is how often a pattern or sequence occurs in the
input (Goldberg, 2013; Goldberg & Casenhiser, 2008, see Bybee, 2007 for an overview of type and token frequency effects).

Turning to the current project, usage-based models predict that children would learn both grammatical gender and plural markings item by item, and would be influenced by how often a specific form occurs. After children have amassed enough input to derive a more general pattern, they would regularize the forms. This means that children may require more input to learn inconsistent forms than consistent forms, as it may be more difficult to find the underlying pattern in the presence of inconsistent input. Importantly, usage-based models may be able to account for how children learn the correspondences between morphological markings and specific nouns. Further, usage-based models may prove useful in explaining how children are able to faithfully acquire a lexically determined system, such as Fering grammatical gender.

While the variational model and the usage-based models are different at their foundation, both of these approaches can offer useful insights. This dissertation was not designed to test either of these models, but I will return to each of them when I discuss the results in Chapter 7.

2.2 Acquisition of morphology

The empirical domain of this dissertation is the acquisition of morphology in Fering, a dialect of North Frisian. Specifically, I investigate how children acquire grammatical gender and plural marking in Fering. Before I turn to previous research on how the type of language input and the quantity of language input influence acquisition, I will present previous studies on the acquisition of grammatical gender and plural
markings. While I will not present a full cross-linguistic survey, I discuss the timeline of acquisition of morphology across several languages and show how the complexity of the morphological system influences the timeline of acquisition. This section provides a foundation comparison for the results of Frisian acquisition presented in Chapters 4 and 5.

2.2.1 Grammatical gender

The task of grammatical gender acquisition is two-fold. The child must first determine that her target language marks grammatical gender, and then learn which grammatical gender is associated with each noun (Corbett, 1991). The present research focuses on the latter of these two issues, and takes as a starting point that children have determined that their language marks grammatical gender. Acquiring grammatical gender marking requires the child to attend to articles or other cues to grammatical gender information, and track to the co-occurrences of grammatical gender markings on nouns. For example, in Fering, the child must notice that fask ‘fish’ is often associated with the masculine definite article de. By following these co-occurrences, the child can learn which gender is associated with each noun.

Recent experimental research has shown that young toddlers are sensitive to and make use of grammatical gender information. Children begin to show sensitivity to grammatical gender information around 24 months. For example, French learning toddlers orient towards pictures faster during a preferential looking task when they are preceded by a definite article which is correctly marked for grammatical gender rather than an incorrect determiners, showing that by about 25 months, toddlers are using
grammatical gender information (van Heugten & Shi, 2009). Cyr and Shi (2013) investigated grammatical gender knowledge in toddlers of different ages to determine when they begin to use grammatical gender information. They familiarized toddlers (14 months, 17 months, 20 months, and 30 months) to pseudo-nouns with French grammatical gender information, and then tested whether they reacted to grammatical gender violations. They found that toddlers at 20 months first begin to use grammatical gender information during a preferential looking task, and that the effect was robust by 30 months.

Even when children are not given grammatical gender information in the prompt, toddlers are aware of the grammatical gender of familiar nouns. Bobb and Mani (2013) showed that 24-month-old German-learning toddlers oriented faster to target nouns when the two pictures have different gender. Lew-Williams and Fernald (2007) showed similar results using an eye-tracking paradigm in older Spanish-learning children, ages 34-42 months.

Children begin producing grammatical gender markings similarly early in development, although the exact age of acquisition differs across languages. In languages where there are phonological cues to grammatical gender, and where there is no overlap with other features, such as case markings, grammatical gender is mastered early, often before age three (see Berman, 1985 for Hebrew; Karmiloff-Smith, 1979 for French; Perez-Pereira, 1991 for Spanish). In contrast, in languages where gender marking is opaque with few phonological cues to gender assignment and when gender overlaps with case marking, grammatical gender acquisition can take much longer, (see Henzl, 1975 for Czech; Mills, 1986 for German; Popova, 1973; Zakharova, 1973 for Russian).
Grammatical gender is acquired early in languages where the system is independent from other morpho-syntactic categories, such as case marking, and when there are clear phonological and morphological cues to gender assignment. For example, Berman (1985) used diary reports and free speech language samples from Hebrew to show that children make few errors in gender agreement by the time they are three years old. These children did, however, continue to make gender agreement errors into the fourth and fifth years when the marking is more complex and there are conflicting cues to grammatical gender, such as a conflict between the plural marking and gender marking.

In French, where grammatical gender is transparent, children acquire gender early. Karmiloff-Smith (1979) used a set of experiments to test whether children learning French used phonological suffixes to assign gender marking to novel nouns. In one experiment, children saw two drawings of the same novel object in different colors. The experimenter then performed an action on one of the objects (e.g., placed an object on the drawing, or rotated it) and the child was asked what the experimenter did, as seen in the example in (1). In this way, the prompt did not provide any grammatical gender information, other than the noun suffix, and prompted the child to produce a grammatical gender marking on the definite determiner and the color adjective.

(1) Experimenter | Child’s expected response
--- | ---
*Voici deux fasines. Qu’est-ce que c’est?* | *Deux fasines*
‘Here are two fasines. What is here?’ | Two fasines
*Qu’est-ce que j’ai fait?* | *Vous avez caché la fasine blanche*
‘What did I do?’ | ‘You hid the white fasine’
The results showed that already by 3;0 children provided the expected gender based on the ending of the noun, but that when the ending is not informative, the children favored the masculine gender. This shows that in addition to knowing that their language marks grammatical gender, the children in this study used phonological cues to derive the correspondences between grammatical gender and nouns.

Similarly, children also acquire grammatical gender in Spanish early. Using a similar method as Karmiloff-Smith (1979), Perez-Pereira (1991) replicated these results in Spanish. In this study, 160 children ages 4;4 -11;6 were shown cartoon drawings of novel objects. The semantic, morphological, and syntactic cues to grammatical gender were manipulated. The semantic cues to grammatical gender were given through novel characters that looked either female or male. Morphological cues were given through typical feminine suffixes (e.g., loden-a) or typically masculine suffixes (e.g., linol-o), and syntactic cues were given through gender-marked indefinite articles (e.g., una lodena vs. un linolo). The results showed that children have mastered gender marking in Spanish before 4;0, as all children were able to use grammatical gender markings. Moreover, children attend more to morphological and syntactic cues to grammatical gender than the semantic cues. However, more recently Larrañaga and Guijarro-Fuentes (2012) argued on the basis of free speech samples that children attend to semantic cues over morphological cues. Regardless of whether children attend to morphological, syntactic, or semantic cues, these two studies indicate that young children are able to derive the grammatical gender of nouns from their input, and use this information when producing gender marking.
In contrast to French and Spanish, acquisition tends to be slightly later in languages where grammatical gender overlaps with case-marking and there are few phonological cues to grammatical gender. For example, grammatical gender is acquired slightly later in German, where it overlaps with case, but has fairly robust phonological and semantic patterns. Children first begin producing grammatical gender in German between 1;5 and 1;8, and the frequency of errors falls below 10% by about age three (Szagun, Stumper, Sondag, & Franik, 2007). Grammatical gender assignment in German is more complex than French or Spanish, but may be aided by several factors including phonological regularities, derivational morphology, and semantic patterns (e.g., Köpcke, 1984).

Similarly, grammatical gender is acquired later in Russian (Popova, 1973; Zakharova, 1973). Interestingly, despite the apparent similarities between the Polish and Russian grammatical gender systems, Polish is acquired earlier than Russian. Smoczyńska (1985, p. 645) suggests that two minor differences between Polish and Russian contribute to the later acquisition in Russian. First, in Russian, unstressed –o, which is a marker for neuter gender, is pronounced like –a, which is a feminine gender marker. Secondly, diminutive masculine forms that end in –a are declined like feminine forms. This means that common diminutive names for boys, such as Zhenya, Sasha, and Kolya, are declined like feminine nouns. Smoczyńska (1985) suggests that these minor differences lead to overall less reliable gender markings in Russian compared to Polish, and results in later acquisition.

In Dutch, several factors conspire to make the Dutch grammatical gender system very complex, and as a consequence, Dutch grammatical gender is acquired fairly late.
Dutch has a two-gender system, where nouns are marked as either common gender or
neuter gender. The grammatical gender system is opaque in that it lacks phonological
cues to grammatical gender. Further, common gender words are much more frequent in
Dutch than neuter words, which leads to a skewed distribution between genders.
Consequently, while children first produce gender marked determiners in Dutch at a
similar age as children learning French, Spanish, or German, they initially rely on
common gender, and studies have shown children continue to make grammatical gender
errors on neuter words until they are between six and ten years old (Blom, Polišenská, &
Unsworth, 2008; Cornips & Hulk, 2008).

Overall, the research on grammatical gender acquisition shows that children are
sensitive to grammatical gender very early in infancy, produce grammatical gender
markings fairly early, and they master their language’s grammatical gender system
around two or three years of age. While grammatical gender is acquired fairly early
across most languages, the exact timeline of complete acquisition, and the exact noun-to-
gender correspondences, depends on factors such as the transparency of gender marking,
the prevalence of phonological cues to grammatical gender, and whether the gender
system interacts with case marking.

From these studies, it is difficult to predict the expected timeline of gender
acquisition in Fering. Grammatical gender does not overlap with case markings in Fering,
and there are two genders in Fering, which suggests that it would be acquired early, like
in French. However, there are no phonological cues to Fering grammatical gender, which
suggests that it would be acquired relatively late, like in German or Dutch. Further, the
studies reviewed here are concerned with whether children acquire the grammatical
gender feature, rather than whether they have acquired which gender is associated with each noun. As such, they do not make specific predictions about how children learn which grammatical gender is associated with which noun.

2.2.2 Plural marking

Similar to grammatical gender acquisition, plural marking acquisition requires children to determine that their language marks plural, and then to determine the correspondences between plural markings and nouns (Corbett, 2000). Recent research has shown that children may be aware of plural marking early, even before they are able to productively produce plural markings (Kouider, Halberda, Wood, & Carey, 2006; Wood, Kouider, & Carey, 2009; Zapf & Smith, 2007; 2009). As with grammatical gender marking, this dissertation is primarily concerned with the acquisition of specific plural marking to noun associations, rather than how children associate plural markings with a “more than one” interpretation. If a language has multiple plural markings, the child must determine which plural marking is connected with each noun. For example, a child learning German must learn to associate the plural marking, -s, with the noun Auto-s ‘car-s’, and the plural marking, -en, with the noun Katze-n ‘cat-s’. This dissertation investigates how children learn which plural marking is associated with a given noun, rather than how children learn that plural marking indicates “more than one”.

Much of the early research on plural marking acquisition was conducted on English, which has a fairly simple plural marking system. The English plural marking system does not overlap with other features, such as grammatical gender or case marking, there are few plural markings (e.g., /-z/ vs. /-s/) and there are clear phonological cues to
plural marking. Early longitudinal and cross-sectional studies on language acquisition documented that plural marking is the first bound morpheme to be acquired in English (Brown, 1973). Plural markings first appear around 18-20 months, and children produce plural markings consistently and accurately by about age two (Brown, 1973; Cazden, 1968; de Villiers & de Villiers, 1973; Mervis & Johnson, 1991).

Figure 2-1. Wug-test (Berko, 1958)

Experimental tasks have shown similar results. In a landmark study, Berko (1958) showed that English speaking children, as young as four, acquire the rules governing plural marking selection using novel nouns. This study is particularly useful for the current investigation because it addresses the selection of plural markings, rather than the acquisition of the category of plural.

Children were shown drawings of real and novel words, followed by two of the same objects, and were asked to complete sentences, as seen in Figure 2-1. The children
in this study were able to use all three plural markings with real nouns, and productively used the [s] and [z], but not the [əz] plural markings, for non-sense words. Thus, children were able to determine not only which plural markings often occur with real nouns, but were also able to extract abstract patterns about the plural marking from their input to use on novel nouns.

More recently, Zapf and Smith (2003, 2007) used a type of Wug-test methodology with younger children in order to test when children begin to use the phonological rules to plural marking in English. They familiarized young children (ages 17.5 months to 28 months) with both common and novel nouns over six weekly sessions, and found that children produce correct plural markings as early as the first session. Over the six week period, the proportion of correctly marked plurals on novel words and real words increased. These results show that children are able to correctly generalize some plural markings for novel nouns long before they use plural marking rules fully productively in English.

Pérez-Pereira (1989) showed that Spanish plural markings are acquired similarly early using a similar task. In this study 130 children, ages three to six, completed a Berko-type sentence completion task for both real words and artificial words, as seen in (2).

(2) *Esto es un gobo. Hora hemos puesto otro más. Ahora hay dos _____.*

‘This is a balloon. Now there’s another. Now there are two _____’

The results show that Spanish speaking children produce the –s plural marking over 92% of the time for both real words, such as *naranja* ‘orange’, and artificial words, such as *lando*. The accuracy of –es plural marking for real words, such as *flor* ‘flower’ was 76%
(for three year-olds) and 98% - 100% for children ages four to six. However, correct production of –es for artificial words, such as astor, was much lower. Three year olds were 36% accurate and six year olds were 55% accurate. These results are similar to the results from Berko (1958) for English, which show that while some plural markings are generalized later in acquisition, plural markings are acquired early when the plural system contains clear phonological cues and does not overlap with other features, such as case.

In German, when there is a larger set of plural markings, children acquire plural markings slightly later than in English or Spanish. Szagun (2001) used spontaneous speech samples from 21 children to demonstrate that children learning German produce plural markings slightly later than in English, around 1;6. Although there was not a significant decrease in their errors between ages 2;10 and 3;8, the types of errors they produce, and the overall frequency of their plural markings becomes more adult-like overtime. At age 2;10, children differ significantly from adult speakers in the frequency of plural type production, with children producing infrequent plural markings more often than adult speakers. By age 3;8, however, children produce plural markings with adult-like frequency distributions. Overall, the results suggest that plural markings are acquired slightly later in German, where the plural system includes a large inventory of plural markings and an overlap between case and plural marking, as compared to other languages with less complex plural systems.

Berman (1985) provides further evidence that the complexity of the plural marking system influences the timeline of acquisition. While the acquisition of plural marking occurs around age three in Hebrew, specific plural markings are acquired at
different ages, depending on their complexity. The plural marking for masculine nouns, -\textit{im}, is acquired first, since this marking indicates only plurality and attaches to unmarked masculine words. The --\textit{ot} plural marking for feminine nouns combines both the feminine marking and the plural marking and is thus acquired later. Thus, even within one language, evidence from Hebrew shows that children acquire simpler plural markings forms more quickly than complex plural markings.

To summarize: the acquisition of plural marking entails both mapping of plural on to a “more than one” interpretation, and deriving the plural marking to noun correspondences. That is, once the child has determined that her language marks plural, she must decide which plural marking is associated with each noun. This is the focus of this dissertation. Cross-linguistically, in languages with more complex plural marking systems, that is, many plural markings and an interaction with other morpho-syntactic features, such as case marking or grammatical gender, plural marking tends to be acquired later than in languages with simpler plural marking systems. From these studies I predict that Fering plural marking will be acquired fairly early. The Fering gender system uses predominantly two plural markings and plural markings do not also mark grammatical gender.

2.3 Type of Input

Until fairly recently, studies on language acquisition have assumed that language input is categorical and systematic. However, several studies have focused on how children acquire different types of input, as a means of learning about language acquisition more generally. These studies are based on the following logic. If all typically
developing children acquire language in the same way, then investigating how language outcomes change when children have different types of language input would shed light on the process of language acquisition more generally. In this section I present the previous research on two different types of input: variable input, and inconsistent input. Variable input refers to predicable alternations in language which are constrained by sociolinguistic factors, and is prevalent in all natural languages. In contrast, inconsistent input is rare in natural languages, and is characterized by unpredictable and probabilistic patterns.

2.3.1 Variable Input

Recent research has investigated variability as one way the type of language input can influence language acquisition. Variable input refers to ambiguous language input which is probabilistic and sociolinguistically constrained. That is, social factors, such as class, gender, and age, change the probability of whether a feature will surface or not. This body of research can be separated into two areas: first, studies have investigated how children master these sociolinguistic constraints, that is, how they learn to use features following adult-like rules (Foulkes, Docherty & Watt, 2005; Miller, 2013; Roberts, 1994; Smith, Durham, & Fortune, 2007; 2009). Fewer studies have investigated how variable input affects children’s acquisition of morphology. Those studies that do suggest that until the child has learned the sociolinguistic patterns, the input may initially appear inconsistent (Johnson, 2005; Miller 2007, 2012; Miller & Schmitt 2012).

Miller (2007) compared the acquisition of plural marking across two dialects of Spanish using comprehension and production tasks. Plural marking is produced
categorically in Mexican Spanish, while it is produced variably in Chilean Spanish. In Chilean Spanish, a final /-s/ lenition process leads to three realizations of plural marking: [s], [h], or [ø]. The variation is sociolinguistically constrained, and /-s/ lenition is more prevalent among working class speakers compared to middle-class speakers. These studies compared the production and comprehension of plural markings across three groups of children ages four-five: working-class Chilean children, whose input is most variable, middle-class Chilean children, whose input is less variable than working-class Chilean children, and Mexican children, whose input is not variable. The Mexican children produce plural markings categorically, while working-class and middle-class Chilean children produce variable plural markings. Further, the middle-class children’s production differs across tasks, indicating that they are sensitive to register differences, which suggests they have two grammars in competition. These results show that variable input allows children to maintain competing grammars for a longer period of time than children with categorical input. These studies are primarily concerned with how children learn to associate plural markings with a “more than one” interpretation, rather than focusing on the selection of a specific plural marking.

These recent studies show that comprehension is delayed in cases where children receive variable input which is sociolinguistically constrained. These delays in comprehension suggest that children require longer to converge on the adult-like grammar. While little is known about how children acquire variable input, considerably less is known about how children are influenced by inconsistent input, which is not constrained by sociolinguistic factors.
2.3.2 *Inconsistent Input*

Unlike variable input, inconsistent input is not predictable from linguistic or extra-linguistic features. Since inconsistent input is arguably less reliable than variable input, it is possible that children will require additional input in order to acquire inconsistent input than variable input. That is, features which are used inconsistently are not generally used more often by certain speakers or in certain circumstances. Because inconsistency is rare in natural languages, studies of the acquisition of inconsistency generally fall into two types: studies on the acquisition of input from L2 speakers, and studies using artificial languages. These studies are of particular interest, because this dissertation tests the acquisition of inconsistent input in a natural language context. In this section I discuss studies on the acquisition of inconsistency using signed languages and artificial languages.

*Acquisition of L2 input*

It is rare that a child learns her language exclusively from L2 speakers. One exception to this is the case of Simon, reported in Singelton and Newport (2004). Simon was a seven year old boy who was profoundly deaf, and only had input from his parents who learned American Sign Language (ASL) as teenagers. Because his parents learned ASL later in life, they produced grammatical morphemes inconsistently, meaning that their patterns of use were unpredictable from either linguistic or extra-linguistic variables. Thus, the input Simon received was inconsistent and included errors which were characteristic of non-native signers. Singelton and Newport (2004) reported that, despite his inconsistent input, Simon performed within the normal native ranges for most forms.
Simon did not perform within native norms for morphological markers conveyed via hand shape, and Singleton and Newport argued that the family was using a non-standard sign and Simon had partly regularized this non-standard sign. Singelton and Newport described the results, saying that Simon “boosts the frequency”, that is, he regularized the input available to him.

While this study is one of the few studies that looks specifically at how children learn languages when they are given inconsistent input, there are several limitations which make it difficult to generalize the results. First, the study, while broad in depth, is limited by the fact that it only investigated one child’s acquisition. From these data it is not possible to determine whether Simon’s pattern of acquisition is typical of children acquiring inconsistent input, or whether the pattern is unique to him. Secondly, Simon’s input contained errors that were typical of second language learners, and therefore it is possible that the results are specific to the acquisition of language when given input with errors, rather than reflecting the acquisition of inconsistent input more generally. As inconsistent input is rare in natural languages, this is one of the broad challenges of studying inconsistent input. Finally, Simon only had access to input from two signers. Thus, while he had a large quantity of input, in terms of hours of input, his input was limited in diversity. It is possible, then, that he regularizes inconsistent input because he had not yet received a diverse enough sample to be able to determine what adult-like inconsistency looks like.

There are several differences between the Simon study and this dissertation. First, the children in this study are slightly younger than Simon, and the children in this study are all bilingual, while Simon was a monolingual ASL signer. Importantly, one key
difference between this dissertation from Singleton and Newport (2004) is that it tests the acquisition of inconsistent input in a larger sample, the participants have had input from a diverse group of adult native speakers. Thus, the current dissertation will be able to test whether the Simon’s pattern of results is typical of language acquisition with inconsistent input.

**Acquisition of artificial languages**

Given that inconsistency is uncommon in natural languages, the limited research on acquisition with inconsistent input has relied on artificial languages. For example, Hudson Kam and Newport (2005, 2009) designed a series of studies using a mini-artificial language to test how children and adults acquire morphology from inconsistent input. In their first experiment, focusing on adult acquisition, Hudson Kam and Newport (2005) taught participants a mini-artificial language with 51 words, including 2 determiners. Groups of adults had input from one of eight versions of the language. The languages differed only in their use of determiners. Half of the versions included semantically meaningful determiners, reflecting a mass/count distinction, and half contained purely grammatical determiners, similar to those used with grammatical gender. Within each of these types of languages, there were four sub-groups of languages, which differed in the frequency that determiners occurred. Determiners were present 45%, 60%, 75%, or 100% of the time. For example, in the 45% condition, 45% of the time a noun occurred, a determiner occurred along with it, and 55% of the time the determiner was absent. For the 75% condition, 75% of the time a noun occurred, a determiner occurred with it. However, there were no linguistic or extra-linguistic cues as
to whether a determiner would occur with the noun or not. Participants were exposed to the language for a short period of time (a total of about three hours over six sessions). After the six exposure sessions, participants completed a series of experimental tasks to test their acquisition of consistently and inconsistently produced grammatical structures and vocabulary.

The results show that adults, regardless of the input language, were able to learn the vocabulary and consistently produce grammatical structures in the language. Participants who had input from different versions of the language did not differ in accuracy on an overall test of the grammar of the language. However, both in a forced-choice comprehension task and a sentence completion task, the adults’ determiner responses closely parallel the input they were exposed to. That is, adults whose input contained determiners 45% of the time also produced determiners about 45% of the time, while those whose input contained determiners 60% of the time produced determiners about 60% of the time. Importantly, reflecting their input, adults used determiners inconsistently, meaning that they did not use determiners with either all of the nouns, or none of the nouns. Further, 31 out of 37 participants did not adopt novel rules, such as using determiners for mass nouns, but not count nouns. The type of inconsistency (semantic versus grammatical) did not affect how the adults in this study learned the inconsistent determiners.

In contrast to adults, the results of a second study show that children did not match the input they were exposed to, but produced determiners in a more systematic way than available to them in their input. In the second study presented in Hudson Kam and Newport (2005), children (average age 6:4 to 10) learned a simplified version of the
mini-artificial language presented in the adult study. This mini-version contained 17 words, including one determiner. As with the adult study, children were exposed to languages with differing degrees of consistency in definite article production. There was no semantic difference between the conditions (i.e., mass/count vs grammatical gender). In the consistent condition, the determiner appeared 100% of the time, while in the inconsistent condition, they appeared 60% of the time and were absent 40% of the time. Within the inconsistent condition, whether the determiner appeared or not was not conditioned by any linguistic or extra-linguistic cues. Children were exposed to the language for a short period of time (between one and two hours over six sessions).

Like the adults in the first study, children in both the consistent and inconsistent conditions were able to learn the consistently produced features in the language, as seen in their high accuracy on the general grammar test. Unlike the adults, children in the inconsistent condition produced determiners systematically. That is, even when children were exposed to inconsistent input, meaning determiners were present 60% of the time, they produced regularized output. There were three types of regularized production. Systematic users were participants who always used the determiner. Systematic nonusers never used the determiner. Systematic other included participants who used determiners in other systematic ways, such as always with subject noun phrases, but never with object noun phrases. The results showed that roughly the same percentage of children in the inconsistent and consistent conditions were systematic users. This means that speakers were either producing the determiner for all of the nouns, none of the nouns, or according to a novel rule, such as only with subjects. Hudson Kam and Newport (2009) finds similar results when the inconsistency comes from alternations between forms of the
determiner, rather than the presence versus absence of the determiner (Hudson Kam & Newport, 2009).

One limitation of these studies is that in order to use an artificial language, the children are by necessity L2 learners. The children were on average six years old when they were exposed to the artificial language. Thus, it is unclear whether children who are exposed to inconsistent input from birth would also regularize inconsistent input. Hudson Kam and Newport (2005, 2009) argue that their results show that child language acquisition drives languages change, but the results may be limited to early second language acquisition. It is necessary to test these results in a natural language context in order to determine their validity.

Importantly, the limited exposure time also makes it difficult to generalize these results to natural language contexts. Adults and children were each exposed to this mini-language for a very short period of time, adults for six hours total, and children for between one and two hours. It is possible, then, that with more exposure, children will approximate their input and match the frequency of input they were exposed to. While participants seemed to learn the language generally, as shown by vocabulary tests and general grammar tests, and previous studies have shown that miniature artificial languages can be learned by young children with short exposure times (Moeser & Olson, 1974), it is unclear whether these findings are the result of the early stages of acquisition, or a more general result.

Perfors (2012) suggested that the difference between adults and children’s performance seen in Hudson Kam and Newport (2005, 2009) stems from how children and adults approach the task. Adults’ previous language learning experience leads them
to believe there is an underlying pattern within the inconsistency, and their attempt to find the pattern results in probability matching. In contrast, children feel less pressure to perform well on the task and adopt a regularization strategy to best meet the task demands. This is another reason why an investigation of inconsistent input in a natural language context may add to the current literature on language acquisition.

To date, there have been no studies that I am aware of that test these results given larger quantities of input or using a natural language context. Chapters 4 and 5 of this dissertation address this gap in the literature by testing these hypotheses with a natural language, where children have varying degrees of language exposure.

2.3.3 Discussion and Summary

While there are studies on the acquisition of variable input, research on inconsistent input is rare. Research on inconsistent input from Singelton and Newport (2004) on the acquisition of ASL, and the experiments using artificial languages in Hudson Kam and Newport (2005, 2009) form notable exceptions to this general trend. The results of that series of studies show that adults mirror the statistical patterns in the input they are exposed to, but that children regularize their input. However, these studies are limited by their short exposure times, and it is difficult to determine whether their pattern of results will hold when children are exposed to the language for a longer period of time. This dissertation tests the acquisition of inconsistent input in a natural language context, in order to determine whether these results are generalizable.
2.4 Quantity of Input

It is a truism that children need some language input in order to acquire their language. However, the question of how much input is needed is less obvious. Children are well equipped to acquire two or more languages, even when their input is split between these languages. However, research on minority language acquisition finds that in some cases, children do not successfully acquire the minority language. With this in mind, one line of research investigates how much language input is necessary for children to successfully acquire their languages.

Before discussing the previous literature on the quantity of input, it is important to distinguish frequency from quantity of input. In this dissertation I focus on the quantity of input rather than input frequency. The quantity of input refers to the amount of time children receive language input. That is, quantity of input can be thought of in terms of hours per day. The frequency of a feature refers to how often the feature occurs in the language. For example, present perfect markings are more frequent in American English compared to Scottish English, meaning that they occur more often. Both the American English speaking child and the Scottish English speaking child may receive the same quantity of language input, but the frequency of this feature is different across the two languages (Gathercole, 1986). Usage-based theories of language acquisition often refer to the frequency of features, and argue that more frequent structure are acquired more quickly than less frequent structures. This is based on the logic that children have access to more instances of frequent structures than less frequent structures and are able to derive the general pattern more quickly for these structures. In contrast, I focus on the amount of input in terms of the amount of time that children hear the language.
There is also a growing body of research that investigates how the amount of input influences how children learn languages that are not endangered. For example, research by Sharon Unsworth and colleagues discusses how the type of input and amount of input influence the acquisition of Dutch in the Netherlands by bilingual children (Unsworth, 2005; 2008; 2013). These studies have also focused on amount of input in terms of the amount of time the child hears the target language. For the purposes of this dissertation, I focus on the literature as it relates to the acquisition of endangered minority languages, as some studies have suggested that endangered minority languages are uniquely affected by the interaction of amount of input and input type (Gathercole & Thomas, 2009; Thomas & Gathercole, 2007; Thomas, Williams, Jones, Davies & Binks, 2013).

The effects of quantity of language input have primarily been studied in bilingual acquisition. In this section I will describe acquisition in cases of balanced bilingualism and discuss the research on minority language acquisition as a special case of unbalanced bilingualism. Overall, the research suggests that the quantity of input interacts with other factors, and should be taken into account when investigating the acquisition of inconsistent input.

2.4.1 Balanced bilingual acquisition

Bilingual children provide an interesting opportunity to explore the role of the quantity of input. Bilingual children, by definition, have less input in each of their languages than their monolingual peers (David & Li, 2008; Yip & Mathews, 2007), because they split their time between two (or more) languages and the hours in a day are
finite. The research on bilingual language acquisition is inconclusive as to whether the
difference in the quantity of input influences the timeline of acquisition.

Some studies have shown that bilingual children acquire their languages slower
relative to monolingual children. Gathercole (1997) finds that Spanish-English bilingual
children acquire mass-count noun distinctions later than monolingual children. A task
with novel words tested children’s ability to infer mass/count noun classifications from
syntactic input showed that at age 7, the monolingual children were able to use syntactic
information to infer whether novel words were mass or count nouns, while the bilingual
children could not. However, by age 9, bilingual children with the highest proficiency in
the L2 paralleled monolingual children, while children with lower L2 proficiency
continued to lag behind. The initial differences in bilingual development disappear by
mid-to late childhood, with children who have more balanced exposure in both languages
catching up to monolingual peers more quickly than those with unbalanced exposure
(Gathercole, 1997).

Similarly, Nicoladia, Song, and Marentette (2012) find that bilingual children
acquire English past tense markings later than monolingual children. In this study,
children ages 5-11 recounted stories from short clips from Pink Panther films. All of the
children in the study were simultaneous bilinguals living in Canada, and heard either
French or Chinese from at least one parent at home. The authors argue that French-
English and Chinese-English bilinguals acquire English past-tense markings similar to
monolingual English children, but that bilinguals fully acquire the feature later than
monolingual peers.
Several studies have found mixed results with respect to vocabulary acquisition. Poulin-Dubois, Bialystock, Blaye, Polonia and Yott (2012) found that at 24-months old, there was a difference in the expressive vocabulary between bilingual and monolingual toddlers, but no difference in terms of receptive vocabulary. Bialystok, Luk, Peets, and Yang (2010) reported that school age bilinguals had a smaller vocabulary than monolinguals, but that was limited vocabulary connected to topics associated with home. In this study the bilinguals and monolinguals had similar sized vocabularies for school-related topics. These studies show the mixed results for vocabulary size between bilinguals and monolinguals.

David and Li (2008) used a more fine grain comparison to show the influence of the quantity of input on bilingual acquisition. They find a positive correlation between vocabulary size and the ratio of language exposure in French-English and English-French bilinguals. For both groups of bilingual’s the more exposure the child had to French, the larger her French vocabulary size. More specifically, differences between the quantity of input from month to month also affects vocabulary size. Moreover, the study finds that children reflect their parents’ language use in that there is a correlation between the Mean Length Utterance (MLU) used by parents in child directed speech and the child’s own MLU.

Other studies have shown no differences in terms of when features are acquired between bilingual children and monolingual children. Rodina and Westergaard (2013) conducted a corpus analysis of the grammatical gender acquisition by two monolingual Norwegian speakers, and two Norwegian-English bilinguals. They find no differences between the acquisition by bilinguals and monolinguals. They argue that the opaque
gender marking system in Norwegian leads to overall late acquisition of grammatical gender for bilingual and monolingual children alike, and as a result, bilingual children face no additional delays relative to monolingual peers.

Because bilingual children must split their time between two or more languages, they provide an interesting example for testing the role of the quantity of language input. Together these studies underscore the importance of taking the amount of language input into account when investigating bilingual language acquisition.

2.4.2 Unbalanced bilingual acquisition: Minority languages

While bilingual children with roughly balanced language input typically acquire both of their languages successfully, children acquiring minority languages may not have sufficient input to acquire both of their languages. The acquisition of minority languages provides an opportunity for investigating how much input is required for successful acquisition, and which features are particularly sensitive to the quantity of input a child receives. Much of the research on minority languages has focused on intergenerational transmission as a measure of language acquisition. These studies explore which factors are important in passing a language on to the next generation.

**Descriptive Studies on Minority Language Transmission**

Many studies on minority languages have been primarily interested in broad language outcomes and which features contribute to successful intergenerational transmission (De Houwer, 2007; Jones, 2012; Morris & Jones, 2007; Nolan, 2008; Steensen, 2003). These studies, although broad, can provide information about the quantity of input needed to acquire a minority language.
Morris and Jones (2007) investigates the patterns of generational transmission of 12 Welsh families. They find the language used by parents at home to be an important predictor for language transmission. However, it is not easy to predict whether a child will be socialized in Welsh or in English, as many factors, including the language of the maternal grandmother, the language of each of the parents, the language of daycare and future decisions about the language of education influence the chances that the child will be socialized in Welsh. The results of this in-depth qualitative study add to the field of minority language acquisition and they further complicate the picture of intergenerational transmission by showing that a complex set of factors influence intergenerational transmission. It is important to keep these factors in mind when assessing the input available to the child learner. However, Morris and Jones (2007) did not assess the relative importance of each of these factors for child language development.

Steensen (2003) is one of the few studies on North Frisian language transmission. This study examined transmission patterns and attitudes towards Frisian language instruction in a small school in Fahreoft, near Dagebüll in mainland North Friesland. The Hans-Momsen-Schule in Fahreoft uses the “Patenschaftsmodell”, where children meet with a Frisian native speaker (or near-native speaker) of the grandparent generation and practice and learn Frisian. Steensen (2003)’s questionnaire data show that different sources of language input, including a range of language partners, classroom language lessons, and the language spoken at home, can contribute to minority language acquisition.

While these studies highlight the importance of intergenerational language transmission, they are limited by the fact that they rely entirely on descriptive analysis or...
self-reported data. Peter and Hirata-Edds (2008) and Gorter (2008; 2012) highlight that it is best to incorporate both assessments and questionnaires to best understand how the quantity of input influences the transmission of a minority language.

**Experimental Studies on Minority Language Acquisition**

Few studies have investigated minority language acquisition experimentally. The research on Welsh and Irish are two important exceptions to this trend. Both of these experimental studies emphasize the importance of the quantity of language input in minority language acquisition. Virginia Gathercole and Enlli Thomas and colleagues have studied various aspects of Welsh acquisition and investigated the role that home language plays in the acquisition of a minority language (Gathercole & Thomas, 2007; 2009; Thomas et al., 2013). Although Welsh is considered one of the successes of language revitalization and there has been an overall rise in the number of Welsh speakers over the last decade, Welsh continues to be considered a minority language. Within Wales, English is the dominant language, and thus children acquiring Welsh may require more Welsh input than English input in order to fully acquire Welsh (Gathercole & Thomas, 2009).

Gathercole and Thomas (2009) and, more recently, Thomas et al. (2013) provide a comprehensive picture of research on Welsh acquisition and the importance of both the type of input and the amount of input. These studies demonstrate across a wide range of features, that the formal complexity of the feature interacts with the amount of linguistic input children receive (see Gathercole & Thomas, 2009 for vocabulary acquisition, grammatical gender, and word order effects, and Thomas et al. 2013 for plural marking).
Welsh contains a highly opaque system of grammatical gender, where gender is marked via phonological mutations of word initial consonants. For example, feminine singular nouns undergo soft mutation following definite articles: \( y \text{gath} \ [\text{gaθ}] \Rightarrow [\text{kaθ}] \) ‘the.FEM cat’ (Gathercole, Thomas, & Laporte, 2001). As not all words begin with a mutable consonant, it is not possible to mark every noun for gender, and there is considerable overlap between form and function. Previous studies indicate grammatical gender in Welsh is acquired very late, even among monolingual speakers (Gathercole et al., 2001).

Gathercole and Thomas (2009) used an experimental comprehension task, in which children heard pairs of sentences with pictures, and in the first sentence the noun was clearly marked for grammatical gender. In the second sentence, the children heard a possessive pronoun with either aspirate mutation, meaning the antecedent was feminine, or soft mutation, meaning the antecedent was masculine. Children saw two pictures and picked which of the pictures matched the second sentence. The target items included humans, animals, and inanimate objects.

Overall, there was a main effect of home language, indicating that the quantity of Welsh that children heard at home significantly affected their comprehension of grammatical gender markings. More specifically, children from Only Welsh Homes (OWH) used grammatical gender accurately for human items by age five, while children from Welsh-English Homes (WEH) were accurate by age seven, and children from Only English Homes (OEH) continued to struggle at age nine. For animal items, the OWH children were overall highly accurate by age nine, but upon closer investigation, this was limited to words with feminine antecedents. Both the WEH children and OEH children
continued to struggle with animal words at age nine. In terms of inanimate objects, all three groups had difficulties at age nine.

Gathercole and Thomas (2009) directly attributed these differences between home language groups to the overall amount of input available to the children. The data from Welsh indicate that acquisition of the minority language is disproportionally affected by input factors. Children acquire the majority language regardless of the home language, but acquisition of the minority language depends on access to input. Similar results have been found when the minority language is not endangered, as in Spanish in the United States (Pearson, Fernandez, Lewedeg, & Oller, 2008).

Similarly, Thomas et al. (2013) showed that children require a large amount of Welsh input to acquire the complex plural marking system. In this study, plural marking was testing in bilingual children (ages 7-11) with different amounts of Welsh input. While all children were bilingual, they differed in terms of how much exposure they have to Welsh. The results of a plural production task showed that children who hear only Welsh at home from birth produced plural markings with nearly adult-like accuracy. The children who hear some Welsh at home, and those who hear Welsh only in school lagged behind their higher-input peers. Thomas et al. (2013) argue that complex morphology requires a large quantity of input, even when it is produced consistently. The results of this study provide further evidence that the properties of the input, in this case opaque compared to transparent cues, and the amount of input both influence acquisition of minority languages.

Hickey (1997) reports on a large scale empirical study of the acquisition of Irish by children in Irish immersion preschools. The study investigated which external factors
were predictors of children’s Irish acquisition, including the factor of home language, and the relative importance of these factors. Tests were developed to measure children’s comprehension, production, and imitation of Irish. These tests were administered to a representative sample of children attending Irish immersion preschools in Ireland. The set of tests was administered to three and four year old children after one year of Irish immersion preschools.

The comprehension task consisted of a picture identification task and an act-out task. In the picture identification task the researcher asked the child a question, such as “where is the dog?” and the child pointed to one of three pictures. For four items in the picture identification task the experimenter described a picture and then presented another picture of the same item with the antonym descriptor, which elicited vocabulary, for example, “this is big, and this is _____?” Two additional items were used to elicit free speech descriptions of common events in the preschool. These were rated on a six-point scale of fluency and comprehensibility. In the act-out ask, children were asked to act-out common activities in the preschool, such as “close your eyes”. In the production task the researcher pointed to objects, pictures, or body parts and asked the child to name the object or the color of the object.

Some of the children came from homes where Irish was spoken daily, and some only began speaking Irish a year prior to starting preschool. Therefore, there was a large range of production skills among the children. The tasks were designed to measure the skills of both lower and higher proficiency children. Within the free speech task, the tester was instructed to ask the broad question, “What is in this picture?”, but if the child remained silent for over 10 seconds, the tester further prompted the child by pointing to
an object in the picture and asking, “What is he/she doing?” If the child continued to be silent, or spoke only in English, the tester pointed to items in the picture and ask, “Look at this- What is this?” In this way, the children were encouraged to perform at their highest level, but the task was not too difficult as to elicit no responses from some children. Beyond the language tasks, children also completed a series of general cognitive tests, which were conducted in their native language, either English or Irish. The general cognitive tasks were based on well-established cognitive tests for children, but were shortened so that the overall length of the session was appropriate for children.

For the comprehension task and production task, the children who received only Irish at home perform the best, then the children who heard both English and Irish at home, and finally those who only heard Irish in the preschool. The results of a regression analysis indicate that a number of factors were significant predictors of the children’s Irish skills. Child specific variables, such as age, and general cognitive ability were positively associated with Irish abilities. Within the home-level variables, there were positive correlations between the child’s Irish proficiency and parents’ Irish ability, the use of Irish as the home language, the frequency of use of Irish at home, and siblings attending a preschool or all Irish school. For the preschool-level variables, there was a significantly positive correlation between the school director’s competence in Irish, the location of the preschool (a private home, a public hall, or a school), the size of the preschool, the ratio between the number of teachers and students, and the location within Ireland (traditionally Irish speaking areas versus non-Irish speaking areas).

Taken together, these results suggest that there are many factors which influence children’s ability to learn minority language, and that both home-level variables, such as
the amount of input the child receives, as well as school-level variables, such as the location of the preschool, are important. More broadly, these results show that the quantity of language input contributes to children’s acquisition of minority languages.

2.4.3 Discussion and Summary

Research on balanced bilingual acquisition and unbalanced bilingual acquisition in the case of minority language acquisition shows that the quantity of language input influences how quickly children acquire their languages. In the case of minority languages, the importance of the quantity of input is magnified, and in some cases, children may not have enough access to minority language input to acquire the language at all. These studies suggest that minority language acquisition is a useful environment for studying the role of quantity of input in language acquisition. However, most of the studies on minority languages are descriptive in nature and focus on whether children use their language as a reflection of general language acquisition.

2.5 Discussion and Summary

This chapter provided a summary of the timeline of acquisition of grammatical gender and plural markings and used cross-linguistic evidence to show that children acquire more complex morphological systems later than simpler systems. After establishing a timeline of morphological acquisition, I summarized research about two features of input which affect acquisition of morphology. First, I discussed studies investigating how the type of input, specifically variable or inconsistent input, influence child language acquisition. Miller (2007) and Miller and Schmitt (2010) show that when children are given variable input they can allow multiple competing grammars to exist
longer before acquiring the adult-like grammar. Hudson Kam and Newport’s studies (2005, 2009) on inconsistent input suggest that while adults reproduce the statistical patterns in their input, children tend to regularize inconsistent input. However, their studies used small artificial languages with short exposure times, which make it difficult to determine whether their results show how children acquire inconsistent input in general, or represent the early stages of acquisition.

Then I summarized research concerning how the quantity of language input—in terms of hours of exposure per day—influences language acquisition. This section focused on balanced bilingual acquisition, and unbalanced bilingual acquisition in the case of minority languages. Minority languages are particularly useful in studies on the quantity of input, but are often descriptive and focus on language use as a reflection of general language skills.

Together this body of research shows that morphological acquisition depends heavily on language input—both type and quantity of input. However, there is insufficient research tackling both of these features in the same study. Studies on inconsistent input fail to account for quantity of input, while studies on the quantity of input fail to account for the type of language input. This dissertation addresses this gap by merging an investigation of inconsistent input with the quantity of input. In Chapters 4, 5, and 6, I present data from experimental tasks investigating how children with different amounts of language input acquire morphology in Fering, a minority language spoken in northern Germany. As grammatical gender marking is produced inconsistently and plural marking is produced consistently in Fering, I am able to test the type of input in addition to the quantity of input. In so doing, this dissertation explores the type and the quantity of
language input simultaneously, and thus provides a more complete picture of language acquisition overall.
Chapter 3. Description of Fering

3.1 Introduction

This dissertation investigates how two features of input, the quantity and type, influence the acquisition inconsistent morphology that is, probabilistic alternations between two forms that cannot be predicted by linguistic or extra-linguistic factors. This dissertation uses the acquisition of grammatical gender in Fering, a dialect of North Frisian spoken on the island of Föhr in northern Germany, as a test case for inconsistent morphology. This chapter provides a description of the grammatical gender and plural marking systems in Fering, and presents the results of an online-survey which establishes a base-line of adult production. These results will inform the experimental tasks of grammatical gender acquisition (see Chapter 4) and plural marking (see Chapter 5).

Fering, a dialect of North Frisian spoken on the island of Föhr in northern Germany, provides an excellent context for testing each of these two features in a natural environment. With regards to the amount of target language exposure, the historical bilingualism, recent language shift, and universal Fering instruction for children at rural elementary schools mean that children on Föhr have differing degrees of exposure to the target language. Turning to the type of input, grammatical gender and plural marking differ in the extent to which they are produced inconsistently in the input. Taken together, the linguistic context on Föhr allows for an investigation of how both the amount of input and inconsistency influence child language acquisition. Through a series of experiments testing grammatical gender and noun plural marking acquisition in Fering, a dialect of North Frisian, the dissertation tests the hypotheses put forth in Hudson Kam and Newport
(2005, 2009) in a natural environment. This chapter anchors the experimental tasks described in Chapters 4 and 5 by describing the linguistic situation on Föhr and presenting a description of grammatical gender and plural marking systems in Fering.

Fering, a dialect of North Frisian spoken on the island of Föhr in northern, and is one of ten mutually unintelligible dialects of the North Frisian language. North Frisian is considered an acutely endangered language (Bohn, 2004); however, Fering remains relatively strong compared to other North Frisian dialects (Århammar, 2007). With regard to the amount of target language exposure, a combination of circumstances has contributed to the current differentiation in children’s levels of Fering input. Föhr has a long history of bilingualism, with Fering spoken alongside German, Low German, and Danish. Recent changes, including a shift toward German, and universal Fering instruction for children at only rural elementary schools mean that children on Föhr have differing degrees of exposure to Fering.

This chapter is organized as follows: Section 3.2 describes the historically bilingual background leading to the current situation on Föhr in which speakers differ in terms of how much access they have to Fering. Sections 3.3 and 3.4 review the previous literature on Fering grammatical gender and plural marking respectively. This is important because it establishes the differences in the level of inconsistency between these two linguistic features. The natural differences between these two morphological features distinguish Fering as a prime environment for testing the acquisition of inconsistent input. Section 3.5 discusses how Fering grammatical gender and noun plural marking inform in the investigation of acquisition of inconsistent input more generally.
3.2 Bilingualism on Föhr

Föhr’s historical linguistic context makes it uniquely suited to testing the role of quantity of input in the acquisition of inconsistent input. A very brief summary of the historical setting on Föhr demonstrates how Föhr’s bilingual context, with children’s differing degrees of Fering input, emerged.

3.2.1 Languages spoken on Föhr

Archeological records suggest that Frisian has been spoken in North Friesland since as early as about 700 CE. Föhr was connected to the mainland until a massive storm, the first Grote Mandranke, separated the island from the mainland. Low German—a Germanic language spoken in northern Germany, and the lingua franca of the Hanseatic League from the 13th to 17th century—gained significant ground during the Hanseatic League period, due to increased maritime trade in the southern and eastern portions of Föhr. Low German continues to be spoken in some areas in southern and eastern Föhr.

Beginning around the turn of the 19th century, Föhr became well known as a tourist location, especially the beach towns in the East, such as the city of Wyk. Most of these tourists spoke German, and consequently the eastern portions of Föhr were heavily influenced by German. Following WWII, a group of immigrants came from mainland Germany to settle mostly on Eastern Föhr, increasing the German presence on the island. These patterns have created a strong geographical distribution of the languages on the island. Fering is spoken most widely in the western areas of the island and the use of Fering decreases from west to east. Low German is spoken primarily in the south, along
with Fering, and German is spoken everywhere on the island, but most widely in the eastern portions of the islands, especially in Wyk.

3.2.2 Status of Fering

German is universally spoken on Föhr, and some people speak Fering, Low German, or Danish in addition. The current number of Fering speakers is difficult to determine. Current estimates range from Roeloffs (2012), who estimates that in 2010 there were 3,500 native speakers, to Grützmacher (2010), who projects 2,000 speakers, and Bohn (2004) who reports there are 1,500 native speakers living on Föhr and an unknown number of non-native speakers and speakers not living on Föhr. Walker (personal communication) explains the wide range in estimates stems from community members’ worries that a negative outcome of a survey may discourage language revitalization efforts. Moreover, given the limited resources for language revitalization, determining the number of speakers has not been a priority (see Walker, 2001; 2007 for an overview).

The shift towards German use has caused concern for Fering's vitality and future. In 1994 Ebert published an article titled Fering, eine todkranke Sprache? [Fering, a deathly ill language?], in which she outlined dramatic shifts in the speech of a younger generation compared to older speakers. For example, she describes the loss of the distinction between A-articles and D-articles (seen in 1a and 1b below; this structure is explained in greater detail in Section 3.3 and the increase in German loan words into Fering.

(1) a. a san skiintj (Traditional Fering) ‘the A-series sun shines’
b. * di san skiintj (Young Fering) ‘the.D-series sun shines’

Ebert explains that she phrased the article title as a question because the fate of Fering was in question. While Ebert’s conclusions about the specific changes to Fering may have been correct, her prognosis of the future of Fering may have been premature. As of 2013, some children on Föhr continue to speak Fering as their native language.

3.2.3 Child Bilingualism on Föhr

The historical context, paired with recent language revitalization efforts, result in a situation where children today differ in terms of how much access they have to Fering. On Föhr, a survey of high school students in 2010 indicates that 29% of children on Föhr speak Fering (Roeloffs, 2012). This underscores the diversity on Föhr in terms of children’s level of exposure to Fering in that some, but not all, children learn Fering as an L1. Additionally, Fering language instruction contributes to further differentiations in groups of bilinguals on Föhr. North Frisian is covered under the European Charter for Regional and Minority Languages, and following the Schleswig-Holstein education policy from 2008, and renewed in 2013, Frisian instruction must be offered if there are a minimum of 12 students who are interested, but participation is voluntary. Fering is taught at both of the rural elementary schools for two hours a week in grades one through four, but it is not at all at the elementary school in the city of Wyk. At the high school level, students can choose to resume Fering education in the eleventh grade at the Gymnasium (college preparatory secondary school), where they can study Fering as one of their required language courses. Students can choose Fering as one of their subjects for their Abitur, or high school exit exam. There are also plans to extend Fering instruction to
grades 7-10 and as well as adding Fering courses at the *Regionalschule (regional secondary school)* in addition to the *Gymnasium* (Faltings, personal communication).

While participation in Fering instruction is voluntary under Schleswig-Holstein\(^1\) policy, all elementary school students at the two rural elementary schools attend Fering classes, and it is assumed that all students will participate. The students who only learn Fering in school represent a group of learners with the lowest level of exposure. This hypothesis will be addressed in Chapters 4 and 5 with a set of experiments that test the acquisition of grammatical gender and plural markings by groups of bilinguals with varying degrees of exposure to Fering.

To summarize, Föhr is currently linguistically diverse, with children differing in terms of their access to Fering input. Groups of bilinguals with diverging degrees of Fering dominance have emerged naturally from a combination of historical bilingualism, recent language shift, and community support for Fering instruction. A child on Föhr may hear only Fering at home, both Fering and German, or only German at home. The historical developments and current climate have produced differences in the amount of target language input children receive, which makes Fering an exceptional environment for testing how the amount of exposure influences acquisition of inconsistent input.

### 3.3 Fering grammatical gender

It has long been noted that Fering grammatical gender is produced inconsistently by adult speakers, meaning that adults do not agree as to the grammatical gender of nouns, and some speakers produce both genders interchangeably with certain nouns. For

---

\(^1\) The state where Föhr is located.
example, this means that some speakers say *de brek* ‘the. Masc pants’ while others say *det brek*, ‘the. Com pants’. In this section I begin with a description of the Fering grammatical gender system before describing grammatical gender inconsistency. Incorporating handbook descriptions and new data, I establish that grammatical gender is produced inconsistently in adult speech, and that items differ with regards to the level of inconsistency.

3.3.1 *Fering grammatical gender system*

Contemporary Fering has a two-way gender distinction. Table 3-1 illustrates the contemporary grammatical gender system in Fering. Gender marking surfaces throughout the adjective and nominal system, such as on definite determiners and possessive adjectives (Wilts, 1995). In singular definite noun phrases, masculine gender is marked by the determiner *a* ([a, ə]), and common gender (sometimes called feminine/neuter) is marked by the determiner *at* [at]. However, there is no gender marking on indefinite determiners. There is no gender distinction in the plural. Gender marking, however, does not surface on the noun itself. While grammatical gender assignment patterns have not been studied in depth, the Fering gender system appears to be largely lexically based and opaque. As such, there are no phonological or morphological cues to grammatical gender assignment, and the grammatical gender must be learned for each noun.
Table 3.1. Grammatical Gender and Plural Marking in Fering

<table>
<thead>
<tr>
<th>Phrase Type</th>
<th>Masculine</th>
<th>Common</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Singular</td>
<td>Plural</td>
</tr>
</tbody>
</table>

**Definite determiners**

<table>
<thead>
<tr>
<th></th>
<th>Masculine</th>
<th>Common</th>
</tr>
</thead>
<tbody>
<tr>
<td>a/de</td>
<td>boosel</td>
<td>a/dön</td>
</tr>
<tr>
<td>DET.MASC table</td>
<td>DET.PL</td>
<td>table.PL</td>
</tr>
<tr>
<td>DET.COM</td>
<td>DET.PL</td>
<td></td>
</tr>
<tr>
<td>‘the table’</td>
<td>‘the tables’</td>
<td>bike</td>
</tr>
<tr>
<td></td>
<td></td>
<td>bike.PL</td>
</tr>
<tr>
<td>‘the bike’</td>
<td>‘the bikes’</td>
<td></td>
</tr>
</tbody>
</table>

**Indefinite determiners**

<table>
<thead>
<tr>
<th></th>
<th>Masculine</th>
<th>Common</th>
</tr>
</thead>
<tbody>
<tr>
<td>een</td>
<td>boosel</td>
<td>nian</td>
</tr>
<tr>
<td>DET.COM</td>
<td>DET.PL</td>
<td>table.PL</td>
</tr>
<tr>
<td>‘a table’</td>
<td></td>
<td>‘a bicycle’</td>
</tr>
<tr>
<td>‘no tables’</td>
<td></td>
<td>‘no bicycles’</td>
</tr>
</tbody>
</table>

**Possessive Adjectives**

<table>
<thead>
<tr>
<th></th>
<th>Masculine</th>
<th>Common</th>
</tr>
</thead>
<tbody>
<tr>
<td>man</td>
<td>boosel</td>
<td>min</td>
</tr>
<tr>
<td>DET.MASC table</td>
<td>my.PL</td>
<td>table.PL</td>
</tr>
<tr>
<td>DET.COM</td>
<td>my.PL</td>
<td>bike.PL</td>
</tr>
<tr>
<td>‘my table’</td>
<td>‘my tables’</td>
<td>bike</td>
</tr>
<tr>
<td></td>
<td></td>
<td>‘my bikes’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>‘my bike’</td>
</tr>
</tbody>
</table>

**A-Articles and D-Articles**

It is important to note that Fering distinguishes two series of definite articles. These are most commonly referred to in the Fering literature as A-articles and D-articles. Although the current project does not focus on the distinction between the two series of
articles, I provide a brief description of the uses of A-articles and D-articles. Ebert (1971, 1998) provides a description of the complex system driving selection of A-articles versus D-articles. Traditionally the A-articles are the citation forms found in dictionaries, and used in general contexts, whereas D-articles are used in more specific contexts where the speakers is providing additional information about a concept which has already entered into the discourse. 2a,b below illustrate the difference between A-articles and D-article use.

(2) a. *Ik haa an wüf faan Berlin bi strun meet. Det wüf saad, hat hed an duaden siahünj fiinjen.*

‘I met a woman from Berlin on the beach. The woman said, that she found a dead seal.’


‘Did you already mend the pants? — Which pants? — The pants that I ripped during hay turning.’

In 2a, the introduction of the woman into the discourse with an indefinite, requires a D-article in the following sentence for a more specific reference to the woman. In 2b, *brek* ‘pants’ first enters the conversation as a general reference, but when the speaker asks for more specific information about which pants the speaker is referring to, *brek* ‘pants’ appears as the D-article. To summarize: Fering has traditionally had two series of definite articles— A-articles, which are used in more general contexts and citation forms, and D-articles, which are used in more specific contexts.
The primary goal of this dissertation is to determine which factors influence children’s acquisition of grammatical gender, rather than the distinction between A-articles and D-articles. Thus, I follow Parker (1993) in disregarding the fine grain semantic distinctions between these two article series. Moreover, previous research reports and local informants confirm, that the distinction is largely lost in contemporary Fering. Parker (1993) shows that younger speakers favor the D-article series over the A-article, although the A-articles were traditionally general references, and the D-series was used when the speaker intended to specify a particular reference or provide additional information. As of 1996, Hoekstra confirms that the distinction between the two series is lost in younger speakers (Hoekstra, 1996). Despite this, Fering teachers continue to cite the distinction but report it as a problematic concept for students of Fering (Roeloffs, personal communication). Contemporary textbooks and dictionaries still maintain the distinction and use the A-articles as the citation form. For the purposes of this dissertation, I will assume that the distinction between A-articles and D-articles is not robust in contemporary Fering, and I consider the articles, *a*, and *de*, to both mark masculine gender while articles, *at*, and *det*, mark common gender.

**Pronoun Agreement**

Pronouns in Fering are marked for the gender of their antecedent. The contemporary pronominal agreement system is presented in Table 3-2 (Arfsten & Tadsen, 2009; Hoekstra, 1996).
Table 3-2. Pronoun Agreement in Contemporary Fering

<table>
<thead>
<tr>
<th></th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Masculine</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Common</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Definite Article</td>
<td><em>de</em></td>
<td><em>det</em></td>
</tr>
<tr>
<td>Pronoun</td>
<td><em>hi</em></td>
<td><em>hat</em></td>
</tr>
</tbody>
</table>

**Historical Changes in Fering Grammatical Gender**

In this section I turn to the historical developments in the Fering grammatical gender system, to provide context for the investigation of current inconsistency in Fering. Within the last 100 years, Fering distinguished between three grammatical genders, masculine, feminine, and neuter, seen in Table 3-3 (Hoekstra, 1996). The contemporary two gender system arose via synchronism between feminine and neuter genders (Hoekstra, 1996). Handbooks, such as Wilts (1995), outline a distinction between the three grammatical genders in the definite article system but note that the feminine is optional.

Table 3-3. Historical Gender System in Fering

<table>
<thead>
<tr>
<th></th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Masculine</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Feminine</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Neuter</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D-Article</td>
<td><em>di</em></td>
<td><em>jü</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>det</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>do, dö(n)</em></td>
</tr>
<tr>
<td>A-Article</td>
<td><em>a</em></td>
<td><em>a</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>at</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>a</em></td>
</tr>
</tbody>
</table>
Grammatical gender syncretism is common throughout Germanic languages (Trudgill, 2012). In modern standard Germanic languages, only German, Icelandic, and Norwegian retain a three-way gender distinction—a system which is assumed to come from Indo-European (Jobin, 2011). Two-way gender distinctions are common in West Germanic languages, including Modern Dutch, East Low German, West Frisian, North Frisian, and in North Germanic languages, including Swedish and Danish (Braunmüller, 1999). In each of these languages, the masculine and feminine genders have fallen together, leaving a distinction between common gender and neuter gender. In contrast, the combination of feminine and neuter genders is unique to Fering (Hoekstra, 1996; Trudgill, 2012). In all other documented cases of gender syncretism in Germanic languages, the masculine and feminine genders fell together. Hoekstra (1996) concludes that this uncharacteristic outcome of gender syncretism stems from a replacement of the feminine pronoun with the neuter pronoun, which is common throughout German dialects. Thereafter, the syncretism in the pronominal system intruded on the definite article system, where the neuter article replaces the feminine article (Hoekstra, 1996).

The historical 3-way gender distinction is lost in contemporary Fering, leaving a two-way distinction in grammatical gender. Based on free speech interviews, word lists, sentence completion tasks, sentence translations, and sentence pairs, Parker (1993) confirms that as of 1981 Fering gender is “essentially” a two gender system, and all of the 45 informants in his study use the contemporary two gender system.

While Ebert (1998) does not find evidence of a three gender system, she asserts that the inconsistency seen in the contemporary two gender system is evidence of a covert feminine gender. That is, she asserts that the use of both masculine and common genders
interchangeably is a marker of the historical feminine gender. Hoekstra (1996) rejects Ebert’s claims saying, “Synchronically speaking, I don’t see any reason to consider these occurrences of the A-article [alternations between a and at] as ‘Manifestationen der Weiblichkeit’ [manifestations of the feminine] (Ebert 1971:97-101). Modern Fering lacks feminine gender.” (1996, 62-63). Thus, although historically Fering had a three gender system, the current system includes two grammatical genders, masculine and common gender. Thus, a contemporary learner of Fering must determine whether a given noun is associated with masculine or common gender.

To summarize: Although historically Fering had three grammatical genders, the contemporary Fering grammatical gender system marks definite articles and pronouns with one of two grammatical genders (ie., masculine and common). The learner’s task, then, is to determine from the input which gender is associated with each noun and to mark determiners and pronouns for grammatical gender accordingly in their own speech.

3.3.2 Inconsistency in Fering Grammatical Gender

While textbook descriptions of Fering show gender marking to be categorical, in reality there is significant inconsistency in gender marking in everyday speech. In Fering, inconsistency refers to the use of both masculine and common gender with a given noun. For example, *brek* ‘pants’ may appear as either *de brek* ‘the.MASC pants’ or *det brek* ‘the.COM pants’. Thus, while definite noun phrases require a definite determiner, which gender marked definite determiner appears with a noun cannot be predicted from the context, and is therefore inconsistent.
Previous research has attempted to determine the nature of this inconsistency. Parker (1993) is one of the most comprehensive studies of the Fering gender system to date. In his study of 45 Fering speakers, he showed that younger Fering speakers exhibit a higher level of inconsistency in gender marking than older speakers, and he suggests that this increase in inconsistency may stem from lower proficiency among younger speakers.

Ebert (1998) posits several accounts to explain the inconsistency in current Fering grammatical gender. Ebert asserts that the alternations in grammatical gender relate to semantic differences. For example, nouns that are inherently connected, such as body parts, use the masculine gender when they are in their normal inalienable state, but common gender when the part is removed, such as the nose of a child’s painting of his father in examples 3a-b.

(3) a. Fering: *Dü heest an drebel bi a nöös* Ebert (1998)
   You have a drop by the.MASC nose
   ‘You have a drop on your nose’ (connected body part)

   b. Fering: *Dü heest at nöös jo ferjiden* Ebert (1998)
   You have the.COM the nose forgot
   ‘You forgot the nose’ (unconnected body part)

Another category of nouns that alternate between masculine gender and common gender are words for parts of the house and household objects. When the noun in question refers to the object within the house, the common gender article is used, and when the function of the object is referenced the masculine gender article is used. This can be seen
particularly well in recent loan words from German (4a,b), such as maskiin ‘machine’
from German ‘die Machine’.

(4) a. Fering: \textit{Det haa ik me a maskiin säid} Ebert (1998)
That have I with \textbf{the.Masc Machine} sewed

\begin{tabular}{ll}
\text{German:} & ‘Das habe ich mit der Maschine genäht’
\text{– function}
\end{tabular}

b. Fering: \textit{Det schuar lait bi’t säimaskiin} Ebert (1998)
The scissors lie next to \textbf{the.COM sewing machine}

\begin{tabular}{ll}
\text{German:} & ‘Die Schere liegt neben der Nähmaschine’
\text{– object}
\end{tabular}

Ebert (1998: 273) posits that nouns with inconsistent gender use may also
represent separate lexical entries, such as \textit{a koon} ‘the.COM coffee pot’, but \textit{at/det koon}
‘the.FEM can’. In these cases, where the semantic referent is significantly different, this
may be a reasonable conclusion. However, there are many words that cannot be
explained by these semantic differences. Because Ebert is not able to find a clear
explanation for grammatical gender alternations, Ebert (1994) posits that it may be
overall low proficiency among younger speakers that drives the inconsistency. However,
even if more younger speakers use grammatical gender interchangeable, it is unclear
whether this reflects lower proficiency or language change.

Both Parker (1993) and Ebert (1998) associate inconsistent use of grammatical
gender with lower proficiency. Ebert (1998) sees this as a general pattern of language
death and asserts that speakers, even the most proficient speakers, are often unsure of the
gender. She argues inconsistently marked gender among younger speakers indicates the decline of the language in general.

To summarize, previous corpus and interview data from Parker (1993) and Ebert (1998) demonstrate that grammatical gender is produced inconsistently in Fering, thereby providing children with inconsistent language input, potentially impeding grammatical gender acquisition in Fering. However, the data from Parker and Ebert were collected over 30 years ago and may not reflect current language use or input available to children learning Fering today. Therefore, it is necessary to confirm that grammatical gender is produced inconsistently in contemporary Fering. Moreover, Parker (1993) and Ebert (1998) provide examples of inconsistency, but do not provide any specific information about the frequency of grammatical gender inconsistently or whether grammatical gender is limited to a few highly inconsistent speakers or items. The corpus study and adult survey presented investigates grammatical gender inconsistency in a detailed and systematic way, which allows them to serve as a baseline for the experimental research presented in Chapter 4.

**Corpus Study**

To further investigate inconsistency in Fering grammatical gender marking, a corpus study of common words was conducted using *My First 1000 Words*, a children’s picture dictionary which is available for all dialects of North Frisian and Low German (Amery, 1998). The nouns were selected from the Low German version², and then the Fering translations were entered into a database using the German translations, the Fering

---

² At the time the database was created I did not have access to the Fering version of the book. However, the items were the same across the Low German and Fering versions.
translations (Wilts, 2011). *My First 1000 Words* was selected because it has been used in North Frisian language revitalization efforts, all items are highly frequent, and it is available in German, Low German, and Fering, as well as all other dialects of North Frisian. As the words were all frequent in child vocabulary, the corpus study was able to serve as a source for the materials in the experimental tasks administered to children. A database was created with 597 Fering nouns\(^3\) along with the Fering grammatical gender and plural marking, which were determined in consultation with two Fering-German dictionaries (Wilts, 1982, 2011), and German and Low German translations. The database was used to examine the frequency of grammatical gender and plural markings was analyzed, as well as the frequency of word with matching grammatical gender and plural marking between German and Fering.

The results revealed that roughly half of the Fering words in this corpus study are masculine gender (44.2%) and half are common gender (47.2%). The remaining 51 items (8.5%) were marked as both common and masculine gender, as determined by the Fering-German dictionaries. The distribution indicates that although feminine and neuter genders have fallen together, the frequency of masculine and common gender in contemporary Fering is balanced.

I next turn to comparisons between Fering and German grammatical gender. This is important for three reasons. Psycholinguistic studies have shown that bilinguals are sensitive to similarities between words with the same grammatical gender between two languages (Bordag & Pechmann, 2007). As all speakers of Fering are also L1 German

---

\(^3\) 597 of the 1000 original words were found in the Fering-German dictionaries. This excluded verbs, and uncommon nouns, such as Judo, which was in *1000 Words* book, but not found in the Fering-German dictionary.
speakers, it was important to take German into consideration. Secondly, cross-linguistic interaction is especially strong when there is a high degree of lexical overlap in genders, meaning that translations are marked with the same gender across languages (Sabourin & Stowe, 2008; Sabourin, Stowe, & de Haan, 2006). As such, it is important to establish how often words in Fering and German share the same grammatical gender. Finally, previous studies on Fering, including Ebert (1994; 1998) and Parker (1993) suggest that Fering speakers transfer the German gender for Fering, and that words with cross-linguistically conflicting gender are highly inconsistent. Comparing Fering to German, there were 69 words of the 597 (12%) which are masculine in German, but were marked as common gender in Fering. There were 59 (10%) words which are feminine or neuter in German, but are not common gender (previously feminine and neuter) in Fering. Therefore, in this sample of Fering 21% of words do not share the same grammatical gender with their German translation.

To summarize: In contrast to the handbook and textbook descriptions which show grammatical gender as categorical, the results of this corpus study indicate gender is marked inconsistently. 8.5% of the words in this study were marked as both genders by dictionaries, meaning that speakers may use both genders interchangeably, as compared to 0% of the German words (Wilts, 2011). From this study, it is not possible to determine whether these the nouns that are marked as one gender are used exclusively with this gender, or whether they are produced inconsistently to a lesser degree. The corpus study investigated the scope of inconsistent grammatical gender in formal accounts, in that it investigated how many words use interchangeable gender, but it could not determine the number of speakers who use interchangeable gender with a given item. This is the goal of
the adult survey. While the corpus study can identify the overall distribution of masculine versus common gender in modern Fering, it was not possible to determine whether all items are equally inconsistent or whether the inconsistency stems from intra- or inter-speaker variation.

3.3.3 Adult Survey

A grammatical gender assignment task was designed to investigate the level of inconsistency in grammatical gender marking and to determine whether inconsistency is driven by few highly inconsistent speakers or is a general aspect of the Fering gender system.

Methods and Design

Fifty-four items from the 1000 Words corpus words were included in the grammatical gender assignment task. As such, all items were highly frequent. Items were inanimate or animate (but non-human) nouns that were non-cognates with German and Low German. Moreover, a native speaker of Northern German vernacular (NGV, see Reershemius, 2009, for a discussion of the syntactic features; Schröder, 2004 for phonological features of this variety) noted that none of the items were cognates with NGV. Fering grammatical gender was determined in consultation with My First 1000 Words and two German-Fering dictionaries (Wilts, 1982, 2011). For these 54 words, all three sources agreed on the gender of the nouns. The order of presentation of items was randomized across participants. All participants also completed a plural marking task (See Section 3.4.3) and the order of tasks was randomized across participants.
Previous research suggests that for both child and adult second language (L2) learners, the acquisition of grammatical gender is facilitated when both the L1 and the L2 assign the same gender to a noun (e.g., Bordag & Pechmann 2007; Sabourin & Stowe 2008; Sabourin, Stowe & de Hann 2006). This suggests that words with the same grammatical gender in Fering and German, such as *a tüs* ‘the.Masc tooth’ compare German *der Zahn* ‘the.Masc tooth’, would be easier to acquire than words with incongruent grammatical gender between Fering and German. Since the participants in the present study are all bilingual in German, the gender of each target word in German was taken into consideration when creating item categories. As seen in the examples in Table 3-4, the final set of 54 items was separated into four categories based on Fering grammatical gender, and the gender overlap with German. Overall, 31 items had matching gender with German, 19 items had mismatching gender, and four items were listed as both masculine and common gender in Fering, according to the (according to Wilts (2011) and the *1000 Words Corpus*) and were considered neither matching nor mismatching items.

### Table 3-4. Categories of Items for Gender Assignment Task

<table>
<thead>
<tr>
<th>Fering Grammatical Gender</th>
<th>Fering-</th>
<th>Fering</th>
<th><em>de</em></th>
<th>boosel</th>
<th>Fering</th>
<th><em>det</em></th>
<th>schört</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fering-MASC</td>
<td>the</td>
<td>the.MASC</td>
<td>table</td>
<td></td>
<td>the.COM</td>
<td>shirt</td>
<td></td>
</tr>
<tr>
<td>Gender Match</td>
<td>German</td>
<td>der</td>
<td>Tisch</td>
<td>German</td>
<td>das</td>
<td>Hemd</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>--------</td>
<td>-----</td>
<td>-------</td>
<td>--------</td>
<td>-----</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>the.MASC</td>
<td>table</td>
<td></td>
<td>the.NEUT</td>
<td>shirt</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| ‘the table’ | ‘the shirt’ |

<table>
<thead>
<tr>
<th>Fering-</th>
<th>Fering</th>
<th>de</th>
<th>dask</th>
<th>Fering</th>
<th>det</th>
<th>höös</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>the. MASC bowl</td>
<td>the.COM</td>
<td>stocking</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gender Mismatch</th>
<th>German</th>
<th>die</th>
<th>Schüssel</th>
<th>German</th>
<th>der</th>
<th>Strumpf</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>the.FEM</td>
<td>bowl</td>
<td>the. MASC</td>
<td>Stocking</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| ‘the bowl’ | ‘the stocking’ |

**Participants**

Participants were recruited through local contacts living on Föhr, including two Fering teachers and the Ferring Stiftung, a local library for Fering studies. All participants who were over 18 years old and speak Fering were encouraged to participate. The sample included both early bilingual speakers, and second language learners. A link to the online survey was also published in a local newspaper and posted on the Ferring Stiftung website and the main page of the Frisian language Wikipedia page\(^4\). 62 participants completed the survey. I report here on the results of 45 speakers who began speaking Fering before the age of two, and current used Fering in their everyday interactions. On average, participants were 39 years old (Range: 18-71 years of age) and one participant did not provide age information.

\(^4\) frr.wikipedia.org
Procedure

Participants were presented with a Fering noun and were asked to select which gender they would use on a novel 7-point scale ranging from 1 ‘a’ (masculine) to 7 ‘at’ (common), with 4 corresponding to ‘a/at’ (interchangeable. Participants were instructed to mark 1 if they always use masculine gender with the noun, 7 if they always use common gender. They marked 4 if they would use both interchangeably and 2 or 3 if they use both genders but masculine gender more often than common gender, and 5 or 6 if they used both genders but common more often than masculine. Marking an item as interchangeable was taken to indicate that the speaker uses the gender marking inconsistently.

Using this likert scale, participants were able to mark their gender preference for an item and the relative frequency with which they would use each gender. Importantly, the participants were instructed not to consult a dictionary but to choose the gender that fit their “gut feeling”. The task required participants to reflect on their grammatical gender use. For the purposes of this study, the level of awareness does not undermine the results. If the results are affected by speaker’s level of awareness, I expect that they will underestimate their level of variability, and thus any results would represent a minimum level of variability. Comparisons between participants give a measure of the inter-speaker variability, and an item analysis shows the amount of intra-speaker variability.

\[5\text{ Because the A-Articles are typically used in the Frisian literature, they were used for this scale. However, since this pilot project was conducted, it has become clear that the D-articles are used more often in contemporary Fering, and therefore I have adopted the D-article notation throughout the dissertations.}\]
Results

On the 7-point likert scale, responses 1 and 2 were combined and considered consistent masculine gender responses. If participants marked 3, 4 or 5, this indicated that they use both genders with an item, called an interchangeable response. As with the masculine gender, responses 6 and 7 responses were combined and considered consistent common gender responses.

I first examine the results by speaker before turning to look at responses by item. To investigate each speaker’s level of grammatical gender inconsistency, the percentage of items that were marked as ‘interchangeable’ was calculated for each speaker. 80% of participants \( (n = 36) \) reported that they use both genders for at least one item. Interchangeable gender responses were common across all speakers and widespread. The average percentage of speakers’ interchangeable gender responses was 14% (range 0%-47\%, \( SD = 11.6\%) \). As Table 3-5 shows, about half of all responses were masculine gender—meaning that participants chose 1 or 2 on the scale, while 32% of all responses were common gender—meaning that participants chose 5 or 6 on the scale. Thus, intra-speaker inconsistency cannot be accounted for by a few highly inconsistent speakers, but rather appears to be a general pattern across most speakers of contemporary Fering.

Table 3-5. Participant Responses by Grammatical Gender

<table>
<thead>
<tr>
<th>Response Type</th>
<th>Mean</th>
<th>SD</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Masculine (Responses 1 and 2)</td>
<td>.54</td>
<td>.17</td>
<td>.17</td>
<td>1.00</td>
</tr>
<tr>
<td>Interchangeable (Responses 3, 4, and 5)</td>
<td>.14</td>
<td>.12</td>
<td>.00</td>
<td>.47</td>
</tr>
</tbody>
</table>
The results are similar if I look at the responses by item. Here I analyzed the percentage of the speakers in the study who indicated they use interchangeable gender for the item. All but three items show some level of inconsistency in gender marking. For 94% of items ($n = 51$) included in the survey, at least one speaker indicated that she would use both genders with the item. For these 51 items, the average percentage of interchangeable gender responses (responses 4, 5 or 6) for each item was 15% (range 0% - 50%, $SD = 12%$). As with speaker level results, the item analysis shows a general pattern of gender inconsistency.

While most items show some level of inconsistency, some items are consistently marked. For 31% ($n = 17$) items included in the survey, over 75% of the participants agreed on the gender they would use. Conversely, 69% of the items ($n = 37$) showed a high level of inter-speaker inconsistency, with less than 75% of speaker agreement. Six items were consistently marked as masculine, with more than 90% of participants agreeing. Those were *tuun* ‘fence’, *strun* ‘beach’ *kai* ‘key’, *boosel* ‘table’, *dööbel* ‘dice’, and *tus* ‘tooth’. The highest level of consensus for common gender items was 80%, and on five items at least 70% of participants reported they always use common gender with the item: *bleed* ‘newspaper’, *wel* ‘bike’, *knif* ‘knife’, *fooder* ‘hay’, *schört* ‘shirt’, and *kop* ‘cup’. Results for all items are presented in Appendix A.

No significant differences were found between items with matching gender in German and those with conflicting gender. That is, items for which both the German and
Fering translations shared the same gender and those with different genders had similar levels of interchangeable gender marking. On average, 13.8% of participants reported that they use interchangeable gender with items with matching gender compared to 16.3% of participants for items with mismatching gender. A paired samples $t$-test revealed no significant difference between the frequency of interchangeable gender for matching versus mismatching items ($t(43) = .69, p = .492$). Excluding the four items with interchangeable gender in the dictionary, for ten items the majority of participants disagreed with the dictionary grammatical gender. Of these items, five items changed to align with the German gender for this word, four maintained the gender mis-match with German, and one item showed no preference for either gender. Thus, it appears that participants are not shifting towards aligning the Fering gender with German gender, or adopting the German grammatical gender in Fering. Instead, participants may be shifting to using masculine gender for all items. Indeed, participants reporting using masculine gender for significantly more items than common gender, as revealed in a paired samples $t$-test ($t(49) = 3.30, p = .002$).

**Summary and Discussion**

The results of this adult online survey support previous research, showing that grammatical gender is marked inconsistently in Fering (Ebert, 1998; Parker, 1993). Overall inconsistency is common and for over two thirds of the items included in the survey, less than 75% of speakers agree on which gender they use. Moreover, these results expand on previous research by showing that variability in gender markings is not limited to a sub-set of highly inconsistent items or speakers. Inconsistency appears to be
more prevalent than previous research and the corpus study of words from the book *My First 1000 Words* revealed. While the majority of items showed some level of inconsistency, items differ in terms of the level of inconsistency, and common gender words were overall more inconsistent than masculine gender words. This suggests a tendency to use masculine more often than common gender. These naturally occurring differences across items make Fering a prime environment for testing the impact of the inconsistency in child acquisition.

### 3.4 Fering plural marking

While grammatical gender marking proved to be inconsistently produced based on the survey results presented in 3.3.3, there are no previous reports of inconsistency in noun plural marking. In section 3.4, I outline the linguistic description of plural marking in Fering, drawing on previous literature, and present new data from a corpus study of common Fering words and an adult online survey. Taken together, these results show that while plural marking is not unanimous in Fering, it is indeed more consistent than grammatical gender.

#### 3.4.1 Previous Literature

Fering plural is marked by affixing a plural marking onto the root noun. As seen in Table 3-1, all plural nouns use the same definite determiner regardless of the plural marking. In Fering, only nouns and determiners are marked for plural, not adjectives. There are five regular plural markers in Fering: *-er, -en, -in, -n*, and *-s*, all of which attach to the end of the noun (Wilts, 2011). For the plural markings, *-er* and *-en*, unstressed vowels in the
Coda are dropped when the plural marking is added. Examples of regular plural marking are presented in Table 3-6.

Table 3-6. Regular Plural Marking in Fering (Wilts, 2011)

<table>
<thead>
<tr>
<th>Regular Plural Marking</th>
<th>Fering Singular</th>
<th>Fering Plural</th>
<th>German Plural</th>
<th>English Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>-er</td>
<td>a hünj</td>
<td>a hünjer</td>
<td>Hunde</td>
<td>dogs</td>
</tr>
<tr>
<td></td>
<td>at wöning</td>
<td>a wönger</td>
<td>Fenster</td>
<td>windows</td>
</tr>
<tr>
<td>-en</td>
<td>at buk</td>
<td>a buken</td>
<td>Bücher</td>
<td>books</td>
</tr>
<tr>
<td></td>
<td>at foomen</td>
<td>a foomen</td>
<td>Mädchen</td>
<td>girls</td>
</tr>
<tr>
<td>-in</td>
<td>at baantje</td>
<td>a baantjin</td>
<td>Ämter</td>
<td>offices</td>
</tr>
<tr>
<td>-n</td>
<td>a mooler</td>
<td>a moolern</td>
<td>Maler</td>
<td>painters</td>
</tr>
<tr>
<td>-s</td>
<td>a koptein</td>
<td>a kopteins</td>
<td>Kapitäne</td>
<td>captains</td>
</tr>
</tbody>
</table>

In addition to the regular plural markings, there are also irregularly marked nouns in Fering. Wilts (1995) notes seven categories of irregularly marked noun plurals, but Wilts (2011) presents a simplified classification system with three categories of irregular noun plural markings. Table 3-7 provides examples of these categories of irregular plural marking based on the simplified system in Wilts (2011).

Table 3-7. Irregular Plural Marking in Fering (Wilts, 2011)

<table>
<thead>
<tr>
<th>Irregular Plural Marking</th>
<th>Fering Singular</th>
<th>Fering Plural</th>
<th>German Plural</th>
<th>English Plural</th>
</tr>
</thead>
</table>
### Marking

<table>
<thead>
<tr>
<th>[zero]</th>
<th>det bian</th>
<th>a bian</th>
<th>Beine</th>
<th>legs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consonant Change</td>
<td>det knif</td>
<td>a kniiwer</td>
<td>Messer</td>
<td>knives</td>
</tr>
<tr>
<td>Vowel Change</td>
<td>det skap</td>
<td>a skeb</td>
<td>Schiffe</td>
<td>shoes</td>
</tr>
<tr>
<td>Word Change</td>
<td>det kint</td>
<td>a jongen</td>
<td>Kinder</td>
<td>children</td>
</tr>
<tr>
<td></td>
<td>de maan</td>
<td>a lidj</td>
<td>Leute</td>
<td>people</td>
</tr>
</tbody>
</table>

#### 3.4.2 Corpus Study

The previous literature outlines a complex picture of Fering plural marking with a large inventory of plural markings. However, the frequency of each plural marking is not presented in handbook accounts, meaning that one or two plural markings may dominate the system. Secondly, many of the regular markings are also plural markings in German, but it is not clear how often the translations equivalents share the same plural marking between languages. To determine the distribution and frequency of each plural marking, a corpus was created using the items from the 1000 Words sample, described in Section 3.3.2.
Within the *1000 Words* sample, 508 words are marked for plural. Of these words, 38 (7.5%) fall into the category of “irregular” plurals as indicated by Wilts (1995). Five plural markings were coded, -*er*, -*en*, ø, -*in*, -*s*, which were categorized as regular markings, following Wilts (1995). The –*er* and –*en* endings were by far the most common. The –*er* ending accounted for 32% (*N* = 163) of the nouns, and the –*en* ending accounted for 53% (*N* = 270) of the nouns. Each of the other markings, -ø, –*in*, and –*s* accounted for 3% or less of the words.

One of the striking differences between German and Fering plural marking is the high rate of unstressed vowel deletion in Fering plurals. In nouns ending in an obstruent + unstressed vowel + sonorant consonant cluster, such as *goobel* ‘Gabel, fork’, the medial unstressed vowel is deleted, giving the form *goobler* ‘forks’. Unstressed vowel deletion occurs with the markings –*er*, -*en*, and –*in*. This is seen in the examples, *taarep* (sg) ‘Dorf village’ ~ *taarpen* (pl) ‘Dörfer villages’ and *tante* (sg) ‘Tante aunt’ ~ *tantin* (pl) ‘Tanten aunts’.

To summarize: the results of the corpus study confirmed previous research showing no inconsistency in noun plural marking. Of the 508 items included in the corpus study, none are marked with more than one marking. That is, there are no cases of interchangeable plural.

### 3.4.3 Adult Survey

While the corpus study and previous literature show noun plural as consistently marked, a plural marking task administered via an online adult survey was conducted to determine the current use of Fering plural markings. Forty-two items were included in the
plural assignment task, which were divided into three categories. Items differed in their plural marking and whether the German translation had a matching plural marker or not.

Table 3-8 provides examples for each category. The Plural Match category included items with either –en or –er in Fering, and the same plural marker in German \( (n = 10) \). The Plural Mismatch-er category included item with –er marking in Fering, and a different plural marker in German \( (n = 13) \). The Plural Mismatch-en category included items with –en marking in Fering, and a different plural marker in German \( (n = 19) \).

<table>
<thead>
<tr>
<th>Category</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Match</td>
<td>Fering</td>
</tr>
<tr>
<td></td>
<td>kop-en cup-PL</td>
</tr>
<tr>
<td>German</td>
<td>Tasse-n cup-PL</td>
</tr>
<tr>
<td></td>
<td>‘cups’</td>
</tr>
<tr>
<td>Mismatch-er</td>
<td>Fering</td>
</tr>
<tr>
<td></td>
<td>fask-er fish-PL</td>
</tr>
<tr>
<td>German</td>
<td>Fisch-e fish-PL</td>
</tr>
<tr>
<td></td>
<td>‘fish’</td>
</tr>
<tr>
<td>Mismatch-en</td>
<td>Fering</td>
</tr>
<tr>
<td></td>
<td>wel-en bike-PL</td>
</tr>
<tr>
<td>German</td>
<td>Fahrräd-er bike-PL</td>
</tr>
<tr>
<td></td>
<td>‘bikes’</td>
</tr>
</tbody>
</table>

**Participants and Procedure**

The same 45 native speaker participants from the gender assignment task completed the plural completion task. Participants were presented with a word written in Fering and were asked to supply the plural form of the noun. Participants also gave a confidence rating for each plural item on a scale from 1-5 \( (1 = \text{not at all sure}; 3 = \text{a little sure}; 5 = \text{very sure}) \). All participants completed the grammatical gender task (Section
3.2.2) prior to the plural assignment task. The order of items within the plural assignment task was randomized across participants.

**Results**

Results from the plural task show a very different pattern than those from the grammatical gender task. Overall, the participants were much more consistent in their plural markings than in their grammatical gender markings. Responses were coded for the plural marking used. The results of all plural items are presented in Appendix B.

Half of the items (\(n = 22\)) were predominantly marked with –en marking, with over half of participants agreeing on the plural marking (average level of consensus is 79% on these items). Twelve items (22%) were primarily marked with –er marking. Of all responses, less than 1% of responses were marked with –s marking, and about 5% of responses were marked with a null marking, meaning that no ending is added. 9.3% of all responses were left blank or the participants indicated they did not know the word. These items may reveal dialect differences within Fering or items for which the German borrowing is more common than the native Fering word. For example, blijant ‘pencil’ is typical of Fering spoken on the eastern part of Föhr, but most participants came from areas in the western part of Föhr, where the word pokluad ‘pencil’ is more common. 22% of participants responded that they did not know this item or left it blank. For the item steulis ‘Schaukel swing’, 18% of participants indicated they did not know the item or left the plural blank. Local informants report that this is more often produced with the German borrowing, Schaukel ‘swing’ than the native Fering word. In contrast to the gender task results, participants largely concurred with the dictionary plural marking for
all but five items. The average confidence rating for all items was 3.98 on a five-point likert scale (SD = .75), suggesting that participants were on average confident in their plural markings.

Summary and Discussion

The adult online survey revealed much less inconsistency in plural marking than in grammatical gender marking. Additionally, the results confirm the corpus study’s results that the most frequent markings are –er and –en endings, with 28 items (67%) of the items in the survey taking an –en marking, and 12 (29%) taking the –er marking. While plural markings were not unanimous, most speakers agree on which plural marking they would use, and for all but five items, most speakers’ responses corresponded with the plural marking provided by Fering dictionaries. Of the items included in the survey, over two thirds (69%) of items are high consensus, meaning that more than 75% of the speakers agree on the plural marking. These results suggest that the child’s input is more consistent for noun plurals and grammatical gender. Thus, because Fering plural is produced consistently I would predict that children with more Fering input will acquire plural markings more quickly than peers with less Fering input. The results of comprehension and production tasks in plural marking in Chapter 5 test this hypothesis.

The consistency in plural marking in Fering allows plural marking to serve as a contrast to inconsistent grammatical gender marking. Thus, Fering’s consistent marking of plural, taken together with inconsistent grammatical gender, further strengthens the
case of Fering as an exemplary environment for testing the role of level of inconsistency in acquisition of inconsistent input.

### 3.5 Summary and Discussion

This chapter provided a description of the grammatical gender and plural marking system in Fering, which will be used as an example of how both the type and quantity of language input influence the acquisition of inconsistent morphology. I summarized the handbook accounts of grammatical gender and plural marking, and presented a corpus study and online survey of the contemporary Fering morphology system. Through a survey of native Fering speakers, I established that grammatical gender is produced inconsistently, while plural markings are produced much more regularly in modern day Fering. Combining the results of the handbook descriptions, the *1000 Words* corpus, and the adult online survey highlights the differences between grammatical gender and plural markings in Fering. To compare the level of consensus of each feature directly, we can look at the percentage of items in the adult survey where more than 75% of participants agree on the gender or the plural marking. For grammatical gender, just 31% of the items meet this minimum level of consensus, compared to 64% of plural items.

By using Fering as a case study, this dissertation tests the theory proposed by Hudson Kam and Newport (2005, 2009) that children regularize inconsistent input. Investigating Fering allows for testing two factors—consistent versus inconsistent input, and quantity of language input. A comparison between consistently produced plural markings and inconsistently produced grammatical gender tests the role of input type in acquisition. The wide range of Fering input available to children on Föhr builds on
Hudson Kam and Newport (2005, 2009) where all children had very limited exposure to the target language.

The prediction is that children will acquire the consistent feature, noun plural markings, before acquiring the inconsistent feature, grammatical gender. Moreover, I predict that low input children will regularize the inconsistent input, following results reported by Hudson Kam and Newport. However, I predict that high-input children will be able to acquire both consistently and inconsistently produced morphology. That is, I predict that the regularization seen in previous literature reflects the early stages of acquisition, and that with additional input children will be able to overcome the initial stage of regularization. Chapters 4 and 5 test these predictions through a series of grammatical gender and plural acquisition experiments.
Chapter 4. Results: Grammatical Gender

4.1 Introduction

Recent research on the acquisition of inconsistent morphology—morphology which is produced probabilistically and is not conditioned by linguistic or extralinguistic factors—suggests that children regularize inconsistent input (Hudson Kam & Newport, 2005; 2009). These studies are some of the first to investigate the acquisition of inconsistent input and leave several questions unanswered. These studies used artificial languages with very short exposure times, and it is not clear whether children regularize inconsistent input when learning natural languages and how the amount of input influences acquisition. In this dissertation I investigate how the amount of language input influences how children acquire inconsistent input.

This chapter presents the results of three experiments, which investigate the influence of the quantity of input on the acquisition of an inconsistently produced feature. Fering grammatical gender provides a useful test-case for inconsistent input, as previous research (e.g., Ebert, 1998; Parker, 1993) and the adult online survey (presented in Chapter 3), suggested that grammatical gender is produced inconsistently by adult speakers of Fering. Moreover, children on Föhr have a wide range of Frisian language input, which facilitates the investigation of how the quantity of input influences the acquisition of inflectional morphology in Fering.

This chapter addresses two broad research questions:

1. How does the quantity of Frisian language input influence participants’ acquisition of grammatical gender?
2. Do participants regularize inconsistently produced grammatical gender?

Throughout the chapter, two sets of comparisons test the effect of the quantity of input on acquisition. First, in the Home Language comparison, three groups of elementary school students with different amounts of Frisian input will be compared. Secondly, in the Age Group comparison, high-input participants in three age groups: elementary school students, high school students, and adults will be compared to test for the effect of cumulative language exposure. The assumption is that because all of these participants have high input, that over time the adults have had more cumulative Frisian input than the high school students, who have had more cumulative Frisian input than elementary school students.

Across both comparisons, the prediction is that participants with more Frisian input will show more adult-like acquisition of grammatical gender. This means that I predict that high-input elementary school students and high school students will produce grammatical gender following adult-like patterns, regularize grammatical gender less often than lower input peers, and produce and comprehend gender-marked pronouns in an adult-like way. I will return to this broad prediction throughout the chapter when discussing the specific predictions for each task and comparison.

Grammatical gender acquisition was tested in three tasks: a definite article elicitation task, a gender-marked pronoun elicitation task, and a gender-marked pronoun comprehension task. I first present the participants, materials, and procedures in Sections 4.1.2-4.1.4. Section 4.2 presents the results of the definite article elicitation task, followed by the gender-marked pronoun task and gender comprehension task, in sections 4.3 and 4.4, respectively. Section 4.5 summarizes and discusses all three tasks.
4.1.1 Fering Grammatical Gender

Fering marks nouns with one of two grammatical genders: masculine gender or common gender. Common gender includes the historically feminine and neuter gender words. In this Chapter I focus on grammatical gender marking on definite articles and on gender-marked pronouns, as seen in Table 4-1. Refer to Chapter 2 for a more in depth discussion of the Fering grammatical gender system.

Adult native speakers of Fering produce grammatical gender inconsistently. This means that adults use both masculine gender and common gender interchangeably with the same noun. For example, about half of adults produce *brek* ‘pants’ as *de brek* ‘the.MASC pants’ while about half produce it as *det brek* ‘the.COM pants’. Importantly, these alternations do not appear to be conditioned by linguistic or extra-linguistic factors. Therefore, from the learner’s perspective, the choice of grammatical gender may appear random. However, this does not mean that there are no statistical patterns in adult speech. As the results of adult survey showed (see Chapter 2), individual nouns differ in their level of consensus. That is, there are some nouns that adult speakers largely agree on which grammatical gender to use and nouns for which fewer adults agree. In this chapter these differences are referred to as high-consensus and low-consensus. With this in mind, successful acquisition of grammatical gender entails learning these statistical patterns.

Table 4-1. Fering grammatical gender system

<table>
<thead>
<tr>
<th>Phrase Type</th>
<th>Masculine</th>
<th>Common</th>
</tr>
</thead>
<tbody>
<tr>
<td>Definite determiners</td>
<td><em>a/de</em></td>
<td><em>boosel</em></td>
</tr>
<tr>
<td></td>
<td><em>DET.MASC</em></td>
<td><em>table</em></td>
</tr>
</tbody>
</table>
4.1.2 Participants

Three groups of participants completed the grammatical gender tasks: elementary school students, high school students, and adult parent and community members. The elementary school participants were recruited from both of the rural elementary schools on Föhr. All parents were given an invitation letter and a consent form, and elementary school students with a returned signed consent forms were asked to participate. High school participants were recruited from Frisian language classes in grades 11-13. Parents of students under 18 years old consented for the student, while students over 18 consented for themselves. All high school students were compensated for their participation. All parents of elementary school students were asked to participate in the same experiments as their children, and additional adult community members were recruited through local contacts. All adult participants were compensated for their participation. As this dissertation investigates child language acquisition, the elementary school students are of primary interest. The high-input high school students and adult speakers serve as comparison groups.
**Level of Input**

Participants were separated into groups based on the amount of Frisian spoken at home using the Utrecht Bilingualism Exposure Calculator (UBiLEC) (Unsworth, 2011). While this relies on self-reporting, parent questionnaires have been shown to accurately reflect actual child language input (Gutierrez-Clellen & Kreiter, 2003). Parents of elementary school participants were interviewed either in person or on the telephone, and when it was not possible to contact the child’s parents, the child was consulted (Unsworth, 2011). High school participants were interviewed directly.

The UBiLEC provides detailed information about the quality and quantity of language input a child is exposed to over time, while also synthesizing this information into broader indicators. For the purposes of this dissertation, a sub-set of the full questionnaire was used. Participants were asked about the language spoken at home, how often each family member speaks Frisian with the child, the age of acquisition for Frisian, the parents’ age of Frisian acquisition, the language spoken with peers at school and after school, and the language spoken during extra-curricular activities. The researcher entered the responses directly into the Excel file which accompanies the UBiLEC and tabulates the responses.

Participants were split into three groups based on their exposure to Frisian at home. Participants from homes where Frisian is spoken 75% -100% of the time were categorized as *high-input* participants. Participants from balanced bilingual homes, where Frisian is spoken between 25% and 75% of the time, were considered *middle-input* participants, and participants who spoke very little or no Frisian at home were considered *minimal-input* participants. Of the 45 elementary school students who participated, 20
were high-input, 12 were middle-input, and 13 were minimal-input. The results of the middle and minimal-input high school students and adult L2 speakers are not presented here.

4.1.3 General Materials

All grammatical gender tasks tested the same 24 items, which were separated into four categories based on the grammatical gender and level of consensus. There were four categories of items: High consensus masculine, low consensus masculine, high consensus common and low consensus common. Sample items are presented in Table 4-2 along with the level of consensus. Items were non-human, non-cognates with German, Low German, and Northern German Vernacular, and mimicked the overall gender distributional pattern in Fering. The adult survey showed that dictionaries may not reflect current use (Wilts, 1982; 2011). Thus, the grammatical gender was determined by the adult survey results rather than the dictionary gender. Because grammatical gender is produced inconsistently, and adults do not agree on the grammatical gender of nouns, it can be difficult to categorize nouns by grammatical gender. For the creation of categories of items, the grammatical gender for a noun was determined as the gender with the most responses from the adult survey (See Chapter 3 for a full description of the online-survey). All but one item was selected from the adult survey.

Table 4-2. Example stimuli for grammatical gender tasks

<table>
<thead>
<tr>
<th>Condition</th>
<th>Example item (% consensus)</th>
<th>Consensus range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Masculine – High Consensus</td>
<td><em>a boosel 'the.Masc table'</em> (96% masc.)</td>
<td>80-96% masculine</td>
</tr>
<tr>
<td>Masculine – Low Consensus</td>
<td><em>a sink 'the. MASC sink'</em> (48% masc.)</td>
<td>46-62% masculine</td>
</tr>
</tbody>
</table>
Common – High Consensus  
**at knif 'the.COM knife' (79% common)**  
68-79% common

Common – Low Consensus  
**at skaab 'the.COM closet' (50% common)**  
44-63% common

Unless otherwise noted, all items were presented as clipart images. To ensure that all stimuli were recognizable and familiar to community members, all clip-art images and picture stimuli were shown to ten adult native speakers on Föhr. Speakers wrote in the Frisian translation of the item. On average, 8.8 of the 10 speakers provided the same Frisian word for each picture (range 4-10). The most common type of other response were German borrowings. As much as possible, the number of items with matching grammatical gender in German was controlled across categories.

4.1.4 General Procedure

All participants were tested individually with a native speaker (Frisian-German bilingual) research assistant and the primary researcher. Students were tested at school in a quiet room. Elementary school students participated during class time and high school students participated after school. Adult participants were tested in a quiet, convenient location, most often in their homes. To ensure that even minimal-input participants understood the task procedure, all instructions were given in German by the native speaker research assistant.

Adult and high school participants completed all tasks in a one hour long session including the plural marking tasks (see Chapter 5) and general fluency tasks (see Chapter 6).

---

8 Because these items were selected from the adult online survey (see Chapter 3), the average consensus was lower for common gender items than for masculine gender items. This may represent a tendency to use masculine gender as a default, which is explored in the results section.
(6). The tasks were separated into two sessions for elementary school participants to ensure elementary school students remained attentive. The elementary school participants completed the second session within four and six and a half weeks after session one. Forty-five elementary school participants completed session one and of those a sub-test of 30 students completed session two as well. All participants completed the tasks in the same order. High school and adult participants completed session two tasks immediately following session one tasks. The order of tasks was as follows:

Session 1:
Task 1: Grammatical gender elicitation OR plural elicitation *
Task 2: Plural elicitation OR grammatical gender elicitation *
Task 3: General fluency story-telling

Session 2:
Task 4: Pronoun production task
Task 5: Pronoun comprehension task
Task 6: Plural acceptability task

The order of the definite article elicitation task and plural marking task was counterbalanced across participants.

Each task began with between three and seven practice items and during the practice items, and participants were given corrective feedback on the expected type of response, but only positive feedback on content. For example, for the pronoun production task, if a participant used a full noun phrase instead of the expected gender-marked pronoun, the participant was told that the puppet only understood if she used the correct format, and the research assistant modeled the correct format. However, if the participant
produced the correct format, but not the correct grammatical gender, only positive feedback was given, such as “Good job!”, or “Well done!” During the experimental items, the research assistant gave only positive feedback.

All stimuli were presented on a laptop using PowerPoint 2010, and all sessions were recorded using an external microphone and Audacity 2.01. If the participant or participant’s parent consented, the sessions were also video recorded. The primary researcher noted the participants’ responses, and these were later confirmed in consultation with the audio and video recordings. The responses to the gender elicitation task were transcribed by a native speaker. Children picked out a sticker half-way through each task and at the end of each task to ensure they remained focused. Adults were told when they were half-way through the task, but were not given a break or sticker. The experimental items, wording of instructions, and sample items are presented in Appendix B.

4.2 Definite article elicitation task

4.2.1 Background

The goal of the definite article elicitation task is to determine how the amount of Frisian input participants receive influences their production of Frisian grammatical gender. In this task, participants completed a picture naming task designed to elicit gender-marked definite articles. This task addresses each of the two broad research questions, though four specific research questions, which are numbered here to correspond to the broad research questions.
1a. Does the quantity of Frisian input participants receive influence how accurately participants use Frisian vocabulary?

1b: Does the quantity of Frisian input participants receive influence how often they produce masculine gender?

2a: Do elementary school students regularize grammatical gender?

2b. Does elementary school students’ pattern of regularization reflect adult-like production of grammatical gender?

The broad prediction is that participants with more Frisian input will show more adult-like production of grammatical gender. Specifically, I predict that participants with more Frisian input (high-input elementary school students and high school and adult participants) will produce fewer word errors, and that even when considering only accurate word responses, participants with more input will produce grammatical gender in a more adult-like manner. With regard to regularization, I predict that because grammatical gender is produced inconsistently, some elementary school students will show regularization, similar to the results of Hudson Kam and Newport (2005, 2009). However, I predict that high-input students will show less regularization than middle-input and minimal-input students. Furthermore, participants with cumulatively more Frisian input, that is, high-input high school students and adults, will show less regularization than high-input elementary school students.

4.2.2 Methods and design

This task elicits gender-marked definite articles using a design based on Schuele and Tolbert (2001). Participants saw two pictures on a computer screen, one of which
was circled. A research assistant told the participant that they would help the researcher learn Fering by helping her to fill out a worksheet. The participant was first asked to say what she saw on the computer screen. This elicited an indefinite article and entered the objects into the discourse, thus allowing for the felicitous use of the definite article in the follow-up question. The participant was then asked to tell the researcher which one of the objects was circled. A sample item is presented in Figure 4-1 and sample wording is presented in 1a-d. All participants completed three practice trials followed by 24 experimental trials.

![Figure 4-1. Stimulus for the Gender Production Task](image)

All items appeared twice in the task, once as the target item and once as the distractor item. The position (left or right) of the target item, and whether an item appeared first as a target item or a distractor item was counter balanced across four lists, and participants were randomly assigned to one of the four lists.

1a) Assistant:  *Wat schochst dü heer?*

   ‘What do you see here?’
That is, while the elicited production task was more structured than spontaneous speech, participants could produce word errors— that is, words other than the target word. Participants were given corrective feedback if they did not produce the expected form, meaning a full noun phrase (NP) including the definite article. However, if the participant produced any definite article and noun, even the wrong articles or nouns, they were always given positive feedback.

**Coding responses**

Responses were transcribed by a native speaker research assistant and then coded by the primary investigator for the type of word and the grammatical gender. Items were coded as either correct word responses or word errors. There are several types of word errors, including German translations of the Frisian word, such as the German word *Tasse* ‘cup’ instead of the expected Frisian word *kop*, phonologically adapted borrowings, such as *Tass* ‘cup’, based on the German form, *Tasse* ‘cup’, other Frisian words for similar concepts, such as *skütel* ‘small bowl’ in place of the expected form *dask* ‘large bowl’, and non-translation German words, such as *Feuerzeug* ‘lighter’ for the Frisian *swaavelstock* ‘match’. There were a small number of responses that did not fit
into these categories, including other language responses, such as Low German or English. Plural responses were categorized based on the single form, e.g., *tes* ‘teeth’ was counted as a correct vocabulary response, because the singular form *tus* ‘tooth’ was the correct word. The vocabulary accuracy analysis examines the frequency of word errors between groups of participants and categories.

4.2.3 Results

The results of the definite article elicitation task are separated into four sections. I first present the vocabulary accuracy, which compares the number of word errors in the gender production task. To isolate grammatical gender marking from vocabulary accuracy, I compare the production of masculine gender for only correct word responses. The percentage of masculine gender responses was selected as the dependent variable because it is difficult to determine what should be considered accurate responses in an inconsistent system. Therefore, rather than defining one grammatical gender as correct and the other as incorrect, the groups were compared based on the percentage of items produced with masculine gender. Next, following the methods in Hudson Kam and Newport (2005, 2009), the number of participants from each Home Language group who regularized inconsistent input was counted. Finally, the pattern of responses across items in each participant group was correlated with adult response patterns, to investigate whether the group responses reflect adult-like patterns of inconsistency. For each analysis, the influence of the amount of input was tested through two comparisons. The Home Language comparison compared elementary school students with different
amounts of Frisian input, and the Age Group comparison compared high-input bilinguals from three age groups.

**Vocabulary Accuracy**

Vocabulary accuracy refers to whether participants produced the correct word, independently of whether they produced the correct grammatical gender. The vocabulary accuracy analysis answers the first specific research question- 1a. Does the quantity of Frisian input participants receive influence how accurately participants use Frisian vocabulary? This analysis separates vocabulary acquisition from grammatical gender acquisition. As reported in the coding section (Section 4.2.2), participants did not always produce the correct vocabulary item. For example, some participants produced the German translation of the Frisian word, such as the German *Tasse* ‘cup’ instead of the expected Frisian word *kop*. Other responses were Frisian words for similar concepts, such as *skütel* ‘small bowl’ in place of the expected form *dask* ‘large bowl’. This analysis tests whether the frequency of word errors differs between participant groups and word categories.

**Home Language Comparison**

Three groups of elementary school students– high-input, middle-input, and minimal-input– were compared to investigate the role of quantity of Frisian input on the acquisition of Frisian vocabulary. Table 4-3 presents the descriptive results of the vocabulary accuracy by Home Language and Category. Examining the vocabulary accuracy across participants with different amounts of Frisian input, two patterns emerge. Students with more Frisian language input at home produced fewer word errors than
those with bilingual or no Frisian input at home. Secondly, across word categories, the rate of word errors is fairly similar. Overall, common gender items were produced with fewer word errors than masculine items, and high consensus items and low consensus items showed similar rates of word errors.

Table 4-3. Vocabulary accuracy by Home Language and Category

<table>
<thead>
<tr>
<th>Home Language</th>
<th>All</th>
<th>High</th>
<th>Low</th>
<th>High</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Items</td>
<td>Consensus</td>
<td>Consensus</td>
<td>Consensus</td>
<td>Consensus</td>
</tr>
<tr>
<td></td>
<td>$M(SD)$</td>
<td>$M(SD)$</td>
<td>$M(SD)$</td>
<td>$M(SD)$</td>
<td>$M(SD)$</td>
</tr>
<tr>
<td>High-Input</td>
<td>20</td>
<td>0.62 (0.09)</td>
<td>0.63 (0.09)</td>
<td>0.58 (0.13)</td>
<td>0.68 (0.16)</td>
</tr>
<tr>
<td>Middle-Input</td>
<td>12</td>
<td>0.31 (0.20)</td>
<td>0.40 (0.26)</td>
<td>0.24 (0.18)</td>
<td>0.32 (0.25)</td>
</tr>
<tr>
<td>Minimal-Input</td>
<td>13</td>
<td>0.20 (0.17)</td>
<td>0.27 (0.21)</td>
<td>0.19 (0.18)</td>
<td>0.18 (0.20)</td>
</tr>
</tbody>
</table>

A 3 X 2 X 2 mixed ANOVA was conducted using target word accuracy as the dependent variable to test for differences between home language groups and categories of items. In order to control for the number of participants per group, a sub-set of 12 participants per Home Language group were entered into the ANOVA. 12 participants from the high-input group were selected to match the middle-input group with respect to gender, grade in school (split into two age groups: 1\textsuperscript{st} and 2\textsuperscript{nd} grade versus 3\textsuperscript{rd} and 4\textsuperscript{th} grade), and school location (Föhr-West vs. Föhr-East). It was not possible to match participants with the
minimal-input group; thus, twelve minimal-input participants were randomly selected. The average vocabulary accuracy and standard deviations are presented in Table 4-4.

Home Language (High-input vs. Middle-input vs. Minimal-input) was entered as a between-subjects variable, and Consensus (High consensus vs. Low Consensus) and Grammatical Gender (Masculine vs. Common) were entered as within-subjects variables.

Table 4-4. Sub-set of vocabulary accuracy by Home Language and Category

<table>
<thead>
<tr>
<th>Home Language</th>
<th>All Items</th>
<th>High Consensus</th>
<th>Low Consensus</th>
<th>High Consensus</th>
<th>Low Consensus</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-input</td>
<td></td>
<td>0.63 (0.09)</td>
<td>0.58 (0.13)</td>
<td>0.71 (0.14)</td>
<td>0.61 (0.23)</td>
</tr>
<tr>
<td>N = 12</td>
<td></td>
<td>(0.08)</td>
<td>(0.13)</td>
<td>(0.14)</td>
<td>(0.23)</td>
</tr>
<tr>
<td>Middle-input</td>
<td></td>
<td>0.31 (0.20)</td>
<td>0.24 (0.18)</td>
<td>0.32 (0.25)</td>
<td>0.28 (0.19)</td>
</tr>
<tr>
<td>N = 12</td>
<td></td>
<td>(0.26)</td>
<td>(0.18)</td>
<td>(0.25)</td>
<td>(0.19)</td>
</tr>
<tr>
<td>Minimal-input</td>
<td></td>
<td>0.18 (0.17)</td>
<td>0.17 (0.16)</td>
<td>0.18 (0.21)</td>
<td>0.11 (0.15)</td>
</tr>
<tr>
<td>N = 12</td>
<td></td>
<td>(0.22)</td>
<td>(0.16)</td>
<td>(0.21)</td>
<td>(0.15)</td>
</tr>
</tbody>
</table>

The analysis revealed a main effect of Consensus, indicating that participants showed higher vocabulary accuracy on high consensus items than low consensus items: $F(1,33) = 16.851, p < .001, \eta^2 = .338$. There was not a main effect of Grammatical Gender, suggesting that participants showed similar vocabulary accuracy on masculine gender items as common gender items: $F(1, 33) = 1.00, p = .664$. In terms of Home Language, the analysis revealed that the amount of Frisian language input elementary
school students have influenced their vocabulary accuracy. There was a main effect of Home Language: $F(2, 33) = 25.253, p < .001$, partial $\eta^2 = .605$, with follow-up comparisons using Tukey HSD revealing that the high-input group had significantly higher vocabulary accuracy than the middle-input group ($p < .001$) and the minimal-input group ($p < .001$). However, the middle-input group and the minimal-input group did not differ significantly in vocabulary accuracy ($p = .133$) from each other.

These main effects were qualified by a significant interaction between Grammatical Gender and Home Language: $F(2, 33) = 3.334, p = .048$, partial $\eta^2 = .168$. Figure 4-2 presents the means for the Home Language X Grammatical Gender interaction. Pairwise comparisons suggest that the interaction was driven by marginally significant higher vocabulary accuracy on common gender items than masculine gender items in the High-input group ($p = .097$) and marginally significantly lower vocabulary accuracy on common gender items than masculine gender items for the minimal-input group ($p = .066$). Thus, while the simple effects are not significant in either direction, the interaction indicates that the simple effects are in opposite directions.
In order to test for the effects of cumulative language input, vocabulary accuracy was compared across high-input bilinguals in three age groups: Elementary school, High School, and Adults. The assumption is that since all participants have high-input, older participants will have more cumulative language input than younger participants. The general prediction is that participants with more language exposure will show higher vocabulary accuracy than those with less language exposure, with a specific prediction that adults and high school students will produce fewer word errors than high-input elementary school students.

Figure 4-2. Grammatical gender vocabulary accuracy: Home language comparison
As expected, the high school students and the adults showed higher vocabulary accuracy than high-input elementary school students. Table 4-5 presents the results of the descriptive statistics for each group and word category.

Table 4-5. Vocabulary accuracy by Age Group and Category

| Age Group          | All Items |  
|--------------------|-----------|---------|
|                    | Masculine | Common  |
|                    | M(SD)     | M(SD)   |
| High input Elementary N = 20 | 0.68 (0.09) | 0.63 (0.09) |
|                     | 0.58 (0.13) | 0.58 (0.21) |
| High School N = 9   | 0.75 (0.19) | 0.72 (0.25) |
|                     | 0.61 (0.25) | 0.87 (0.16) |
| Adult N = 13        | 0.83 (0.08) | 0.79 (0.12) |
|                     | 0.76 (0.13) | 0.91 (0.11) |

Similar to the Home Language comparison, a mixed 3 X 2 X 2 ANOVA was conducted to compare the vocabulary accuracy across categories and Age Group, using vocabulary accuracy as the dependent variable. The same sub-set of 12 high-input elementary school participants was used in the Age Group comparison as the Home Language comparison. The averages and standard deviations for the sub-set are presented in Table 4-6. Age Group (Elementary vs. High School vs. Adult) was entered as a between-subjects variable. Word Consensus (High consensus vs. Low consensus) and
Grammatical Gender (Masculine vs. Common) were entered as within-subjects’ variables.

Table 4-6. Vocabulary accuracy by Age Group and Category for Matched Participants

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Masculine</th>
<th>Common</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
<td>High</td>
</tr>
<tr>
<td>Items</td>
<td>Consensus</td>
<td>Consensus</td>
</tr>
<tr>
<td>M(SD)</td>
<td>M(SD)</td>
<td>M(SD)</td>
</tr>
<tr>
<td>High-input</td>
<td>0.63(0.09)</td>
<td>0.61(0.08)</td>
</tr>
<tr>
<td>High School</td>
<td>0.75(0.19)</td>
<td>0.72(0.25)</td>
</tr>
<tr>
<td>Adult</td>
<td>0.83(0.08)</td>
<td>0.79(0.12)</td>
</tr>
</tbody>
</table>

Analyses revealed a main effect of Consensus, with participants showing higher vocabulary accuracy on high consensus items than low consensus items: $F(1, 31) = 8.287, p = .007$, partial $\eta^2 = .211$. The main effect of Grammatical Gender was also significant, with participants showing higher vocabulary accuracy on common gender items than masculine gender items: $F(1, 31) = 20.486, p < .001$, partial $\eta^2 = .398$. The analysis showed that there was a significant effect of cumulative language input, as seen by the main effect of Age Group: $F(2, 31) = 8.616, p = .001$, partial $\eta^2 = .357$. Follow-up comparisons using Tukey HSD revealed that the main effect was driven by lower vocabulary accuracy among high-input elementary school students compared to adults ($p$
Vocabulary accuracy was not significantly different between High school students and Adults ($p = .334$). All interactions were non-significant: Consensus X Group: $F < 1$; Grammatical Gender X Group: $F(2, 31) = 1.319$, $p = .282$; Consensus X Grammatical Gender: $F < 1$; Consensus X Grammatical Gender X Group: $F < 1$.

To summarize both comparisons: For the elementary school students, the amount of Frisian input students have access to influences their vocabulary accuracy, with the high-input students producing fewer word errors than the middle-input and low-input. In general, participants showed higher vocabulary accuracy on high consensus items than low consensus items.

When comparing across Age Groups three general points emerge. First, the elementary school students were less accurate on vocabulary than either the high school students or the adults, secondly, high consensus items have higher vocabulary accuracy than low consensus items, and thirdly, common gender items exhibited higher vocabulary accuracy than masculine gender items. The differences in vocabulary accuracy between groups and between categories highlight the importance of examining gender production independently from vocabulary accuracy. For example, it is possible that the low vocabulary accuracy on low consensus words means that elementary school students—but not high school students or adults—were less familiar with the low consensus words. This is further evidence that word errors should be removed from the gender production analysis. As such, the gender production analysis should remove any word errors and only investigate grammatical gender production when the participant produces the correct word.
Gender Production

Transitioning from vocabulary acquisition to grammatical gender acquisition, I now test whether bilinguals with different levels of language input produce grammatical gender differently. The gender production analysis addresses the second specific research question: Does the quantity of Frisian input participants receive influence how often they produce masculine gender? Grammatical gender accuracy is difficult to determine in Fering when grammatical gender is produced inconsistently. For example, the online survey results show that adult native speakers do not always agree with each other or with the dictionary. With this issue in mind, the percentage of masculine gender responses will be analyzed, rather than assuming a correct grammatical gender for a given noun. The percentage of masculine gender responses was calculated for each category of items (high-consensus masculine, low-consensus masculine, high-consensus common, and low-consensus common). Recall that for the categories of items, the gender was determined by the adult online survey (refer to Section 4.1.3 for a description of categories).

As the vocabulary accuracy analysis showed significant differences in the production of word errors between groups with different amounts of Frisian input, all word errors were removed from the gender production analysis. This avoids a confound between vocabulary acquisition and grammatical gender acquisition. Two comparisons were conducted to test for the effects of quantity of language exposure.
**Home Language Comparison**

In order to test the impact of quantity of input, the production of masculine gender was compared in three groups of elementary school students—high-input, middle-input, and minimal-input. The percentage of masculine gender responses was compared for each Home Language group and word category. All word errors were removed from this analysis. Three middle-input were removed due to overall low vocabulary accuracy, as they did not produce any correct word responses. In the minimal-input group the number of participants removed due to low vocabulary accuracy is different across word categories (see Table 4-7 for the number of participants by group). Table 4-7 presents each Home Language group’s percentage of masculine responses split by grammatical gender and level of consensus. As Table 4-7 shows, across all Home Language groups, elementary school students use masculine gender for high consensus masculine items almost exclusively. Overall, the high-input and minimal-input participants produce masculine gender responses at a similar rate, while the middle-input participants appear to favor the masculine gender. Importantly, this analysis is not able to determine whether participants are producing grammatical gender in an adult-like manner, just the general tendency to mark masculine grammatical gender.
Table 4-7. Proportion of Masculine gender responses for correct vocabulary responses by Home Language and Category

<table>
<thead>
<tr>
<th>Home Language</th>
<th>All Items</th>
<th>High Consensus</th>
<th>Low Consensus</th>
<th>High Consensus</th>
<th>Low Consensus</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-input</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N = 20</td>
<td></td>
<td>0.68 (0.27)</td>
<td>0.94 (0.23)</td>
<td>0.53 (0.33)</td>
<td>0.73 (0.34)</td>
</tr>
<tr>
<td>Middle-input</td>
<td></td>
<td>0.79 (0.29)</td>
<td>0.85 (0.19)</td>
<td>0.78 (0.36)</td>
<td>0.83 (0.26)</td>
</tr>
<tr>
<td>N = 9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minimal-input</td>
<td></td>
<td>0.68 (0.31)</td>
<td>0.83 (0.26)</td>
<td>0.43 (0.46)</td>
<td>0.50 (0.47)</td>
</tr>
<tr>
<td>N = 10</td>
<td></td>
<td></td>
<td>(N = 9)</td>
<td>(N = 9)</td>
<td>(N = 8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(N = 10)</td>
<td></td>
<td>(N = 6)</td>
</tr>
</tbody>
</table>

A mixed 3 X 2 X 2 ANOVA was conducted to compare the percentage of masculine responses on correct vocabulary items. The same sub-set of 12 matched participants was used as in the vocabulary accuracy analysis was entered into the ANOVA. The ANOVA procedure removes participants who had no responses in any one category. The descriptive statistics for the remaining participants is presented in Table 4-8. All incorrect vocabulary responses were excluded, as were responses that were not marked for grammatical gender, such as indefinite articles or bare nouns. The remaining correct vocabulary responses were coded as either masculine gender responses, or common gender responses. The definite articles, *de ‘the.MASC’* and *a ‘the.MASC’*, were
considered masculine responses, and *det ‘the.COM’ and at ‘the.COM’ were considered common gender responses. The percentage of masculine gender responses in each category of words was entered as the dependent variable. Home Language (High-input vs. Middle-input vs. Minimal-input) was entered as a between-subjects variable. The within-subjects variables were: Consensus (High consensus vs. Low Consensus) and Grammatical Gender (Masculine vs. Common).

Table 4-8. Sub-set of gender production by Home Language and Category

<table>
<thead>
<tr>
<th>Home Language</th>
<th>Consensus</th>
<th>Home Language</th>
<th>Consensus</th>
<th>Consensus</th>
<th>Consensus</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Masculine</td>
<td></td>
<td></td>
<td>Common</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M(SD)</td>
<td>M(SD)</td>
<td>M(SD)</td>
<td>M(SD)</td>
<td></td>
</tr>
<tr>
<td>High-input</td>
<td>0.90 (0.29)</td>
<td>0.60 (0.43)</td>
<td>0.54 (0.33)</td>
<td>0.69 (0.35)</td>
<td></td>
</tr>
<tr>
<td>N = 12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle-input</td>
<td>0.85 (0.19)</td>
<td>0.73 (0.42)</td>
<td>0.78 (0.36)</td>
<td>0.83 (0.26)</td>
<td></td>
</tr>
<tr>
<td>N = 9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minimal-input</td>
<td>0.77 (0.33)</td>
<td>0.60 (0.43)</td>
<td>0.30 (0.45)</td>
<td>0.60 (0.42)</td>
<td></td>
</tr>
<tr>
<td>N = 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The analysis revealed that participants used masculine gender more often for high consensus items than low consensus items. There was a main effect of Consensus: $F(1, 23) = 4.667, p = .041$, partial $\eta^2 = .169$. There was no main effect of Grammatical Gender ($F < 1$) or Home Language: $F(2, 23) = 1.047, p = .367$. There were no significant interactions between Home Language: Consensus X Group: $F(2, 23) = 1.697, p = .205$;
Grammatical Gender X Group: $F(2, 23) = 1.230, p = .311$; Consensus X Grammatical Gender X Group: $F < 1$. The main effect of Consensus was qualified by a significant interaction between Grammatical Gender and Consensus: $F(1, 23) = 11.046, p = .003$, partial $\eta^2 = .324$. Figure 4-3 illustrates the percentage of masculine responses by category.

![Grammatical gender production: Home language comparison](image)

Figure 4-3. Grammatical gender production: Home language comparison

The interaction between Grammatical gender and Consensus suggests that elementary school students from all Home Language groups favor masculine gender for masculine gender items more for high consensus items than for low consensus items. For common gender items, elementary school students use masculine gender less often for high consensus items than low consensus items.
Age Group

In order to further test the role of the amount of input, three groups of high-input bilinguals from three age groups were compared. Similar to the vocabulary accuracy analysis, the assumption is that since all participants have high input, older participants, that is, high school students and adults, have cumulatively more Frisian input than elementary school students. As with the Home Language comparison, the three age groups were compared on their production of masculine gender on correct vocabulary responses. Incorrect vocabulary responses and responses with no gender marking, such as indefinite determiners, and bare nouns were removed. Table 4-9 presents the frequency of masculine gender responses for high-input elementary school students, high school students, and adults for each word category. Recall that these results reflect how often participants use masculine gender. This means that I would expect to see high percentages of masculine responses for masculine words, and low percentages of masculine responses for common gender items. Overall, as predicted, elementary school students tend to use the masculine gender more often than high school students or adults.
Table 4-9. Masculine gender responses by Age Group and Category

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Masculine</th>
<th></th>
<th>Common</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
<td>High</td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Items</td>
<td>M(SD)</td>
<td>M(SD)</td>
<td>M(SD)</td>
<td>M(SD)</td>
</tr>
<tr>
<td>High-input Elementary N = 20</td>
<td>0.68</td>
<td>0.94 (0.23)</td>
<td>0.53 (0.33)</td>
<td>0.53 (0.42)</td>
</tr>
<tr>
<td>High School N = 9</td>
<td>0.54</td>
<td>0.89 (0.11)</td>
<td>0.63 (0.29)</td>
<td>0.23 (0.36)</td>
</tr>
<tr>
<td>Adult N = 13</td>
<td>0.43</td>
<td>0.80 (0.23)</td>
<td>0.34 (0.23)</td>
<td>0.25 (0.24)</td>
</tr>
</tbody>
</table>

A mixed 3 X 2 X 2 ANOVA was conducted to compare the frequency of masculine gender responses. The same 12 high-input elementary school students were included as in the Home Language comparison. Table 4-10 presents the descriptive statistics which were entered into the ANOVA. The percentage of masculine gender responses was entered as the dependent variable. Age Group (High-input Elementary vs. High School vs. Adults) was entered as a between-subjects variable. The within-subjects variables were Consensus (High vs. Low) and Grammatical Gender (Masculine vs. Common).
### Table 4-10. Sub-set of gender responses by Age Group and Category

<table>
<thead>
<tr>
<th>Age Group</th>
<th>All Items</th>
<th>High Consensus</th>
<th>Low Consensus</th>
<th>High Consensus</th>
<th>Low Consensus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Juvenile</td>
<td>0.69</td>
<td>0.90 (0.29)</td>
<td>0.54 (0.33)</td>
<td>0.60 (0.43)</td>
<td>0.69 (0.35)</td>
</tr>
<tr>
<td>Elementary N = 12</td>
<td></td>
<td>(0.30)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High School N = 9</td>
<td>0.54</td>
<td>0.89 (0.11)</td>
<td>0.63 (0.29)</td>
<td>0.23 (0.36)</td>
<td>0.53 (0.28)</td>
</tr>
<tr>
<td></td>
<td>(0.21)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adult N = 13</td>
<td>0.43</td>
<td>0.80 (0.23)</td>
<td>0.34 (0.23)</td>
<td>0.25 (0.24)</td>
<td>0.38 (0.27)</td>
</tr>
<tr>
<td></td>
<td>(0.19)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The analysis revealed that the high consensus items were more often marked with masculine gender than low consensus items. There was a main effect of Consensus: $F(1, 31) = 7.178, p = .012$, partial $\eta^2 = .188$, suggesting that high-consensus words were marked with masculine gender than low-consensus words. There was also a main effect of Grammatical gender: $F(1, 31) = 48.595, p < .001$, partial $\eta^2 = .611$, suggesting that masculine gender words were marked with masculine gender significantly more often than common gender items. The main effect of Age Group was marginally significant: $F(2, 31) = 3.254, p = .054$, partial $\eta^2 = .173$, and follow-up comparisons using Tukey HSD revealed that the elementary schools students used masculine gender significantly more often than the adults ($p = .041$), but not significantly more often than the high
school participants ($p = .511$). The adult participants and high school participants did not differ in their production of masculine responses ($p = .452$).

These main effects were qualified by a significant interaction between Grammatical Gender and Consensus. This suggests that the percentage of masculine responses depends on both the Grammatical Gender and the Level of Consensus ($F(1, 31) = 2.314, p < .001$, partial $\eta^2 = .597$. The significant interaction suggests that the participants use masculine gender less often for high consensus common gender items, and use masculine gender more often for low consensus common gender items.

There was also significant interaction between Grammatical Gender and Age Group: $F(2, 31) = 6.852, p = .003$, partial $\eta^2 = .307$. The estimated marginal means for the interaction are presented in Figure 4-4.

![Grammatical gender production: Age group comparison](image)
Follow-up pairwise comparisons of the estimated marginal means revealed that participants in both the High School and Adult groups, the frequency of masculine gender responses was significantly different by grammatical gender (both $p < .001$). However, the high-input elementary school students did not show differences in the production of masculine gender between masculine and common gender categories ($p = .225$). Both the High School and Adult groups produced significantly more masculine gender responses for masculine items than common gender items, but elementary school students produced similar rates of masculine responses for both sets of items. This suggests the elementary school students are not sensitive to differences in grammatical gender. However, this global similarity in the rates of masculine gender production cannot address whether students have acquired the lexically determined pattern.

To summarize the gender production comparisons: For elementary school students– when considering only correct vocabulary responses– elementary school students from all Home Language groups tend to use masculine gender similarly. Overall, elementary school students use masculine gender more often for high consensus items than low consensus items. The effect of Consensus, however, depends on the grammatical gender of the item, with masculine gender words showing more of an effect of consensus than common gender words.

When comparing high-input bilinguals from three age groups, the participants produced masculine gender more often for masculine items than common gender items. Moreover, they produced masculine gender more often for high consensus items than low consensus items. The effect of Grammatical Gender was more pronounced for high
school students and adults, while the elementary school students produced more similar responses regardless of the grammatical gender of the item.

Across both comparisons, the results show that participants with less Frisian input tended to use masculine gender more often than participants with more Frisian input. That is, while it was not statistically significant, within elementary school students, middle-input students appeared to use masculine gender more often than high-input students. Similarly, within the age group comparison, the high-input elementary school students— who have cumulatively less input than high school students or adults— tended to use masculine gender more often than the participants with cumulatively more input. This tendency to rely on masculine gender may represent regularization, similar to the results of Hudson Kam and Newport (2005, 2009). This is the focus of the following section.

**Regularization**

The second broad research question was whether participants regularize inconsistently produced morphological markings. That is, when participants are given inconsistent morphology, such as grammatical gender in Frisian, do they over-generalize and produce predominately one grammatical gender marking? To test whether elementary school students regularize inconsistent input, the number of participants who show regularization were counted. Following Hudson Kam and Newport (2005, 2009), a strict criterion for regularization was used. Participants who use the same grammatical gender for all, or all but one, possible responses were categorized as *regularized speakers*. Inconsistent speakers use both grammatical genders, and do not regularize either grammatical gender.
All responses with a Frisian determiner (*de*, *det*, *a*, or *at*), including word errors, factored into this analysis. While *de* ‘the.MASC’ was considered a Frisian determiner, this is complicated by the fact that these responses may in fact be reduced German determiners. German determiners—der ‘the.MASC’ and *die* ‘the.FEM’—can be phonologically reduced, resulting in a form which is phonologically identical to the Frisian masculine determiner, *de* [də]. However, a number of factors make it difficult to determine whether these are reduced German articles or in fact Frisian articles. First, intense contact between German and Frisian on Föhr has led to many borrowings. This blurs the line between “German words” and “Frisian words”. Some borrowings that violated Frisian phonotactics have undergone phonological adaptation, such as *tasch* ‘purse’ which was borrowed from German *Tasche* ‘purse’. However, not all borrowings contain a phonotactic violation, and therefore there was no need for them to be altered.

Given the difficulty in determining whether responses are in fact German words or Frisian words, I consider all *de* responses as masculine Frisian responses. For the purpose of the regularization analysis, I will not attempt to determine whether responses are considered Frisian words or German words, and instead I rely simply on the determiner. Responses with *de*, *det*, *a*, or *at* were considered Frisian articles regardless of the word response. This means that if a word error response contained a Frisian determiner, such as *de skütel* ‘the.MASC little bowl’ in place of the target word *de dask* ‘the.MASC large bowl’, this was still included as a masculine response for the purpose of this analysis.

Participants with very low vocabulary accuracy were removed from the regularization analysis, because they did not produce enough words to confidently
determine whether they are regularized speakers or inconsistent speakers. If they had not been removed, a participant who only produced two words with Frisian articles but who used masculine gender both times would have been considered a regularized speaker. Thus, participants who produced fewer than half of the items with Frisian determiners were included in this analysis. Of the 45 elementary school students, nine did not produce a minimum of 12 of 24 items with a Frisian definite article, and were therefore excluded from the analysis. Responses without Frisian definite articles are predominantly German responses with a German definite article, but also include indefinite articles, and bare noun responses. Regularization was determined using the same strict criteria as Hudson Kam and Newport (2009).

All but one consistent speaker used the masculine definite article for both masculine and common gender nouns. As seen in Table 4-11, some of the elementary school students regularize Frisian grammatical gender. However, the amount of Frisian input influences the percentage of participants who regularize their input. Just under half of the high-input participants regularize grammatical gender, as compared to three-quarters of middle-input participants, while minimal-input students regularize grammatical gender least often of all home language groups. However, a Chi-square test did not reveal any significant differences between the number of regularized speakers in the different home language groups ($\chi^2 (2, N = 36) = 4.410, p = .110$).\(^7\)

\(^7\) All Chi-square comparisons were adjusted for multiple comparisons using the Bonferroni adjustment. The adjusted $\alpha$ level for comparisons with three levels was $p = .017$, and $p = .025$ for follow-up two-way comparisons.

\(^8\) For the purposes of comparisons, regularized masculine speakers and regularized common speakers were combined.
Table 4-11. Regularized Speakers by Home Language Group

<table>
<thead>
<tr>
<th>Home Language</th>
<th>Regularized Masculine Speaker</th>
<th>Regularized Common Speaker</th>
<th>Inconsistent Speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-input N = 20</td>
<td>7 (35%)</td>
<td>1 (5%)</td>
<td>12 (60%)</td>
</tr>
<tr>
<td>Middle-input N = 8</td>
<td>6 (75%)</td>
<td>0 (0%)</td>
<td>2 (25%)</td>
</tr>
<tr>
<td>Minimal-input N = 8</td>
<td>2 (25%)</td>
<td>0 (0%)</td>
<td>6 (75%)</td>
</tr>
</tbody>
</table>

As with the vocabulary accuracy and gender production, groups of high-input bilinguals at three age groups were compared to test for the effects of cumulative language input. As Table 4-12 shows, while some of the high-input elementary school students regularize Frisian grammatical gender, only one of the nine high school participants shows regularization. All adult participants produce Frisian grammatical gender inconsistently.

A Chi-square test revealed that there was a significant difference in how many participants in each Age Group regularized grammatical gender: $\chi^2 (2, N = 42) = 8.211, p = .016$. Follow-up Chi-square tests comparing each of the Age Groups showed that the overall difference in the number of participants who regularize was driven by significantly more regularized speakers in the high-input elementary school student group compared to the adults ($\chi^2 = 6.864, p = .009$), but no significant difference between elementary school students and high school students ($p = .120$) or between high school students and adults ($p = .219$).
Table 4-12. Regularized Speakers by Age Group

<table>
<thead>
<tr>
<th>Home Language</th>
<th>Regularized Masculine Speaker</th>
<th>Regularized Common Speaker</th>
<th>Inconsistent Speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-input</td>
<td>7 (35%)</td>
<td>1 (5%)</td>
<td>12 (60%)</td>
</tr>
<tr>
<td>Elementary N = 20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High School N = 9</td>
<td>1 (11%)</td>
<td>0 (0%)</td>
<td>8 (89%)</td>
</tr>
<tr>
<td>Adult N = 13</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>13 (100%)</td>
</tr>
</tbody>
</table>

To summarize: Following Hudson Kam and Newport (2005, 2009) the regularization analysis counted the number of speakers in each group who use the one grammatical gender for all or all but one response. The results show that while some of the elementary school students regularize grammatical gender, similar to the results from Hudson Kam and Newport (2005, 2009), as predicted, fewer students with high Frisian input regularized grammatical gender markings. This general trend held true for the Home Language comparison with the high-input students regularizing less often than the middle-input students, and the Age Group comparison with only one of the high school students and no adults regularizing grammatical gender. Interestingly, the minimal-input appear to regularize the least. It is important to keep in mind, however, that the regularization analysis cannot distinguish adult-like inconsistency and random inconsistency. This is the focus of the following section.
Adult-like Inconsistency

Adult speakers of Fering produce grammatical gender inconsistently. However, not all inconsistent speakers are adult-like speakers. There are two possible types of inconsistent speakers: adult-like inconsistent speakers, and randomly inconsistent speakers. This means that it is not enough for speakers to produce grammatical gender inconsistently; they must produce the specific pattern of adult-like inconsistency. Recall that some nouns are produced fairly consistently by adult speakers, while others show much higher inconsistency. This results in a specific by-item pattern of adult grammatical gender production. Thus, if elementary school students show adult-like inconsistency, for each item, when adults produce mostly masculine gender, the elementary school students would also produce mostly masculine gender, and when adults produce mostly common gender, the elementary school students would also produce mostly common gender.

To test whether each groups’ pattern of results reflects adult-like inconsistency, a set of correlations between adult responses and each group of elementary school students was conducted. The correlations include the same data as in the regularization analysis, meaning that word errors with Frisian articles were included in this analysis. The percentage of masculine responses by item was used as the dependent variable. For each item the average percentage of speakers who produced a masculine gender response was calculated. Figure 4-6 shows the scatterplots of each group’s percentage of masculine responses compared to adults’ responses to the same item. Spearman’s rho correlations revealed a significantly positive correlation between the high-school student responses and the adult responses ($r_s = .717, N = 24, p < .001$, two-tailed) suggesting that their pattern of responses follows the adult-like pattern. Within elementary school students, the
high-input students show a strong significant correlation with the adult responses: \( r_s = .686, N = 24, p < .001, \) two-tailed), and a significant, but slightly less strong correlation between the middle-input students and adults \( r_s = .423, N = 24, p = .039, \) two tailed). There was no correlation between the minimal-input students and the adults: \( r_s = .347, N = 24, p = .097, \) two-tailed, suggesting that their pattern of responses was not adult-like.

One limitation of the previous analysis is that it potentially compares the grammatical gender of different word responses to each other. For example, if some participants produced a word error, such as skütel ‘small bowl’ in place of the correct word response, dask ‘large bowl’, the percentage of masculine responses for skütel ‘small bowl’ would be compared to the percentage of masculine responses for dask ‘large bowl’. To address this concern, a follow-up analysis was conducted using only percentage of masculine responses for only correct words; all word errors were removed from this analysis. The prediction is that participants with more Frisian input (high-input elementary and high school students) will show significant correlations with adult responses, meaning that they show adult-like patterns of inconsistency, while lower-input participants (middle-input and minimal-input students) will not correlate with adult responses.

The follow-up correlations on only correct word responses using Spearman’s rho correlations showed that high school students’ responses show the strongest correlation with adult responses \( r_s = .663, N = 24, p < .001, \) two tailed). Within the elementary school students, the high-input students’ responses were significantly correlated with the adult production by item \( r_s = .613, N = 23, p = .002, \) two tailed). However, neither the middle-input nor the minimal-input students’ responses were significantly correlated with
adult responses (Middle-input $r_s = .377$, $N = 20$, $p = .102$, two tailed; Minimal-input $r_s = .454$, $N = 15$, $p = .089$, two tailed). The scatterplots for each of the groups compared to the adult responses is presented in Figure 4-6.

To summarize: While the regularization analysis was not able to distinguish between adult-like inconsistent speakers and randomly inconsistent speakers, a set of correlations was conducted between adults responses and each groups’ responses to test for adult-like production of grammatical gender. The results show that when considering all responses, including word errors, the high school students, high-input students, and the middle-input students’ responses are correlated with the adult responses, but the minimal-input students’ responses are not significantly correlated with the adults’ responses. When considering only correct word responses, and thus more directly comparing the responses on each item, the high-input students and high school students show an adult-like pattern of responses, while the middle-input and minimal-input students were not significantly correlated with adult responses. That is, while regularization analysis suggested that both the high-input students and minimal-input students produce grammatical gender in an adult-like manner- the high-input students produce an adult-like pattern of inconsistency, while the minimal-input students produce random inconsistency.
Figure 4-5. Scatterplots of masculine responses by Input Group compared to Adult responses by item regardless of word accuracy.
Figure 4-6. Scatterplots of masculine responses by Input Group compared to Adult responses by item for correct word responses only.
4.2.4 Discussion

The definite article task tested participants’ ability to produce adult-like patterns of grammatical gender markings. This chapter’s two broad research questions were addressed through four more specific analyses. First, the number of word errors was compared, and then the groups were compared based on their frequency of masculine gender responses. Next, the number of speakers who regularized the inconsistent input was counted, and finally, correlation analyses tested whether speakers whose responses at the group level were more inconsistent showed adult-like inconsistency or random inconsistency. To investigate how the amount of input influences grammatical gender acquisition, two comparisons were made for each analysis.

The results of the four sub-analyses provide converging evidence that while some participants regularize grammatical gender, as predicted by Hudson Kam and Newport (2005, 2009), when they are given substantial Frisian input, they are able to overcome regularization and faithfully produce adult-like patterns of inconsistency. Across both comparisons, participants with more Frisian input make fewer word errors and produce grammatical gender in a more adult-like way. High-school students, who have cumulatively higher input than high-input elementary school students, produce significantly fewer word errors and show a strong correlation with adult-like patterns of grammatical gender production. Within elementary school students, the high-input students produced the fewest word errors, and are the only group to show a significant
correlation with adult production of grammatical gender when considering only correct word responses.

While the gender production analysis shows a similar pattern, the results of the Home Language comparison did not show significant differences. The interpretation of these results is limited by the small sample sizes, especially in the gender production analysis. Because many of the participants with low input frequently produced word errors, the number of trials remaining for their results in the gender production task is low. This limitation of statistical power may be why there is not a statistically significant difference between groups in the Home Language comparison for the gender production analysis. Another explanation is the overall similar frequency of masculine gender responses among high-input and minimal-input students. While the groups have overall similar rates of masculine gender responses, the correlation analysis showed that only the high-input students show adult-like patterns.

Despite these limitations, the general finding is that participants are able to successfully acquire inconsistent features, although substantial input is required in order to overcome the initial stage of regularization. This finding is particularly important in a minority language context, where children may not have sufficient access to the minority language in order to successfully acquire difficult features, such as inconsistent morphology. Recall that middle-input students mostly hear Frisian spoken at home about 50% of the time and most began learning Frisian at birth. However, even these students do not have enough input in order to overcome a tendency towards regularization. This suggests that these students need additional input, either at home, in school or in the community, if they are to successfully acquire grammatical gender in Fering.
4.2.5 Summary

The definite article elicitation task tested participants’ ability to associate nouns with a given grammatical gender. The adult responses support the pilot study findings that grammatical gender is produced inconsistently in contemporary Fering, and established a base-line pattern of adult-like inconsistency. The results were analyzed using four sub-analyses: first the rates of word errors were investigated for each group, then the production of grammatical gender. Turning to the second broad research question, the number of speakers who regularized grammatical gender markings was counted, finally a set of correlations was conducted with adult responses in order to distinguish between adult-like inconsistent speakers and randomly inconsistent speakers.

A general pattern emerges across both the Home Language and Age Group comparisons: participants with more Frisian input show more adult-like responses. The high-input group did not differ from the adult group in terms of the number of word errors or their production of masculine gender. Further, only one high-school student regularized grammatical gender, and as a group, the high school students showed adult-like inconsistency. Within the elementary school participants, the high-input group produced fewer word errors and showed adult-like inconsistency. The middle-input group showed the highest level of regularization, and while the minimal-input group showed low levels of regularization, their pattern of inconsistency was random, not adult-like. Taken together, the results support the broad prediction that while some elementary school students would regularize grammatical gender – because they receive inconsistent input – that regularization may reflect the initial stages of learning, as participants with more input are able to faithfully acquire the adult-like pattern of inconsistency.
4.3 Pronoun Production Task

4.3.1 Background

The pronoun production task investigates grammatical gender acquisition by focusing on the production of gender-marked pronouns. In this task, participants were provided with grammatical gender information in the prompt and then asked to produce a gender-marked pronoun which agreed with the antecedent. If participants attend to the grammatical gender given in the task, they can complete this task successfully without knowing the grammatical gender for any specific nouns. Thus, the results reflect participant’s knowledge of grammatical gender independent of the gender assignment for specific nouns.

Following the broad research questions, the specific research questions are:

1a. Does the quantity of Frisian input participants receive influence how often participants use gender-marked pronouns?

1b. Does the quantity of Frisian input participants receive influence how accurately participants mark grammatical gender on gender-marked pronouns?

2a. Do participants regularize grammatical gender markings on gender-marked pronouns?

The broad prediction is that participants with more Frisian input will produce gender markings on pronouns more often, and that when they do use gender-marked pronouns, they will be more accurate in gender agreement. Specifically, this means that I predict high-input elementary school students to produce more gender-marked pronouns than middle-input or minimal-input peers, and to have higher accuracy on gender-marked
pronouns. Testing the effects of cumulative language input, I predict that high school students and adults will produce more gender-marked pronouns and mark gender more accurately than high-input elementary school students.

4.3.2 Methods and design

For this task participants were told that they would be seeing what people are doing, but that a puppet, Kurt, cannot see colors. The participants were asked to help Kurt by telling him what color the objects were. See Figure 4-8 for an example and 2 for sample wording. Participants were instructed that Kurt can only understand simple sentences including a pronoun. Participants first saw a black and white line drawing of either a boy or girl performing an action along with a recorded description. In the next slide, the participant saw the same drawing with the target item in color. Participants were then prompted to tell the puppet what color the target item was. Participants were given grammatical gender information twice, once during the description of the action and once during the prompt. Participants completed three practice items and received corrective feedback on format during the practice section. For example, if a participant produced a full noun phrase or only a color during a practice item, they were reminded that the puppet can only understand when they use the gender-marked pronoun, and the practice item was modeled with the correct format. If the participant produced a gender-marked pronoun, positive feedback was given regardless of whether it was the correct gender-marked pronoun.

‘here is another girl. What is she doing? She is buying a purse. What color is the.\text{MASC} purse? Tell Kurt, what color the.\text{MASC} purse is.’

Child: \textit{hi as gül}

‘It.\text{MASC} is yellow’

Figure 4-7. Stimulus for Pronoun Production Task

The pronoun production task was the first task of the second session, meaning that 13 high-input, 11 middle-input, and 6 minimal-input students completed this task. The same nine high school students and 13 adults as in the gender production task completed this pronoun production task. All participants will be included in all of the comparisons for this task, as the difference in the number of participants was much smaller on this task than the gender production task.

4.3.3 Results

The results are divided into three sections. First, the responses were coded for the form accuracy, that is, whether participants produced a gender-marked pronoun. The
Form accuracy analysis is similar to the vocabulary accuracy analysis for the definite article elicitation task in that it tests whether participants are using the correct format, rather than testing grammatical gender acquisition. Then turning to grammatical gender acquisition, the grammatical gender accuracy was compared for participants with differing amounts of Frisian input. Finally, the frequency of all responses was analyzed in order to test for grammatical gender regularization in gender-marked pronouns. As with the gender production task, the influence of the quantity of input will be tested through two comparisons—a home language comparison and an age group comparison.

**Form Accuracy**

Form accuracy refers to whether a response contained the correct form, that is, a gender-marked pronoun. Responses that included a gender-marked pronoun, *hi* ‘he’ or *hat* ‘she’ were considered correct forms, while other forms, including definite articles, such as *a* ‘the.Masc’ or *at* ‘the.Com’ were considered form errors. In contrast to vocabulary accuracy on the definite article production task (Section 4.2), the form accuracy shows whether the participants produce a pronoun, regardless of the vocabulary accuracy or grammatical gender.

For the Home Language comparison, Table 4-13 presents the percentage of form accuracy for items within each word category for each group of elementary school students. Overall, elementary school students showed low form accuracy, meaning that they rarely produced gender-marked pronouns. A one-way ANOVA was used to test for differences in form accuracy by Home Language. The responses from all word categories were included and all participants were included. Form accuracy was entered as the
dependent variable and Home Language was entered as the between-subjects variable.

The analysis did not reveal a main effect of Home Language, suggesting that the elementary school students showed low form accuracy, regardless of the amount of Frisian input: $F(2, 29) < 1$.

Table 4-13. Form accuracy by Home Language and Category

<table>
<thead>
<tr>
<th>Home Language</th>
<th>All Items</th>
<th>High Consensus</th>
<th>Low Consensus</th>
<th>High Consensus</th>
<th>Low Consensus</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-input</td>
<td>0.23</td>
<td>0.24 (0.43)</td>
<td>0.23 (0.44)</td>
<td>0.23 (0.44)</td>
<td>0.22 (0.42)</td>
</tr>
<tr>
<td>$N = 13$</td>
<td>(0.43)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle-input</td>
<td>0.19</td>
<td>0.18 (0.37)</td>
<td>0.18 (0.36)</td>
<td>0.21 (0.40)</td>
<td>0.20 (0.37)</td>
</tr>
<tr>
<td>$N = 11$</td>
<td>(0.37)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minimal-input</td>
<td>0.17</td>
<td>0.17 (0.33)</td>
<td>0.22 (0.39)</td>
<td>0.17 (0.41)</td>
<td>0.14 (0.34)</td>
</tr>
<tr>
<td>$N = 6$</td>
<td>(0.37)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To test for the impact of cumulative high input, the form accuracy of high-input bilinguals from three age groups—elementary school, high school, and adults, was compared. Table 4-14 presents form accuracy for each age group of high-input bilinguals. Similar to the Home Language comparison, overall form accuracy was low for the high-input bilinguals. Adult speakers produced a gender-marked pronoun for only 35% of the items. While the high school and adult participants produced slightly more gender-
marked pronouns than elementary school participants, all groups produced similar rates of pronouns across categories, as revealed in a one-way ANOVA with form accuracy as the dependent variable and Age Group as a between-subjects variable: $F(2, 34) < 1$.

There were no other significant effects or interactions.

Table 4-14. Form accuracy by Age Group and Category

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Masculine All Items</th>
<th>Masculine High Consensus</th>
<th>Masculine Low Consensus</th>
<th>Common High Consensus</th>
<th>Common Low Consensus</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-input</td>
<td>0.23 (0.43)</td>
<td>0.24 (0.43)</td>
<td>0.23 (0.44)</td>
<td>0.23 (0.44)</td>
<td>0.22 (0.42)</td>
</tr>
<tr>
<td>Elementary $N = 13$</td>
<td>0.36 (0.46)</td>
<td>0.39 (0.47)</td>
<td>0.38 (0.48)</td>
<td>0.35 (0.49)</td>
<td>0.33 (0.43)</td>
</tr>
<tr>
<td>Adult $N = 13$</td>
<td>0.35 (0.40)</td>
<td>0.45 (0.46)</td>
<td>0.32 (0.41)</td>
<td>0.28 (0.43)</td>
<td>0.33 (0.41)</td>
</tr>
</tbody>
</table>

The overall averages, however, obscure a bifurcated pattern of results. Thirty-two of the 52 participants never produced more than one gender-marked pronoun, while an additional 12 participants almost always produced the gender-marked pronoun (22 or more of the 24 items with a gender-marked pronoun). That is, 85% of participants either used gender-marked pronouns for nearly every item or they did not use a gender-marked pronoun for any item. The participants who never produced a gender-marked pronoun are evenly distributed across home language groups (62% ($N = 8$) High-input, 64% ($N = 7$) Middle-input, and 50% ($N = 3$) Minimal-input). These results show that the majority of
participants are either always using a gender-marked pronoun or never using gender-marked pronouns, but that the amount of Frisian input, either within elementary school or the cumulative amount of input, does not influence how likely participants are to use gender-marked pronouns.

**Gender-Marked Pronoun Production Accuracy**

I now turn to grammatical gender accuracy, rather than form accuracy. While in general it is difficult to determine the grammatical gender for a given noun, grammatical gender accuracy was less problematic for the pronoun production task. Participants were given the grammatical gender twice in the prompt. Therefore, the correct gender-marked pronoun is the one that agrees with the grammatical gender given in the prompt. If a participant responded with *hi* ‘he’ after hearing a masculine definite article, this was considered an accurate gender-marked pronoun response. If a participant responded with *hat* ‘she’ after hearing a common definite article, this was considered an accurate response. All form errors were removed from the analysis. As the definite article elicitation task discussed, determining the correct grammatical gender can be difficult given that it is produced inconsistently by adult speakers. In this analysis the correct grammatical gender for each item was determined by the online survey described in Chapter 2.

It is important to keep in mind that the accuracy results reflect only correct form responses. For example, as seen in Table 4-14, adults responded with a gender-marked pronoun on 35% of the items, and the results in Table 4-16 reflect accuracy on only these responses. Elementary school participants of all home language groups performed around
chance level (50%)\(^9\). These results must be interpreted cautiously, however, given the limited amount of data. For example, for high consensus common items, adult speakers produced only 28% of the items with a gender-marked pronoun leaving a total of 23 items. The low production of gender-marked pronouns makes it difficult to determine gender accuracy patterns. Because there are few data remaining after the form errors were removed, the results will only be presented descriptively.

No clear patterns emerge among the elementary school students with regards to the category of items, although there is a tendency for common gender items to be produced more accurately. High school and adult speakers show a similar pattern, with higher accuracy on common gender items than masculine gender items.

Table 4-15. Pronoun production accuracy by Home Language and Category

<table>
<thead>
<tr>
<th>Home Language</th>
<th>All items</th>
<th>High Consensus</th>
<th>Low Consensus</th>
<th>High Consensus</th>
<th>Low Consensus</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Masculine</td>
<td>$M(SD)$</td>
<td>$M(SD)$</td>
<td>$M(SD)$</td>
<td>$M(SD)$</td>
</tr>
<tr>
<td>High-input</td>
<td>0.43 (0.37)</td>
<td>0.67 (0.24)</td>
<td>0.75 (0.12)</td>
<td>0.67 (0.24)</td>
<td>0.47 (0.19)</td>
</tr>
<tr>
<td>Middle-input</td>
<td>0.67 (0.22)</td>
<td>0.33 (0.58)</td>
<td>0.33 (0.58)</td>
<td>0.79 (0.34)</td>
<td>0.67 (0.58)</td>
</tr>
<tr>
<td>Minimal-input</td>
<td>0.50 (0.50)</td>
<td>0.50 (0.71)</td>
<td>0.33 (0.58)</td>
<td>1.00 (0.00)</td>
<td>1.00 (0.00)</td>
</tr>
</tbody>
</table>

\(^9\) 50% is considered chance level because only gender-marked pronouns were included, resulting in two options- a masculine gender pronoun and a common gender pronoun.
Table 4-16. Pronoun production accuracy by Age Group and Category

<table>
<thead>
<tr>
<th>Age Group</th>
<th>All items</th>
<th>High Consensus M(SD)</th>
<th>Low Consensus M(SD)</th>
<th>High Consensus M(SD)</th>
<th>Low Consensus M(SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-input</td>
<td></td>
<td>0.43 (0.37)</td>
<td>0.67 (0.24)</td>
<td>0.75 (0.12)</td>
<td>0.67 (0.24)</td>
</tr>
<tr>
<td>Elementary N = 13</td>
<td></td>
<td>0.76 (0.05)</td>
<td>0.83 (0.02)</td>
<td>0.52 (0.17)</td>
<td>0.93 (0.10)</td>
</tr>
<tr>
<td>High School N = 9</td>
<td></td>
<td>0.76 (0.23)</td>
<td>0.88 (0.35)</td>
<td>0.69 (0.40)</td>
<td>0.87 (0.22)</td>
</tr>
<tr>
<td>Adult N = 13</td>
<td></td>
<td>0.76 (0.23)</td>
<td>0.88 (0.35)</td>
<td>0.69 (0.40)</td>
<td>0.87 (0.22)</td>
</tr>
</tbody>
</table>

**Frequency of responses**

An analysis of the frequency of all responses, including unexpected forms, was conducted, as seen in Table 4-17 by Home Language and Table 4-18 by Age Group. Responses were coded into the following categories: masculine gender-marked pronouns, common gender-marked pronouns, masculine Frisian definite articles, common Frisian definite articles, full noun phrases (NP), and other. For all participants, 9% of items fell into the “other responses” category, including German responses and color only responses.

All groups show similar response patterns. For all groups the most frequent response type was the form *det*, which is both the common definite article and the non-gender-marked demonstrative pronoun in Fering. For high-input elementary school students, the second most common type of response was a full NP, whereas for middle-
input and minimal-input elementary students “other responses” were the second most common response. Together these two response types account for about just over half of the elementary school responses. High school and adult participants, however, almost never responded with a full NP, and favored the gender-marked pronouns as the second most common response, with each gender roughly equally frequent.

Table 4-17. Pronoun production frequency of responses by Home Language

<table>
<thead>
<tr>
<th>Home Language</th>
<th>Masculine Pronoun</th>
<th>Common Pronoun</th>
<th>Masculine Article</th>
<th>Common Article</th>
<th>Noun Phrase</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-input</td>
<td>hi M (SD) 0.13 0.10 0.11</td>
<td>hat M (SD) 0.10 0.10</td>
<td>de/a M (SD) 0.11</td>
<td>det/at M (SD) 0.11</td>
<td>det wel M (SD) 0.32</td>
<td>blä ‘blue’ M (SD) 0.22</td>
</tr>
<tr>
<td>N = 13</td>
<td>(0.24)</td>
<td>(0.20)</td>
<td>(0.19)</td>
<td>(0.39)</td>
<td>(0.38)</td>
<td>(0.28)</td>
</tr>
<tr>
<td>Middle-input</td>
<td>0.08</td>
<td>0.11</td>
<td>0.10</td>
<td>0.42</td>
<td>0.13</td>
<td>0.16</td>
</tr>
<tr>
<td>N = 11</td>
<td>(0.26)</td>
<td>(0.29)</td>
<td>(0.18)</td>
<td>(0.38)</td>
<td>(0.26)</td>
<td>(0.28)</td>
</tr>
<tr>
<td>Minimal-input</td>
<td>0.01</td>
<td>0.16</td>
<td>0.07</td>
<td>0.35</td>
<td>0.22</td>
<td>0.19</td>
</tr>
<tr>
<td>N = 6</td>
<td>(0.03)</td>
<td>(0.37)</td>
<td>(0.08)</td>
<td>(0.39)</td>
<td>(0.37)</td>
<td>(0.40)</td>
</tr>
</tbody>
</table>
Table 4-18. Pronoun production frequency of responses by Age Group

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Masculine Pronoun M (SD)</th>
<th>Common Pronoun M (SD)</th>
<th>Masculine Article M (SD)</th>
<th>Common Article M (SD)</th>
<th>Noun Phrase M (SD)</th>
<th>Other M (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elementary</td>
<td>0.13 (0.24)</td>
<td>0.10 (0.20)</td>
<td>0.11 (0.19)</td>
<td>0.32 (0.39)</td>
<td>0.22 (0.38)</td>
<td>0.12 (0.28)</td>
</tr>
<tr>
<td>High School</td>
<td>0.18 (0.22)</td>
<td>0.18 (0.27)</td>
<td>0.11 (0.18)</td>
<td>0.44 (0.42)</td>
<td>0.00 (0.01)</td>
<td>0.08 (0.15)</td>
</tr>
<tr>
<td>N = 13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adult</td>
<td>0.18 (0.22)</td>
<td>0.17 (0.30)</td>
<td>0.20 (0.21)</td>
<td>0.36 (0.29)</td>
<td>0.07 (0.11)</td>
<td>0.04 (0.06)</td>
</tr>
<tr>
<td>N = 13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4.3.4 Discussion and Summary

The most striking result of the pronoun production task is the low frequency of gender-marked pronoun responses. This is a common trend across all participant groups, although the adults and the high school students tended to use the gender-marked pronouns more often than the elementary school students. The results of the accuracy results reveal a trend toward higher accuracy for common gender items than masculine gender items. However, the low response rate makes it difficult to interpret these results with confidence. The frequency analysis showed that all groups, regardless of age or
home language, favored the *det* response. While this is the common gender definite article, it is more likely that participants intend this as the gender neutral demonstrative pronoun. Instead, participants appear to be using a semantic system based on a natural gender and animacy distinction, similar to English, where masculine pronouns occur with human men, common gender pronouns occur with human women, and inanimate items are marked with a gender neutral demonstrative pronoun.

One limitation of this task is that it is difficult to determine whether participants disfavored using the gender-marked pronoun in general, as the demonstrative pronoun was also felicitous in this context. However, during the examples, participants were given corrective feedback and examples indicating that the task required the gender-marked pronouns. If a participant responded to a practice item with *det as blä* ‘that/the is blue’, she was given explicit feedback, and told *könst dü sai – hat as blä* ‘can you say- She is blue’. Given that participants were given examples and feedback demonstrating that the task required a gender-marked pronoun, it is more likely that the overuse of the demonstrative reflects participants’ reluctance to use gender-marked pronouns generally, rather than not understanding the task requirements. Further, participants reported that this is an unnatural use of pronouns, and that although they use gender-marked pronouns for animate items, inanimate items are not marked for gender in Fering today. Many adult participants remarked after the task that it is unusual to ask what the gender pronoun they use for inanimate objects. The definite article production task shows that participants use grammatical gender with inanimate nouns, but pronoun agreement strikes native speakers as ungrammatical. One elementary school student reported on a radio program that the task was very difficult, which led to community member responses reporting that the task
was difficult because it requires participants to use grammatical gender agreement which is not a feature of contemporary Fering.

The vast majority, 85%, of participants either used gender-marked pronouns on all or almost all items, or never used a gender-marked pronoun. There are two accounts for this bifurcated pattern of results. It is possible that the participants who consistently marked gender on pronouns learned the task demands during the practice items, and therefore successfully completed an unnatural task. Alternatively, it is possible that there are two parallel systems within the community, those who use gender agreement in inanimate pronouns and those who do not use gender agreement in pronouns. Home language and age did not appear to predict whether a speaker marks grammatical gender on pronouns or not. Future research using naturalistic data, which did not have the same artificial task demands, could determine which of these two accounts best reflects the current use of gender-marked pronouns on Föhr. If participants were producing gender-marked pronouns to meet the demands of an artificial task, I predict that they would not do so during a naturalistic task. However, if there are two parallel systems in the community, I expect to see the same pattern of results in a naturalistic task as in this experimental task.

4.4 Pronoun Comprehension Task

4.4.1 Background

The pronoun comprehension task addresses the first broad research question: How does the quantity of Frisian language input influence participants’ acquisition of grammatical gender? This task tested whether participants make use of grammatical
gender information in the comprehension of gender-marked pronouns. The results of the elicitation task (Section 4.2) showed that some elementary school students do not produce grammatical gender with adult-like accuracy. However, this may be due to the fact that they do not know the correct gender to associate with a given noun, or it may mean that they do not use grammatical gender information. In this task, participants were given grammatical gender information on definite determiners and then asked to select a picture using a gender-marked pronoun. In this way, it is possible to determine whether participants use grammatical gender information in pronoun agreement during comprehension.

The pronoun comprehension task answers one specific research question: Does the quantity of Frisian input participants receive influence how accuracy in gender-marked pronoun comprehension? Following the general prediction that participants with more Frisian input will be more accurate in grammatical gender, I predict that high-input elementary school students will be more accurate in the comprehension of gender-marked pronouns than middle-input and minimal-input students. Moreover, examining the effects of cumulative input, I predict that high school students and adults will be more accurate on the gender-marked pronoun comprehension task than high-input elementary school students.

4.4.2 Methods and design

This task tested the comprehension of pronoun grammatical gender agreement, using a task similar to Gathercole and Thomas (2009). The 24 experimental items were distributed into 12 object pairs. The participants saw a magician and were told he can
make objects disappear. All target item stimuli followed the same pattern. The research assistant first introduced the objects, and provided grammatical gender information on the definite articles, as in 3a. On the following slide, the participant participants heard a pronoun with gender marking twice, as in 3b. On the final slide the participants saw two boxes with question marks on them and were asked to point where the item that disappeared was, as in 3c. A common gender item was always paired with a masculine gender item. High consensus items were always paired with lower consensus items to ensure that the difference in the level of consensus between target items was equal across items. Each item pair was shown twice, once using a masculine pronoun in the prompt and once with a feminine pronoun in the prompt. To limit priming effects across the task, half of the items were clipart images and the other half were photographs. Four lists were created to counter balance whether items appeared first as pictures or as clipart images and the whether the target item appeared on the left or right side of the screen, as seen in Figure 4-9. The order of items in all four lists was pseudo-randomized to ensure that no more than two items of any category, gender, or target location appeared in a row.

3a.  *Lukens heer! Heer san en kop uun en brek. Lachs in bilj as det kop, rochs in bilj as de brek.*

   ‘Look here! Here is a cup and pants. Left in the picture is the.COM cup and right in the picture is the.MASC pants.’

b.  *Zauber Zauber! Hi.MASC as ferswünjen! Hi.MASC as wech!*

   Hocus Pocus! It.MASC is gone! It.MASC disappeared!

c.  *Wise mi wat as ferswünjen.*

   Show me what disappeared.
All stimuli were recorded as a set, where the research assistant first introduced the items and followed by the location information. Each item was recorded twice and the best recording was selected. Using Audacity 2.01, the location information was spliced together with the introduction information. A 0.5 second pause was inserted between the introduction of the items and the location information. The left item was always introduced first, regardless of whether the target item appeared on the left or right side of the screen. The recordings were made with natural intonation, resulting in a continuation rise in pitch on the left location item, and falling intonation on the right item. The gender pronoun information and the final prompt, as in 3b,c, was identical regardless of which item was correct to ensure that there were no intonational cues which could have indicated the correct answer.

Pilot research with adults indicated that participants required multiple examples to understand the task procedure. All participants completed eight practice items prior to beginning the target items. The first item was completed by the researcher as an example.

Figure 4-8. Stimulus for Gender Pronoun Comprehension Task
The first two items included obvious size and color contrasts, for example, the blue book versus the yellow book and the big box versus the small box. In these examples the disambiguating information, similar to 3c, included either the color or size of the object. Participants were then told that the next trials were more difficult, and that they should listen carefully. The subsequent two examples included natural gender oppositions with pairs of people, for example, a boy versus a girl, or a man “Jan” and a woman “Julia”. In these examples, the disambiguating information was a gender-marked pronoun, meaning that it was identical to the experimental items. The final set of practice items included grammatical gender oppositions similar to the experimental task, but used animals instead of inanimate objects. As with the previous examples, a gender-marked pronoun was given as disambiguating information. Finally participants were given a short break prior to starting the task with the target items, during which time they could ask questions.

As the second task of the second session, the same participants completed this task as the gender-marked pronoun production task–13 high-input, 11 middle-input, and 6 minimal-input students completed this task. The same nine high school students and 13 adults as in the gender production task and the gender-marked pronoun production task also completed this pronoun comprehension task. All of the participants will be included in all of the comparisons for this task, because the difference in the number of participants was much smaller on this task than the gender determiner production task.

4.4.3 Results

In order to investigate for the effects of the amount of Frisian input, two sets of comparisons were conducted: a Home Language comparison and an Age Group
comparison. First the high-input, middle-input, and minimal-input students were compared on accuracy, then the results of the high-input bilinguals from different age groups was compared to test for the effect of cumulative language exposure. As in the gender-marked pronoun production task, grammatical gender accuracy refers to selecting the gender-marked pronoun that agrees with the grammatical gender information given in the prompt. Following the planned comparisons of accuracy, two post-hoc analyses were conducted to further explore the effect of quantity of language input.

**Gender-Marked Pronoun Comprehension Accuracy**

First, looking at accuracy for the elementary school students split by home language group, we see that overall the all groups appear to be at chance level (50%)\(^{10}\). Further exploring the results split by word type (Level of Consensus and Grammatical Gender) does not reveal large differences, as seen in Table 4-15. That is, the elementary school students performed similarly across items regardless of the level of consensus or grammatical gender. A comparison of high-input participants at different ages showed that high school and adult participants were more accurate than elementary school participants, but all groups perform similarly across categories, as seen in Table 4-20. Overall, it appears that the participants struggled with the pronoun comprehension task across all Home Language and Age Groups, and did not show differences across categories.

---

\(^{10}\) Chance level is 50% because there are two pictures to choose from.
Table 4-19. Pronoun comprehension accuracy by Category and Home Language

<table>
<thead>
<tr>
<th>Home Language</th>
<th>All Items</th>
<th>High Consensus</th>
<th>Low Consensus</th>
<th>High Consensus</th>
<th>Low Consensus</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Masculine</td>
<td>M(SD)</td>
<td>M(SD)</td>
<td>M(SD)</td>
<td>M(SD)</td>
</tr>
<tr>
<td>High-input</td>
<td>0.57 (0.16)</td>
<td>0.60 (0.30)</td>
<td>0.50 (0.27)</td>
<td>0.55 (0.14)</td>
<td>0.62 (0.24)</td>
</tr>
<tr>
<td>N = 13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle-input</td>
<td>0.53 (0.07)</td>
<td>0.64 (0.22)</td>
<td>0.53 (0.16)</td>
<td>0.44 (0.15)</td>
<td>0.50 (0.25)</td>
</tr>
<tr>
<td>N = 11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minimal-input</td>
<td>0.49 (0.14)</td>
<td>0.50 (0.32)</td>
<td>0.55 (0.13)</td>
<td>0.36 (0.22)</td>
<td>0.55 (0.21)</td>
</tr>
<tr>
<td>N = 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4-20. Pronoun comprehension accuracy by Category and Age Group

<table>
<thead>
<tr>
<th>Age Group</th>
<th>All Items</th>
<th>High Consensus</th>
<th>Low Consensus</th>
<th>High Consensus</th>
<th>Low Consensus</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Masculine</td>
<td>M(SD)</td>
<td>M(SD)</td>
<td>M(SD)</td>
<td>M(SD)</td>
</tr>
<tr>
<td>High-input</td>
<td>0.57 (0.16)</td>
<td>0.60 (0.30)</td>
<td>0.50 (0.27)</td>
<td>0.55 (0.14)</td>
<td>0.62 (0.24)</td>
</tr>
<tr>
<td>Elementary</td>
<td>N = 13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High School</td>
<td>0.80 (0.12)</td>
<td>0.80 (0.20)</td>
<td>0.80 (0.11)</td>
<td>0.76 (0.12)</td>
<td>0.83 (0.17)</td>
</tr>
<tr>
<td>Adult</td>
<td>N = 13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.74 (0.17)</td>
<td>0.72 (0.27)</td>
<td>0.71 (0.22)</td>
<td>0.72 (0.18)</td>
<td>0.82 (0.20)</td>
</tr>
</tbody>
</table>
Post-hoc Analysis of Order Effects

While the participants were not given any feedback that revealed the gender-marked pronoun to definite article agreement patterns, participants could have developed a strategy for selecting answers. That is, they could have assumed that *hi* ‘he’ was used with masculine inanimate nouns, and *hat* ‘she’ was used with common gender inanimate nouns, and applied that information to successfully complete the task, despite the fact that they would not use gender-marked pronouns for inanimate nouns in their own speech. To investigate this possibility, the results were reanalyzed comparing the first half of items to the second half of items. Participants were randomly assigned to one of four lists in which items were presented in four pseudorandom orders. For each list, categories of items were evenly distributed between the first half and second half. The results, presented in Table 4-21 and Table 4-22, show that participants do not improve over the course of the task.

Table 4-21. Pronoun comprehension improvement by Home Language

<table>
<thead>
<tr>
<th>Home Language</th>
<th>Improvement (Second Half-First Half)</th>
<th>First Half Accuracy</th>
<th>Second Half Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>M (SD)</strong></td>
<td><strong>M (SD)</strong></td>
<td></td>
</tr>
<tr>
<td>High-input N = 13</td>
<td>0.04 (0.16)</td>
<td>0.54 (0.18)</td>
<td>0.59 (0.18)</td>
</tr>
<tr>
<td>Middle-input N = 12</td>
<td>0.04 (0.24)</td>
<td>0.51 (0.12)</td>
<td>0.55 (0.15)</td>
</tr>
<tr>
<td>Minimal-input N = 7</td>
<td>0.00 (0.33)</td>
<td>0.49 (0.11)</td>
<td>0.49 (0.28)</td>
</tr>
</tbody>
</table>
Table 4-22. Pronoun comprehension improvement by Age Group

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Improvement (Second Half-First Half)</th>
<th>First Half Accuracy M (SD)</th>
<th>Second Half Accuracy M (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-input</td>
<td>0.04 (0.16)</td>
<td>0.54 (0.18)</td>
<td>0.59 (0.18)</td>
</tr>
<tr>
<td>Elementary N = 13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High School N = 9</td>
<td>0.06 (0.13)</td>
<td>0.77 (0.14)</td>
<td>0.82 (0.12)</td>
</tr>
<tr>
<td>Adult N = 13</td>
<td>0.08 (0.19)</td>
<td>0.70 (0.19)</td>
<td>0.78 (0.20)</td>
</tr>
</tbody>
</table>

Post-hoc analysis of practice items

To test whether participants understood the task procedure, a post-hoc analysis was conducted for the practice items. There were three categories of practice items. The first two items had either a color or size distinction, meaning that no knowledge of gender was necessary to complete these two items. The next two items contained natural gender information with human items. Finally, there were two animal items with grammatical gender information. All participants completed the practice items in this same order. Data were available for only 93% of items for the practice items, as this was not a planned analysis. Data were missing from 7% of items due to experimenter errors including not starting the video recording until the experimental items had begun, or not recording the examples by hand. The results revealed that participants performed at ceiling for size and color distinctions, as well as natural gender items with people. Adult participants and
high school students also performed at ceiling for grammatical gender items with animals. Table 4-23 presents the results for practice items by Home Language and Table 4-24 presents the results by Age Group and practice category. Within the elementary school students, all three language input groups were at least 90% accurate for color/size items and natural gender items, but differed on animal items, in which the minimal-input group fell to chance level.

Table 4-23. Gender comprehension practice items by Home Language

<table>
<thead>
<tr>
<th>Home Language</th>
<th>Color/Size</th>
<th>Human</th>
<th>Animals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N = 2$</td>
<td>$N = 2$</td>
<td>$N = 2$</td>
</tr>
<tr>
<td></td>
<td>$M (SD)$</td>
<td>$M (SD)$</td>
<td>$M (SD)$</td>
</tr>
<tr>
<td>High-input</td>
<td>0.96 (0.14)</td>
<td>1.00 (0.00)</td>
<td>0.88 (0.22)</td>
</tr>
<tr>
<td>Middle-input</td>
<td>1.00 (0.00)</td>
<td>0.95 (0.15)</td>
<td>0.86 (0.23)</td>
</tr>
<tr>
<td>Minimal-input</td>
<td>1.00 (0.00)</td>
<td>0.93 (0.19)</td>
<td>0.43 (0.45)</td>
</tr>
</tbody>
</table>

Table 4-24. Gender comprehension practice items by Age Group

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Color/Size</th>
<th>Human</th>
<th>Animals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N = 2$</td>
<td>$N = 2$</td>
<td>$N = 2$</td>
</tr>
<tr>
<td></td>
<td>$M (SD)$</td>
<td>$M (SD)$</td>
<td>$M (SD)$</td>
</tr>
<tr>
<td>High-input Elementary</td>
<td>0.96 (0.14)</td>
<td>1.00 (0.00)</td>
<td>0.88 (0.22)</td>
</tr>
<tr>
<td>High School</td>
<td>0.96 (0.14)</td>
<td>0.93 (0.19)</td>
<td>0.96 (0.14)</td>
</tr>
<tr>
<td>Adult</td>
<td>1.00 (.00)</td>
<td>1.00 (0.00)</td>
<td>1.00 (0.00)</td>
</tr>
</tbody>
</table>
4.4.4 Discussion

Similar to the pronoun elicitation task, participants appear to struggle with the gender-marked pronoun comprehension task. This task required participants to connect grammatical gender information with gender-marked pronouns, an opposition which may not be robust for inanimate nouns in contemporary Fering. A comparison of the first half and second half of items does not show improvement over the course of the task, suggesting that participants did not develop a strategy for responses.

The results of the post-hoc analysis rule out the possibility that participants’ low performance stems from a misunderstanding of the task method. All participant groups were able to successfully complete the task when the items were distinguished by size, color, or natural gender. Despite the fact that the animal items showed some more variability, the general pattern indicates that participants understood the task methodology. In contrast, participants struggled with the inanimate items in the experimental portion of the task.

Importantly, if participants followed the grammatical gender information provided twice, they would have been able to respond accurately to all items, even without any specific knowledge of the grammatical gender of the items. However, in order to successfully complete the task, the participant must know the agreement pattern between gender-marked pronouns and definite articles and attend to the grammatical gender information in the prompt. The low accuracy suggests that either participants do not know the agreement patterns or they do not attend to the grammatical gender
information given on the definite articles in the prompts, or do not use this information in agreement for pronouns for inanimate objects.

4.4.5 Summary

While all groups appeared to understand the procedure of the gender-marked pronoun comprehension task, as demonstrated by their accuracy on the practice items, they did not attend to grammatical gender information on definite articles in the prompt for the comprehension of gender-marked pronouns. Moreover, participants did not improve over the course of the task, suggesting that the participants did not develop an effective strategy over the course of the task. The results of the pronoun comprehension task, taken together with the results of the pronoun production task and participants’ reports that pronouns are not marked for gender agreement, strengthens the possibility that gender agreement on pronouns with inanimate antecedents is not robust in contemporary Fering. Instead, Fering appears to have shifted toward an English-like gender system in which pronouns follow a semantic system based on animacy and sentience. It appears that pronouns with a masculine animate antecedent are marked with hi ‘he’, pronouns with a feminine animate antecedent are marked with hat ‘she’, and pronouns with inanimate antecedents are not marked for grammatical gender.

4.5 Summary and Discussion

This chapter investigated how the amount of input influences the acquisition of inconsistent morphology. As adults produce grammatical gender in Fering inconsistently, the acquisition of Frisian grammatical gender was used as an example of the acquisition of inconsistent morphology generally. I presented three tests of grammatical gender,
including a definite article elicitation task, a gender-marked pronoun elicitation task, and a gender-marked pronoun comprehension task. Throughout the chapter, two comparisons were used to test the role of the amount of input. The Home Language comparison contrasts the results of three groups of elementary school students with different amounts of Frisian input. The Age Group comparison examined the role of cumulative language input by contrasting high-input bilinguals from three age groups.

Across both comparisons, the gender production task showed that participants with more Frisian input make fewer word errors and produce patterns of grammatical gender which are more adult-like. Within elementary school students, the high-input group produced the fewest word errors and was the only elementary school group to show a significant correlation with adult grammatical gender production for correct word responses. Similarly, the high school students, who have more cumulative input than high-input elementary school students, make fewer word errors and showed a strong positive correlation with the adult-like pattern of grammatical gender. Overall, the Home Language comparison and the Age Group comparison provide converging evidence that the amount of Frisian input significantly influences how speakers acquire Fering grammatical gender marking.

With regards to regularization, the definite article elicitation results support previous findings that some elementary school students regularize inconsistent input (Hudson Kam & Newport, 2005; 2009). Overall, the regularization analysis showed that almost half ($N = 16$ of $N = 36$) elementary school students used the same grammatical gender marking for all or all but one item. However, the amount of Frisian input influences how often elementary school students regularize inconsistent input. While
both the high-input group and the minimal-input group show less regularization than the middle-input students, the high-input group produces adult-like inconsistency, while the minimal-input students are randomly inconsistent. While these results support Hudson Kam and Newport’s (2005, 2009) findings, they also highlight the drawbacks of a study with limited exposure time. With increased exposure, children may overcome the initial stage of regularization and faithfully acquire even a complex and inconsistent feature, such as grammatical gender.

Both the pronoun production task and the pronoun comprehension task suggest that marking grammatical gender agreement on pronouns with inanimate antecedents may not be robust in contemporary Fering. In terms of production, participants appeared to prefer the gender neutral demonstrative pronoun, det, over gender-marked pronouns. The pronoun comprehension task supports this hypothesis, in that, if children were attending to the grammatical gender information in the task, they would have been able to respond to all items successfully. However, across all groups and categories, elementary school participants perform at chance. The practice items show that participants understood the task procedure, but struggled with using grammatical gender information in pronoun agreement.

To summarize: This chapter on Fering grammatical gender shows that while some children regularize inconsistent input, the acquisition of inconsistent input is influenced by the quantity of input. While acquiring an inconsistent system requires extensive input, given sufficient input, it is possible to faithfully learn patterns of inconsistency. From these data it is not possible to determine whether the results stem from overall reduced language skills among lower input students, or whether they represent something
particular about the acquisition of inconsistent morphology. Chapter 5 presents the results of two plural marking experiments which can show the role of quantity of input independent from inconsistency.
Chapter 5. Results: Plural Marking

5.1 Introduction

Similar to grammatical gender, plural marking acquisition is influenced by the quantity of input children receive. For monolingual acquisition, children acquire frequent forms more quickly and with less input less frequent forms (Bittner & Köpcke, 2001; Mills, 1986; Szagun, 2001). While grammatical gender and plural marking are similar in this regard, in Fering, they differ in how consistently they are produced by adult speakers. While adults produce grammatical gender inconsistently in Fering (See Chapter 4), pilot studies suggest that plural marking is produced consistently. Adults often disagree on the grammatical gender of a noun, but they rarely disagree about which plural marking is used for a given noun. This means that investigating children’s acquisition of plural markings can test the role of the quantity of input independently from the input type.

Before specifically discussing plural marking acquisition in Fering, it is important to understand the child’s task in the acquisition of the plural marking. In Fering, the acquisition of plural marking requires the child to determine that her language marks plural on nouns, and determine from the input which plural marking is associated with each noun. This chapter focuses on the acquisition of these noun-to-plural marking correspondences, rather than the acquisition of the category of plural. Consider a child learning Frisian plural markings: every time the child hears wel–en ‘bikes’ it contains the –en plural marking. She can quickly determine from the input that the –en plural marking occurs with this noun. In comparison, when learning grammatical gender, she will hear wel ‘bike’ with both masculine and common genders interchangeably. Therefore, the
input for plural marking is more consistent than for grammatical gender. This dissertation investigates how the quantity and type of input influence language acquisition, and in this chapter I address two broader research questions:

1. How does the amount of Frisian language input influence participants’ acquisition of plural markings?

2. Do children regularize plural markings in Fering, as they do grammatical gender?

With regard to the first question, I predict participants with more Frisian input will be more accurate in marking plural. In terms of regularization, if there are differences between acquiring inconsistent input and consistent input, then this predicts that while children regularize grammatical gender in Fering, they will regularize plural markings less often. If, however, input consistency does not influence acquisition, then children will regularize plural markings to the same extent that they do with grammatical gender.

This chapter presents the results of two experiments— one plural marking elicitation task and one acceptability judgment task. The chapter is organized as follows: Section 5.1 begins with an overview of the plural marking system in Fering, and then presents information about the participants, materials, and procedures for the production and comprehension task. Section 5.2 presents the results of the plural marking task. Section 5.3 outlines the results of the acceptability task. Section 5.4 discusses and summarizes the results of this chapter.
5.1.1 Plural Marking in Fering

As seen in Table 5-1, plural is marked in Fering on regular nouns by adding a suffix on to the singular form of the noun. Chapter 2 provides a more detailed description of the Fering plural system.

Table 5-1. Regular Plural Marking in Fering (Wilts, 2011)

<table>
<thead>
<tr>
<th>Plural Marking</th>
<th>Fering Singular</th>
<th>Fering Plural</th>
<th>English Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>–er</td>
<td>a hünj</td>
<td>a hünjer</td>
<td>dogs</td>
</tr>
<tr>
<td></td>
<td>at wöning</td>
<td>a wönger</td>
<td>windows</td>
</tr>
<tr>
<td>–en</td>
<td>at buk</td>
<td>a buken</td>
<td>books</td>
</tr>
<tr>
<td></td>
<td>at foomen</td>
<td>a foomnen</td>
<td>girls</td>
</tr>
<tr>
<td>-in</td>
<td>at baantje</td>
<td>a baantjin</td>
<td>offices</td>
</tr>
<tr>
<td>-n</td>
<td>a mooler</td>
<td>a moolern</td>
<td>painters</td>
</tr>
<tr>
<td>-s</td>
<td>a koptein</td>
<td>a kopteins</td>
<td>captains</td>
</tr>
</tbody>
</table>

Wilts (2011) categorizes plural markings into five regular suffixes, and four irregular plural markings. A pilot survey of basic Frisian vocabulary showed that contemporary Fering uses a more simplified system, where two plural markings, –er and –en, dominate the Frisian plural marking system. In a sample of about 500 words, close to one third of nouns (32%) were marked with –er suffix, and about half (53%) were
marked with the –en suffix. Irregular plurals accounted for 7.5% of the items and the remaining nouns were scattered between the null plural, –in, and –s suffixes, and, of these, no one plural marking accounted for more than 3% of items. The distributional pattern results in a system where, from the learner’s perspective, there are basically two options for plural markings, –en or –er.

Unlike grammatical gender, the results of the adult survey study (See Chapter 3) show that adult native speakers produce plural markings consistently in Fering. That is, adults largely agree on which plural inflection they use. The plural elicitation task in Section 5.2 supports this previous result.

5.1.2 General Participants

Three age groups of participants completed the two plural tasks— elementary school students, high school students, and adults (parents and adult community members). These participants also completed the grammatical gender tasks presented in Chapter 4, and the general fluency task presented in Chapter 6.

Level of Input

As in Chapter 4, Participants were separated into three groups based on the quantity of Frisian input spoken at home. A sub-set of questions from the Utrecht Bilingualism Exposure Calculator (UBiLEC) were used to determine the participants’ home language (Unsworth, 2011). Following procedures used in Unsworth (2011), for elementary school students, the primary investigator interviewed the student’s parents or the student herself when the parent was not available (Unsworth, 2011). High school students were interviewed by the investigator. The interview asked about the child’s
speech partners, which languages the parents use with the child, and how often they speak Frisian with the child. Based on these results, participants were split into three groups: high-input homes where Frisian is spoken between 75-100% of the time at home; middle-input homes where Frisian is spoken at home about half of the time (25-75%), and minimal-input homes where Frisian is rarely spoken (0-25%).

This dissertation focuses child language acquisition, and therefore the results of the elementary school students are of primary interest. High-input high school students and adult native speakers serve as comparison groups. The results of the middle-input and low-input high school students and adult second language learner speakers will not be presented here. Throughout this chapter I will compare elementary school students from different home language groups—high-input, middle-input, and minimal-input—and high-input bilinguals from three age groups—elementary school students, high school students, and adults—to investigate the role of the quantity of input plays in the acquisition of plural markings. For the age group comparison, since all of these speakers are from high-input homes, the assumption is that older participants have cumulatively more input than the younger participants. Therefore, both the Home Language comparisons and the Age Group comparison provide information about the role of quantity of input on the acquisition of plural marking in Fering.

5.1.3 General Materials

Twenty-eight items were separated into four categories. Previous research investigating grammatical gender processing among bilinguals has shown that bilingual speakers are sensitive to the similarities between their two languages and respond
differently when words share inflections in both languages (e.g., Bordag & Pechmann, 2007; Sabourin & Stowe, 2008; Sabourin, Stowe & de Hann, 2006). In other words, if a suffix has the similar forms in both languages as, for example, the German plural marking, \(-n\), and the Fering plural marking \(-en\), it is predicted to be easier to acquire the plural marking in Fering. As all participants are bilingual in both German and Fering, the similarities between plural markings were taken into consideration when designing the categories of items. As seen in 1-4, items were classified into four categories.

1. Match: the same plural marking in German and Fering \((N = 9)\)
   
e.g., Fering: brag–en ‘bridges’ c.f. German: Brücke-n ‘bridges’

2. Mismatch-\(en\): different plural marking in German and Fering and \(-en\) marking in Fering \((N = 6)\)
   
e.g., Fering: wel–en ‘bikes’ c.f. German: Fahrräd–er ‘bikes’

3. Mismatch-\(er\): different plural marking in German and Fering and \(-er\) marking in Fering \((N = 9)\)
   
e.g., Fering: fask–er ‘fish’ c.f. German: Fisch-e ‘fish’

4. Irregular plurals \((N = 4)\)
   
e.g., Fering: tes ‘teeth’ c.f. Zähne ‘teeth’

One match category was created, with words with both \(-en\) and \(-er\) marking in Fering, because matched items are relatively infrequent. Eight of the nine items in the matching category were marked with the \(-en\) suffix in Fering. Words with different plural markings on the Fering word and its German translation were split into two categories based on the plural marking in Fering. Similar to the overall distribution of plural markings from the
pilot corpus study, across all categories, 14 items (50%) were marked with –en suffix, 10 with –er (35%), and there were 4 irregular plurals (14%).

All words except the irregular items were tested in an adult pilot study (See Chapter 2). Because previous research has shown that bilinguals process cognates differently from non-cognates (e.g., Lemhöfer, Spalek, & Schriefers, 2008), all items were non-cognates with German, Low German, or Northern German Vernacular, as judged by a native speaker informant who was not associated with the project. Items were presented as clip-art images. All but six items\textsuperscript{11} were normed with ten adult native speaker community members to ensure that the images were recognizable. For the normed items, an average of 8.78 of the ten speakers agreed on the Frisian word (range 4-10).

5.1.4 General Procedure

All tasks were administered by a Frisian-German bilingual native speaker and the primary investigator. For the elementary school children, the plural tasks, the grammatical gender tasks (see Chapter 4) and the general fluency task (see Chapter 6) were split into two separate testing sessions, to ensure they remained attentive. High school students and adult participants completed the plural tasks in one hour-long session, along with the grammatical gender tasks (see Chapter 4) and the general fluency task (see Chapter 6). The order of the grammatical gender elicitation and plural elicitation task was counter-balanced across participants; the order of the other tasks was the same across all participants. The order of the tasks was as follows:

\textsuperscript{11} The four irregular plural items were added after norming was conducted. 36 items were normed, but 12 items were not able to be used because they are primarily used with a German borrowing or speakers disagree about the word used. A native speaker research assistant confirmed that the images used for the additional six items accurately matched the words, and that the words are currently used on Föhr.
Session 1:
Task 1: Grammatical gender elicitation OR plural elicitation *
Task 2: Plural elicitation OR grammatical gender elicitation *
Task 3: General fluency story-telling

Session 2:
Task 4: Pronoun production task
Task 5: Pronoun comprehension task
Task 6: Plural acceptability task
* order counter-balanced across participants

In this chapter I present the results of the plural elicitation task and the plural acceptability task, highlighted in bold. Participants completed three practice items before each of the tasks to ensure that they understood the task procedure. The items for each task, instructions, and example items are presented in Appendix C.

Students completed the tasks individually with the research assistant and primary investigator in a quiet room at the school. Elementary school students were tested during school hours, most often during Frisian class, and high school students were tested after school. Adults participated at a quiet location convenient for them, most often at their home. All stimuli were presented on a laptop computer using PowerPoint 2010, and all sessions were recorded using an external microphone and Audacity 2.01. The sessions were also video recorded if the participant or the participant’s parent consented. During the session, the primary investigator transcribed the responses, which were later confirmed through video or audio recordings. A native speaker research assistant transcribed the results of the plural elicitation task. For the elicitation task, child
participants were given a short break half-way through and allowed to select a sticker. Adult and high school participants were told when they were half-way through the task, but were not given a break. There was no break during the acceptability task.

5.2 Plural elicitation task

5.2.1 Background

The plural elicitation task addresses each of the two broad research questions. Participants were asked to name sets of pictures in order to elicit plural markings. This task tests which plural marking speakers associate with a given noun. To return to the earlier example, this task tests whether the child learning Fering has accurately learned from the input that wel ‘bike’ is marked with the –en suffix to indicate plural, rather than the –er suffix, and can accurately produce this in her own speech.

The broad research questions are addressed with three more specific research questions (labeled here to correspond with the broad research questions).

1a. Does the quantity of Frisian input participants receive influence how accurately participants use Frisian vocabulary?

1b. Does the quantity of Frisian input participants receive influence how accurately participants produce plural markings?

2. Do elementary school students regularize plural markings, similar to grammatical gender markings?

I predict that the amount of Frisian input will influence participants’ production of plural markings, with students with more Frisian input more accurate in marking plurals.
However, I do not predict that elementary school students will regularize plural markings, as adults produce plural markings consistently.

5.2.2 Participants

As a part of session one, 48 elementary school students completed the plural elicitation task. Of these, 21 were high-input students, 14 were middle-input students, and 13 were minimal-input students. Nine high-input high school students and 13 adult native speakers completed this task, and will be used as a comparison for the elementary school results.

5.2.3 Methods and design

This task uses the same design as the definite article elicitation task (Chapter 4 Section 4.2), based on Schuele and Tolbert (2001). In order to set up a felicitous context for eliciting nouns, participants were told that they were going to help the primary researcher learn Fering. The participant sat with a native speaker research assistant (Frisian-German bilingual) and viewed sets of pictures on a computer screen. The primary researcher sat facing the participants, so that she could not see the computer display. Participants saw two sets of three identical objects on the screen, and one of the sets of objects was circled (See Figure 5-1).
Figure 5-1. Sample elicited production item

The participant was instructed to first name both sets of objects on the screen and then to name the circled set of objects. This was done in order to parallel the definite article elicitation task, which limited the number of different types of tasks overall. All sets of objects appeared twice over the course of the task, once as a distractor item, and once as the target item. Four lists were created to counter-balance whether the object appeared first as a target or distractor item, and for the location of the target item (right side of the screen versus left side of the screen). Participants were randomly assigned to one of the four lists. There were three practice items before the task began. For the practice items, participants were asked to say how many objects appeared in the set so as to elicit a number in their response 1c. This would ensure that they understood that the response was plural. During the practice items participants were given corrective feedback about the correct format. The most common format error was omitting the number in the response, and participants were told by the research assistant to provide the number of
items in her responses. If participants provided the correct format—a number and a noun phrase—only positive feedback was given.

(1) a. Assistant:  *Wat schochst dü heer?*  
   ‘What do you see here?’

b. Child:  *trii stäären an trii dööbler*  
   ‘three stars and three dice’

c. Assistant:  *Sai Alison, wat heer umkreist as.*  
   ‘Tell Alison what is circled.’

d. Child:  *Alison, trii dööbler san umkreisert.*  
   ‘Alison, three dice are circled.’

**Coding**

A Fering native speaking research assistant used the audio recordings to transcribe the responses. The transcriptions were coded by the primary investigator as either a correct word response or a word error. Correct word responses were then subsequently coded for the type of plural marking (e.g., *–en, –er*, irregular, or other).

5.2.4 **Results**

The results are presented as follows: first the results of the adult speakers are presented in order to check that plural markings are produced consistently in Fering. Secondly, vocabulary accuracy—how often participants produce word errors—will be analyzed. Then, isolating plural acquisition from vocabulary acquisition, plural accuracy for correct word responses—how often participants produce the correct plural marking when they produce the correct word—is analyzed. Both vocabulary accuracy and plural
accuracy will be tested through a Home Language comparison and Age Group comparison. Finally, the regularization analysis investigates whether participants regularize plural markings in Fering. For all ANOVA statistics, post-hoc tests used the Tukey HSD test when available and the Bonferonni correction when Tukey HSD was not available. Tukey HSD was used because Bonferonni was considered too conservative for these comparisons following up on main effects. SPSS does not provide a Tukey HSD test for testing post-hoc comparisons for interaction effects, and in these cases a Bonferonni correction was used.

**Adult plural production**

Before turning to child language acquisition, I analyzed adult plural responses to establish a baseline for the child results. This dissertation tests whether participants acquire inconsistent input differently than consistent input through a comparison between inconsistently produced grammatical gender and consistently produced plural marking. The first analysis checks that adult speakers produce plural marking consistently.

The plural responses of the 13 adult native speakers were analyzed by item for the level of consensus of each item. The level of consensus is the percentage of speakers who marked a given noun with the same plural marking. For example, if a word has a 95% level of consensus, this means that 95% of speakers used the same plural marking with this word. Only the responses for the correct words were analyzed, and word errors, such as German borrowings, were removed (Refer to Table 5-5 for the percentage of word errors by category). This step was taken because some correctly marked German borrowings would have been erroneously considered incorrect plural markings in Fering.
For example, a German borrowing, *Tisch-e* ‘tables’, in place of the Frisian *boosl–er* ‘tables’, could have been considered incorrect, because it was not marked with the –er, even though –e is the correct plural marking for the word *Tisch* in German. The percentages presented in Table 5-2 reflect the number of responses with the same plural marking divided by the total of correct word responses for each item. The results are split by Category (Match vs. Mismatch-*en* vs. Mismatch-*er* vs. Irregular).

Table 5-2. Level of adult consensus on plural production

<table>
<thead>
<tr>
<th>Category</th>
<th>Average Level of Consensus</th>
<th>SD</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Match N = 9</td>
<td>0.97</td>
<td>0.08</td>
<td>0.75</td>
<td>1.00</td>
</tr>
<tr>
<td>Mismatch-<em>en</em> N = 6</td>
<td>0.95</td>
<td>0.08</td>
<td>0.80</td>
<td>1.00</td>
</tr>
<tr>
<td>Mismatch-<em>er</em> N = 9</td>
<td>0.97</td>
<td>0.06</td>
<td>0.85</td>
<td>1.00</td>
</tr>
<tr>
<td>Irregular N = 4</td>
<td>0.94</td>
<td>0.04</td>
<td>0.91</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Overall, adult native speakers largely agree with each other as to the plural marking associated with a given noun, and this holds true across categories. These results confirm that plural is marked consistently among adult speakers of Fering, and thus that the plural marking input available to children is consistent. As plural marking is produced consistently, I predict that the child participants will not regularize plural markings to the same extent they did grammatical gender markings. However, the amount of input children have will influence their acquisition of plural markings, with high-input children exhibiting more adult-like production of plural marking than middle- and low-input children.
**Vocabulary Accuracy**

Before turning to plural accuracy, I first analyzed the vocabulary accuracy, specifically how often participants produce the correct word rather than a word error. The goal of this analysis is to separate vocabulary accuracy from plural accuracy. Responses with the expected Frisian word were considered correct words, while all other responses were considered word errors. There were several types of word errors, including German translations (e.g., *Tasse* ‘cup’ for *kop* ‘cup’), phonologically adapted German borrowings (e.g., *Tass* ‘cup’ for *kop* ‘cup’), Frisian words for similar concepts (e.g., *skütel* ‘small bowl’ for *dask* ‘large bowl’), and other concepts (e.g., *Feuerzeug* ‘lighter’ for *swaavelstock* ‘match’).

The vocabulary accuracy analysis tests whether participants differ in how often they produce word errors without considering the plural marking, while the plural accuracy analysis investigates plural marking without considering word errors. First vocabulary accuracy in elementary school students with different amounts of Frisian input will be compared, and then high-input bilinguals from three age groups will be compared. Following the broad prediction that participants with more Frisian input will perform better on the plural marking task, the prediction is that high-input elementary school students will be more accurate on vocabulary than middle-input or minimal-input students, and that elementary school students will produce more word errors than high school students or adults.
Home Language

Vocabulary accuracy among elementary school students from three home language groups was compared to test the effect of quantity of input on vocabulary acquisition. Table 5-3 shows the vocabulary accuracy for each Home Language group and category (Match vs. Mismatch-\textit{en} vs. Mismatch-\textit{er} vs. Irregular). The most striking result is the large difference in vocabulary accuracy between groups with different amounts of Frisian input. As predicted, the high-input students produced the fewest word errors, followed by the middle-input students, and then the minimal-input students. Across categories of items, however, children produce roughly the same number of word errors.

Table 5-3. Vocabulary accuracy by Home Language and Category

<table>
<thead>
<tr>
<th>Home Language</th>
<th>All Items</th>
<th>Match</th>
<th>Mismatch-\textit{en}</th>
<th>Mismatch-\textit{er}</th>
<th>Irregular</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M(SD)$</td>
<td>$M(SD)$</td>
<td>$M(SD)$</td>
<td>$M(SD)$</td>
<td>$M(SD)$</td>
</tr>
<tr>
<td>High-input</td>
<td>0.55</td>
<td>0.57</td>
<td>0.46</td>
<td>0.51</td>
<td>0.77</td>
</tr>
<tr>
<td>$N = 21$</td>
<td>(0.10)</td>
<td>(0.17)</td>
<td>(0.14)</td>
<td>(0.12)</td>
<td>(0.16)</td>
</tr>
<tr>
<td>Middle-input</td>
<td>0.38</td>
<td>0.38</td>
<td>0.30</td>
<td>0.38</td>
<td>0.55</td>
</tr>
<tr>
<td>$N = 14$</td>
<td>(0.20)</td>
<td>(0.24)</td>
<td>(0.19)</td>
<td>(0.25)</td>
<td>(0.28)</td>
</tr>
<tr>
<td>Minimal-input</td>
<td>0.20</td>
<td>0.15</td>
<td>0.13</td>
<td>0.22</td>
<td>0.37</td>
</tr>
<tr>
<td>$N = 13$</td>
<td>(0.15)</td>
<td>(0.17)</td>
<td>(0.11)</td>
<td>(0.15)</td>
<td>(0.32)</td>
</tr>
</tbody>
</table>

A mixed 3 X 3 ANOVA with the percentage of vocabulary accuracy as the dependent variable was conducted. Home Language (High-input vs. Middle-input vs. minimal-input) and category (Match vs. Mismatch-\textit{en} vs. Mismatch-\textit{er} vs. Irregular) were significant factors.
Minimal-input) was entered as a between-subjects variable and Category (Match vs. Mismatch-\textit{en} vs. Mismatch-\textit{er}) was entered as a within-subjects variable. Irregular items were removed due to the small number of items. In order to reduce differences between the number of participants in each group, 12 participants from each Home Language group were compared. The sub-set of participants from the high-input and middle-input groups were matched on elementary school grade (1 and 2 vs. 3 and 4), gender, and school (Föhr-West vs. Föhr-East). It was not possible to match the minimal-input students, as they were, on average, older than the high-input and middle-input students. Twelve of the 13 minimal-input students were selected randomly for this comparison.

The average vocabulary accuracy and standard deviation for the sub-set of data which was included in the ANOVA statistic is presented in Table 5-4.

Table 5-4. Sub-set of vocabulary accuracy by Home Language and Category

<table>
<thead>
<tr>
<th>Home Language</th>
<th>Match</th>
<th>Mismatch-\textit{en}</th>
<th>Mismatch-\textit{er}</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M(\text{SD})$</td>
<td>$M(\text{SD})$</td>
<td>$M(\text{SD})$</td>
</tr>
<tr>
<td>High-input $N = 12$</td>
<td>0.63 (0.15)</td>
<td>0.51 (0.12)</td>
<td>0.54 (0.10)</td>
</tr>
<tr>
<td>Middle-input $N = 12$</td>
<td>0.33 (0.22)</td>
<td>0.29 (0.20)</td>
<td>0.40 (0.25)</td>
</tr>
<tr>
<td>Minimal-input $N = 12$</td>
<td>0.14 (0.16)</td>
<td>0.11 (0.11)</td>
<td>0.20 (0.15)</td>
</tr>
</tbody>
</table>

The ANOVA analysis revealed a main effect of Category: $F(2, 66) = 6.096, p = .004$, partial $\eta^2 = .156$. Post-hoc pairwise comparisons revealed that there were more word errors for the words in the Mismatch-\textit{en} category than the other two categories. Compared to words in the Mismatch-\textit{en} category, vocabulary accuracy was marginally
significantly higher for words in the Match category \((p = .056)\), and significantly higher for words in the Mismatch-\textit{er} category \((p = .008)\). The words in the Match category and Mismatch-\textit{er} category were not significantly different in terms of vocabulary accuracy \((p = 1.000)\).

There was also a main effect of Home Language: \(F(2, 33) = 21.665, p < .001\), partial \(\eta^2 = .568\). Follow-up comparisons indicated a cline of vocabulary accuracy. High-input students produced significantly fewer word errors than either middle-input or minimal-input students (vs. Middle-input: \(p = .003\); vs. Minimal-input: \(p < .001\)), and middle-input students produced fewer word errors than their minimal-input peers \((p = .013)\).

The main effects were moderated by a marginally significant interaction between Category and Home Language: \(F(4, 66) = 2.452, p = .054\), partial \(\eta^2 = .129\). Figure 5-2 illustrates the interaction. The marginally significant interaction reflects that the high-input students were more accurate on vocabulary in the Match category than the other two groups. Thus, vocabulary accuracy is influenced by both the type of plural ending and by the amount of Frisian that students hear at home.
Continuing with the investigation of how the amount of input influences vocabulary accuracy, high-input bilinguals from three age groups—elementary school, high school, and adult—were compared. The age group comparison tests differences in cumulative language input. The average vocabulary accuracy and standard deviations for each Age Group and Category are presented in Table 5-5. In general, as expected, the elementary school students produce more word errors than the high school students or the adults.

Figure 5-2. Plural vocabulary accuracy: Home language comparison

Age Group

Continuing with the investigation of how the amount of input influences vocabulary accuracy, high-input bilinguals from three age groups—elementary school, high school, and adult—were compared. The age group comparison tests differences in cumulative language input. The average vocabulary accuracy and standard deviations for each Age Group and Category are presented in Table 5-5. In general, as expected, the elementary school students produce more word errors than the high school students or the adults.
To test for differences in vocabulary accuracy between high-input bilinguals in different age groups, a mixed 3 X 3 ANOVA was conducted using vocabulary accuracy as the dependent variable. Paralleling the Home Language comparison, the irregular plural items were removed from the analysis due to differences between the numbers of items. The same sub-set of 12 high-input elementary school students was used for the comparison in order to minimize differences in the number of participants in each group. Age Group (Elementary vs. High School vs. Adult) was entered as a between subjects variable and Category (Match vs. Mismatch-\textit{en} vs. Mismatch-\textit{er}) was entered as a within subjects variable. Table 5-6 presents the descriptive statistics for the data included in the ANOVA analysis.

Table 5-5. Vocabulary accuracy by Age Group and Category

<table>
<thead>
<tr>
<th>Age Group</th>
<th>All Items</th>
<th>Match</th>
<th>Mismatch-\textit{en}</th>
<th>Mismatch-\textit{er}</th>
<th>Irregular</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M(SD)$</td>
<td>$M(SD)$</td>
<td>$M(SD)$</td>
<td>$M(SD)$</td>
<td>$M(SD)$</td>
</tr>
<tr>
<td>High-input</td>
<td>0.55</td>
<td>0.57</td>
<td>0.46</td>
<td>0.51</td>
<td>0.77</td>
</tr>
<tr>
<td>Elementary $N=21$</td>
<td>(0.10)</td>
<td>(0.17)</td>
<td>(0.14)</td>
<td>(0.12)</td>
<td>(0.16)</td>
</tr>
<tr>
<td>High School</td>
<td>0.79</td>
<td>0.85</td>
<td>0.71</td>
<td>0.78</td>
<td>0.89</td>
</tr>
<tr>
<td>$N=9$</td>
<td>(0.19)</td>
<td>(0.24)</td>
<td>(0.20)</td>
<td>(0.20)</td>
<td>(0.25)</td>
</tr>
<tr>
<td>Adult</td>
<td>0.85</td>
<td>0.94</td>
<td>0.79</td>
<td>0.78</td>
<td>0.92</td>
</tr>
<tr>
<td>$N=13$</td>
<td>(0.08)</td>
<td>(0.11)</td>
<td>(0.13)</td>
<td>(0.12)</td>
<td>(0.16)</td>
</tr>
</tbody>
</table>
Table 5-6. Sub-set of vocabulary accuracy by Age Group and Category

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Match</th>
<th>Mismatch-en</th>
<th>Mismatch-er</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-input Elementary</td>
<td>M (SD)</td>
<td>M (SD)</td>
<td>M (SD)</td>
</tr>
<tr>
<td>N = 12</td>
<td>0.63 (0.15)</td>
<td>0.51 (0.12)</td>
<td>0.54 (0.10)</td>
</tr>
<tr>
<td>High School</td>
<td>N = 9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.83 (0.24)</td>
<td>0.81 (0.18)</td>
<td>0.70 (0.19)</td>
</tr>
<tr>
<td>Adult</td>
<td>N = 13</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.94 (0.10)</td>
<td>0.77 (0.20)</td>
<td>0.78 (0.10)</td>
</tr>
</tbody>
</table>

The analysis revealed a main effect of Category: $F(2, 62) = 16.546, \ p < .001$, partial $\eta^2 = .348$. Post-hoc comparisons revealed that there were significantly fewer word errors in the Match category than either the Mismatch-en category ($p < .001$) or the Mismatch-er category ($p = .001$). There were similar rates of word errors in the Mismatch-en and Mismatch-er categories ($p = .706$).

There was also a main effect of Age Group: $F(2, 31) = 16.269, \ p < .001$, partial $\eta^2 = .512$. Follow-up comparisons revealed that the main effect was driven by significantly more word errors among elementary school students compared to both high school students and adults (vs. High School: $p = .001$; vs. Adult $p < .001$). The difference between high school students and adults was not significant ($p = .537$), indicating that adults and high school students produce similar rates of word errors. There was not an interaction between Category and Age Group: $F < 1$.

To summarize both comparisons: Among elementary school students, high-input students produced significantly fewer word errors than either middle-input or minimal-input students, and the middle-input students outperformed the minimal-input students.
When comparing high-input bilinguals across three age groups, the high-input elementary school students produced more word errors than either the high school students or adults. The high school students, however, show similar levels of vocabulary accuracy to the adult participants. Together, these results suggest that both the age of the participant and the amount of Frisian spoken at home influence how likely participants are to produce a word error. Vocabulary accuracy is also different across different categories of nouns, with the most word errors occurring in the Mismatch-\textit{en} category.

\textbf{Plural Accuracy}

Continuing with the question of how the amount of Frisian input influences the production of plural marking, plural accuracy—how often participants correctly mark plural on correct vocabulary items—was analyzed. Unlike grammatical gender, plural marking is produced consistently by adult speakers of Fering, which makes it possible to determine if responses are correct or incorrect. For example, adult speakers agree with each other and with dictionary plural markings that \textit{wel} ‘bike’ is marked with the –\textit{en} suffix: \textit{wel-en} ‘bikes’. Therefore, for –\textit{en} words, –\textit{en} responses will be considered correct responses, while all other plural markings, will be considered errors. For all –\textit{er} words, -\textit{er} responses will be considered correct, and all other responses will be considered incorrect. As the vocabulary accuracy analysis indicated that the amount of Frisian input and age group influenced participants’ word errors, the plural accuracy analysis only includes plural responses for correct word responses. This isolates plural accuracy and controls for differences in vocabulary accuracy. As with the vocabulary accuracy
analysis, the role of quantity of input is tested through a Home Language comparison and an Age Group comparison.

*Home Language*

The plural accuracy results for elementary school students with different amounts of Frisian input are presented first. All word errors were removed from the analysis. This excludes 45% of responses by high-input students, 62% of responses by middle-input students, and 80% of responses by minimal-input students (see the vocabulary accuracy analysis for more detail). Four students with no correct word responses were removed, two of whom were in the middle-input group, and two of whom were in the minimal-input group. Table 5-7 presents, descriptively, the results for all participants by Home Language and Category. Overall, two trends emerge from the descriptive results. First, when considering only correct vocabulary responses, students with more Frisian input are more accurate on plural markings than students with less input. Secondly, overall, students are the least accurate on irregular plural items.

Table 5-7. Plural accuracy by Home Language and Category

<table>
<thead>
<tr>
<th>Home Language</th>
<th>All Items M(SD)</th>
<th>Match M(SD)</th>
<th>Mismatch-en M(SD)</th>
<th>Mismatch-er M(SD)</th>
<th>Irregular M(SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-input</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N = 21</td>
<td>0.87 (0.07)</td>
<td>0.99 (0.04)</td>
<td>0.97 (0.07)</td>
<td>0.79 (0.19)</td>
<td>0.70 (0.21)</td>
</tr>
<tr>
<td>Middle-input</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N = 12</td>
<td>0.65 (0.29)</td>
<td>0.76 (0.39)</td>
<td>0.78 (0.33)</td>
<td>0.64 (0.35)</td>
<td>0.39 (0.27)</td>
</tr>
</tbody>
</table>
A mixed 3 X 3 ANOVA was conducted using plural accuracy on correct vocabulary items as the dependent variable. As with the vocabulary accuracy analysis, the irregular items were removed due to the difference in the number of items per category. The same sub-set of 12 participants per group were included in the analysis, however, due to the calculations of the ANOVA, participants who produced no accurate vocabulary in any one category (Match or Mismatch-*en* or Mismatch-*er*) were removed. Thus, the final results reported here include data from 12 high-input students, 10 middle-input students and five minimal-input students.\(^{12}\) The descriptive statistics for the data entered into the ANOVA are presented in Table 5-8. Home Language (High-input vs. Middle-input vs. Minimal-input) was entered a between-subjects variable, and Category (Match vs. Mismatch-*en* vs. Mismatch-*er*) was entered a within-subjects variable.

Table 5-8. Subset of plural accuracy by Home Language and Category

<table>
<thead>
<tr>
<th>Home Language</th>
<th>Match</th>
<th>Mismatch-<em>en</em></th>
<th>Mismatch-<em>er</em></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M(SD)</td>
<td>M(SD)</td>
<td>M(SD)</td>
</tr>
<tr>
<td>High-input N = 12</td>
<td>0.98 (0.06)</td>
<td>0.95 (0.09)</td>
<td>0.81 (0.17)</td>
</tr>
<tr>
<td>Middle-input N = 10</td>
<td>0.82 (0.34)</td>
<td>0.80 (0.33)</td>
<td>0.70 (0.32)</td>
</tr>
<tr>
<td>Minimal-input N = 5</td>
<td>0.80 (0.30)</td>
<td>0.60 (0.55)</td>
<td>0.45 (0.30)</td>
</tr>
</tbody>
</table>

\(^{12}\) The results are similar if all participants are included, rather than just matched participants. As the results were similar, the results of the matched participants were presented to parallel the vocabulary accuracy analysis.
The analysis revealed a main effect of Category: $F(2, 48) = 7.990, p = .001$ partial $\eta^2 = .248$. Follow-up pair-wise comparisons show that participants are significantly less accurate in marking plural on words in the Mismatch-er category than the Match category ($p = .002$) or the Mismatch-en category ($p = .048$). There was not a significant difference on plural accuracy between words in the Match category and the Mismatch-en category ($p = .504$).

The amount of Frisian input had a marginally significant impact on participants’ plural accuracy: $F(2, 24) = 3.335, p = 0.054$, partial $\eta^2 = .218$. Follow-up comparisons showed that high-input students showed similar plural accuracy as middle-input students ($p = .307$), but high-input students were significantly more accurate in plural markings than minimal-input students ($p = .049$). Middle-input students and minimal-input students showed similar plural accuracy ($p = .429$). There was not a significant interaction between Home Language and Category: $F < 1$.

**Age Group**

High-input bilinguals’ plural marking was analyzed by comparing the plural accuracy from three groups of high-input bilinguals: high-input elementary school students, high school students and adults. As with the vocabulary accuracy analysis, the age group comparison tests the role of the cumulative amount of language input. The descriptive results for each age group and category are presented in Table 5-9.

---

13 A post-hoc non-parametric comparison of the plural accuracy using Kruskal-Wallis test was conducted and revealed a similar significant effect of Home Language $\chi^2 (2, N = 44) = 16.143, p < .001$. 
All word errors were removed. This removed 45% of responses by high-input elementary school students, 21% of high school responses, and 15% of adult responses (see the vocabulary accuracy analysis for further details). No participants were removed due to low vocabulary accuracy. The descriptive results show that high-input elementary school students, high school students, and adults all show similar plural accuracy overall. However, elementary school students show lower plural accuracy in the Mismatch-er and irregular categories than the adults. The high school students also show lower accuracy on the Mismatch-er category, but are at or above 80% accuracy for all categories.

Table 5-9. Plural accuracy by Age Group and Category

<table>
<thead>
<tr>
<th>Age Group</th>
<th>All Items</th>
<th>Match</th>
<th>Mismatch-en</th>
<th>Mismatch-er</th>
<th>Irregular</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M(SD)$</td>
<td>$M(SD)$</td>
<td>$M(SD)$</td>
<td>$M(SD)$</td>
<td>$M(SD)$</td>
</tr>
<tr>
<td>High-input Elementary</td>
<td>0.87 (0.07)</td>
<td>0.99 (0.04)</td>
<td>0.97 (0.07)</td>
<td>0.79 (0.19)</td>
<td>0.70 (0.21)</td>
</tr>
<tr>
<td>High School $N = 9$</td>
<td>0.86 (0.28)</td>
<td>0.94 (0.17)</td>
<td>0.85 (0.33)</td>
<td>0.80 (0.32)</td>
<td>0.83 (0.33)</td>
</tr>
<tr>
<td>Adult $N = 13$</td>
<td>0.96 (0.04)</td>
<td>0.99 (0.03)</td>
<td>0.95 (0.10)</td>
<td>0.96 (0.08)</td>
<td>0.94 (0.11)</td>
</tr>
</tbody>
</table>

A mixed 3 X 3 ANOVA was conducted using plural accuracy as the dependent measure. As with the vocabulary accuracy analysis, the irregular items were removed due to differences in the number of items. The same subset of high-input elementary school students was used as in the Home Language comparison, in order to minimize differences in participant numbers. The descriptive statistics for the data included in the ANOVA are presented in Table 5-10.
Table 5-10. Sub-set of plural accuracy by Age Group and Category

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Match</th>
<th>Mismatch-en</th>
<th>Mismatch-er</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M(SD)</td>
<td>M(SD)</td>
<td>M(SD)</td>
</tr>
<tr>
<td>High-input elementary</td>
<td>0.98 (0.06)</td>
<td>0.95 (0.09)</td>
<td>0.81 (0.17)</td>
</tr>
<tr>
<td>High School</td>
<td>0.94 (0.17)</td>
<td>0.85 (0.33)</td>
<td>0.80 (0.32)</td>
</tr>
<tr>
<td>Adult</td>
<td>0.99 (0.03)</td>
<td>0.95 (0.10)</td>
<td>0.96 (0.08)</td>
</tr>
</tbody>
</table>

The analysis revealed a main effect of Category: $F(2, 62) = 12.533, p < .001$, partial $\eta^2 = .288$. Follow-up pair-wise comparisons revealed a cline of accuracy, with words in the Match category marginally more accurate than words in the Mismatch-en category; words in the Mismatch-er category were the least accurate: Match versus Mismatch-en $p = .059$; Match versus Mismatch-er $p < .001$; Mismatch-en versus Mismatch-er $p = .037$.

When comparing plural accuracy on only accurate word responses there is not a significant difference among Age Groups: $F(2, 31) = 1.289, p = .290$. However, the main effect of Category is moderated by a significant interaction between Age Group and Category: $F(4, 62) = 3.071, p = .023$, partial $\eta^2 = .165$. The interaction is driven by significant differences between categories for the elementary school students and high school students. Figure 5-3 illustrates the interaction between Age Group and Category for plural accuracy. The elementary school students had significantly higher plural accuracy in the Match category and Mismatch-en category than the Mismatch-er category: Match versus Mismatch-er ($p < .001$); Mismatch-en versus Mismatch-er ($p =$
Elementary school students show similar plural accuracy in the Match category and Mismatch-\textit{en} category ($p = .416$). For high school students, the words in the Match category show higher plural accuracy than both the Mismatch-\textit{en} words ($p = .036$) and the words in the Mismatch-\textit{er} category ($p = .005$). There was not a significant difference between words in the Mismatch-\textit{er} and Mismatch-\textit{en} categories ($p = .292$). The adults show similarly high plural accuracy across categories (all $ps > .05$). Thus, participants’ plural accuracy depends both on the category of word and on their age group.

To summarize both comparisons: for elementary school students, when considering only accurate word responses, students are more accurate in marking plural
when the words have either the same plural marking in both Frisian and German or have the –en marking in Frisian than when words have the –er ending in Frisian. The amount of Frisian input for elementary school students also influences how accurately they use plural markings, with both high-input and middle-input students more accurate in plural markings than minimal-input students. When comparing the high-input bilinguals in three age groups, all three groups show similar overall plural accuracy, but elementary school students and high school students are also influenced by the particular plural marking category for a word. The elementary school students showed lower plural accuracy on the words in the Mismatch-er category than either the Match or the Mismatch-en category. The high school students were more accurate on the words in the Match category than either the words either of the Mismatch categories. Overall, even when controlling for differences in vocabulary accuracy, participants with more Frisian input produce plural markings more accurately than those with less Frisian input.

**Regularization**

The second broad research question asks whether participants regularize plural markings, similar to the results of the grammatical gender production task (see Chapter 4). The goal was to determine whether participants over-use one plural marking for all, or nearly all, items. In order to investigate whether the participants regularized plural markings, a similar method was used as in the grammatical gender regularization analysis. The number of participants who used one plural marking almost exclusively was calculated. If a participant marked all, or all but one, item with the same plural marking, she was considered a *regularized speaker*. If a participant used a range of plural
markings, she was considered an inconsistent speaker. All responses with a Frisian plural marking (–en, -er, or irregular) were included, regardless of whether the response was a correct word or word error. All responses without Frisian plural markings were removed, such as singular responses, such as skap ‘ship’ for skep ‘ships’. Only participants who produced more than half of the items with a Frisian plural marking were included in the analysis. Two middle-input and one minimal-input elementary school students were removed because they did not produce at least half of the items with Frisian plural markings. The goal of this analysis is to investigate whether participants show a general tendency to mark plurals with one particular plural marking.

Following these criteria, only two participants qualify as regularized speakers. Both of the participants are elementary school students, and one is a minimal-input student and one is a middle-input student. The minimal-input student produced only one correct word and the rest of responses were German borrowings, and therefore should not be considered a regularized speaker of Frisian. Of the 11 correct word responses that the middle-input student produced, ten were marked with the –en plural marking, and one with the –er marking. Therefore, there was one participant who is considered a regularized speaker of Frisian, notably fewer than the number of regularized speakers in the grammatical gender analysis.

Despite the fact that only one participant qualified as a regularized speaker according to the same criteria used on the grammatical gender production task, the use of such stringent criteria may hide more subtle tendencies to favor one plural marking over another. Additionally, because the items in the Match category were predominately –en words, it is difficult to determine from the accuracy results whether participants are
particularly good at producing plural markings when they match the plural marking of the corresponding German word, or if participants favor the–en marking as a default plural marking overall (see also their higher accuracy on Mismatch-en items). A more detailed analysis of the frequency of the type of responses across items will investigate these possibilities. Responses were separated into three categories based on the plural marking: -en responses, -er responses, and irregular plural responses. All responses with a Frisian plural marking, including German borrowings, code-switches, and Frisian word errors, are included in the regularization analysis. For example, Hose-n ‘pants’ was included in the –en category because it is marked with a –en suffix, even though it is a German code-switch.

The proportion of each response type was calculated for each speaker. Figure 5-4 shows the proportion of all responses for each Home Language and Age Group. There appears to be a tendency for speakers with less Frisian input to rely on –en plural markings more than speakers with more Frisian input. However, even the speakers in the minimal-input group do not exclusively mark plural with the –en ending, using –en marking for 78% of items, and –er marking for 21% of items. Overall, the elementary school students favor the –en marking, and the students with more language input rely on –en marking less than those with more input. However, the elementary school students did not use one marking exclusively, as many students did in the grammatical gender task.
To summarize: Following the methodology in the grammatical gender production task, there was only one middle-input speaker who qualifies as a regularized speaker, far fewer than the number of regularized speakers in the grammatical gender task. A follow-up analysis of the proportion of each plural marking showed a tendency for lower-input speakers to use –en markings more often than higher-input speakers, however, they do not use any one plural marking exclusively.

5.2.5 Discussion

The plural elicitation task addressed both of the two broad research questions by answering three more specific research questions. The results of the plural elicitation task supported the adult survey results (See Chapter 3) in showing that adults produce plural markings consistently in contemporary Fering. This suggests that the inconsistency in
grammatical gender marking does not extend to other morphological features of Fering, but rather is specific to grammatical gender. Moreover, the consistency in adult plural marking means that the input available to children is also consistent, in that certain nouns always occur with –en and other nouns always occur with –er.

The results of both the Home Language comparison and the Age Group comparison provide converging evidence that, as predicted, participants with more Frisian input produce fewer word errors than participants with less Frisian input. The influence of quantity of input can also be seen in the plural accuracy analysis, in which only the plural markings on correct word responses are considered. When considering correct words, high-input students were most accurate on plural markings, followed by the middle-input students and then the minimal-input students.

Overall, when considering only correct word responses, higher input participants’ responses show that they are on their way to producing the adult-like pattern of plural markings. The general pattern can be seen in the Home Language comparison and the Age Group comparison. The high-input students and middle-input students did not differ in plural accuracy, which suggests that between 25% and 50% language input may be sufficient for acquiring consistently produced morphology, such as plural markings in Fering.

The comparison between high-input participants at three different age points – elementary, high school, and adult – shows similar results across age-groups. However, there was a significant interaction between category and age group, which suggests that by elementary school, high-input bilinguals are producing plural markings in adult-like ways for words in the Match and Mismatch-en category, but are not yet adult-like for the
Mismatch-er category. This pattern of results suggests that the students with less cumulative Frisian input rely on –en plural markings as a default more than the elementary school students with more Frisian input. These results contrast with the grammatical gender results, which found that elementary school students were not yet adult-like in their production. Thus, when given consistent input, high-input bilinguals are successful at producing adult-like patterns.

One of the key findings of the plural marking task is that children do not regularize consistently produced morphology, using the same strict criteria as the grammatical gender task. That is, they do not use one plural marking exclusively. Only one of the participants regularized the plural markings, according to the same criteria as the grammatical gender task and Hudson Kam and Newport (2005, 2009). However, the follow-up analysis of the frequency of responses type showed that this strict criteria for regularization masks a tendency to over-use the –en marking, and that the lower-input participants tended to use –en more often than higher-input participants. This suggests that elementary school students may initially over-use the –en marking before fully acquiring the adult-like plural system.

Researchers and textbooks have posited a strong connection between plural markings and grammatical gender in Fering (Arftsen & Tadsen, 2009; Wilts, 2011). Therefore it is possible that the changes in the grammatical gender system also influence the plural marking system. It is important to note, however, that this change runs counter to the historical patterns. Traditionally, masculine words were marked with –er plural marking and common gender words were marked with –en plural marking. The data in this dissertation shown that speakers favor the masculine gender, which was previous
associated with the –er ending, while speakers also favor the –en ending, which had been associated with the common gender, which may be disappearing in Fering. Therefore it is possible that as the grammatical gender system changes, the connection between plural marking and grammatical gender will diminish—a possibility that will be discussed in greater detail in Chapter 7.

Taken together, the results of the plural marking task show that with more input, speakers become more adult-like in vocabulary accuracy and plural-marking accuracy. Unlike the grammatical gender results, even though the lower participants struggle to produce the correct vocabulary and accurate plural markings at similar rates as higher-input participants, they do not use exclusively one plural marking. Therefore, regularization, according to the strict criteria in Hudson Kam and Newport (2005, 2009), appears to be limited to inconsistently produced input, rather than representing a general trend in Frisian acquisition.

5.2.6 Summary

The plural elicitation task probed which plural markings participants use with specific nouns. In doing so, it tested participants’ acquisition of noun to plural marking correspondences. Supporting the adult survey results, the adult responses show that they produce plural markings consistently and that they largely agree on the plural marking associated with a given noun. Across both the Home Language comparison and the Age Group comparison, higher input speakers produce fewer word errors and are more accurate in producing plural markings than lower input speakers, suggesting that the
amount of input children have access to affects both their acquisition of vocabulary and plural markings.

A comparison between the three elementary-aged participant groups showed that students with more Frisian input at home are more accurate than middle or minimal-input students. Similarly, the analysis by age group showed that high-input elementary school students are indistinguishable from both high school and adult participants in their patterns of –en production, but that they still may lag behind in the accurate production of –er plural markings. The high-input elementary school students’ pattern of production suggests that they over-use the –en plural marking, although they do not use it exclusively. That is, while students tend to regularize inconsistent input, such as grammatical gender, they do not meet the strict criteria of regularization for consistently produced morphology.

5.3 Acceptability Task

5.3.1 Background

In addition to production tasks, investigating how children understand language is important for fully understanding child language acquisition (Hirsh-Pasek & Golinkoff, 1998; Miller & Schmitt, 2010). Recent studies have shown that children may be able to comprehend language features prior to being able to produce the same structure. This suggests that studies that only investigate production may underestimate the child’s language skills. With this in mind, the acceptability task investigates participants’ comprehension of plural markings, specifically, whether they are sensitive to the differences between correctly and incorrectly marked plural forms.
The acceptability task addresses the first broad research question through two more specific research questions:

1a. How does the amount of Frisian input influence participants’ ability to identify correct and incorrect plural markings?

1b. How does the amount of Frisian input influence the extent to which groups of participants differentiate between correct and incorrect plural markings?

5.3.2 Participants

Thirty-two elementary students completed this task—13 high-input students, 12 middle-input students, and seven minimal-input students\(^\text{14}\). The results of nine high-input high school students and 13 adult native speakers will be presented as comparisons.

5.3.3 Methods and design

Following methodology from Ambridge, Pine, Rowland and Young (2008) and Ambridge (2010), participants heard a series of words and were asked to point on a five-point smiley face scale, depicted in Figure 5-5, whether they thought the word was correct or incorrect. One of the advantages of this task is that it allows for testing the comprehension of single words (Ambridge, 2010).

The instructions for the task were given by the research assistant, and the stimuli, seen in 2a-c, were recorded by the research assistant and played via imbedded audio on the PowerPoint slides. Participants first saw a clipart image of a single object, along with a description, as outlined in 2a. Next, the participant saw the same clipart image in

\(^{14}\) Elementary school students completed the acceptability task on Day 2, meaning that eight high-input, two middle-input, and six minimal-input students did not complete this task.
triplicate, along with a description which used a plural marking, as in 2b. On the final slide, participants saw the five-point smiley face and were asked to judge whether the description the research assistant provided for the plural depiction was correct, as in 2c. If participants asked for clarification, they were told to base their judgment on the final word in 2b, but were not told explicitly to focus on plural markings. Participants pointed to the face that corresponded with their judgment. Before the task began, the research assistant provided a description of the meaning of each face, and the primary researcher repeated the description while pointing to each face. The complete instructions for this task are presented in Appendix A. Participants were then asked if they had any questions.

(2) a. *Heer as een tus.*

Here is a tooth


And then here? Do you see now? Now there are three teeth.

c. *Haa ik det rocht saad?*

Did I say that right?

Figure 5-5. Acceptability Scale

The audio stimuli were recorded by the same native speaker research assistant who interacted with participants. Across participants, each item appeared in both correct
and incorrect versions. Items were counter-balanced across two lists, such that no
participant ever heard the same item as both incorrect and correct. Participants were
randomly assigned to one of the two lists. Over the course of the task, participants heard
an equal number of correct and incorrect plural markings. The same recording was used
for the initial verbal introduction of the item 2a and the acceptability prompt 2c,
regardless of whether the trial contained a correct or incorrect plural marking.

5.3.4 Results

Acceptability Accuracy

In order to perform accurately on the task, a participant must accurately judge
correct items as correct and reject incorrect items as incorrect. In order to calculate
accuracy, the responses on the five point smiley-face scale were converted into three
responses: acceptable judgments (including the two blue happy faces), unacceptable
judgments (including the two red sad faces), and indeterminate judgments (includes the
half red/half blue face). Accuracy was determined as follows: Both acceptable judgments
for correct versions and unacceptable judgments of incorrect versions were considered
accurate responses. Unacceptable judgments of correct versions and acceptable
judgments of incorrect versions were considered inaccurate responses. Indeterminate
responses were considered incorrect responses.

In order to test for the influence of amount of input, participants’ accuracy will be
compared within elementary school students with different amounts of input, and within
high-input bilinguals from three age groups. Following the general prediction that
students with more input will be more adult-like in plural accuracy, I predict that higher
input students will be more accurate at identifying correct and incorrect forms than lower input students.

**Home Language**

Elementary school students with three levels of home language input were compared to test for differences based on the quantity of input. As seen in Table 5-11, elementary school students are at or close to chance level (50%) for each category across all three home language groups.

Table 5-11. Acceptability accuracy by Home Language and Category

<table>
<thead>
<tr>
<th>Home Language</th>
<th>All Items</th>
<th>Match</th>
<th>Mismatch-en</th>
<th>Mismatch-er</th>
<th>Irregular</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M(SD)$</td>
<td>$M(SD)$</td>
<td>$M(SD)$</td>
<td>$M(SD)$</td>
<td>$M(SD)$</td>
</tr>
<tr>
<td>High-input</td>
<td>(N = 13)</td>
<td>0.59 (0.18)</td>
<td>0.58 (0.22)</td>
<td>0.55 (0.26)</td>
<td>0.60 (0.20)</td>
</tr>
<tr>
<td>Middle-input</td>
<td>(N = 12)</td>
<td>0.51 (0.17)</td>
<td>0.51 (0.20)</td>
<td>0.57 (0.18)</td>
<td>0.49 (0.20)</td>
</tr>
<tr>
<td>Minimal-input</td>
<td>(N = 7)</td>
<td>0.51 (0.10)</td>
<td>0.48 (0.19)</td>
<td>0.50 (0.19)</td>
<td>0.54 (0.13)</td>
</tr>
</tbody>
</table>

A 3 X 3 Mixed ANOVA using acceptability accuracy as the dependent variable was conducted. As all participants were included in this comparison, the data presented in Table 5-11 represent the data in the ANOVA analysis. Irregular plural items were removed due to the differences in the numbers of items in each category. The main effect of the between subjects variable of Home Language group (High-input vs. Middle-input
vs. Minimal-input) was not significant $F < 1$. The within subjects variable of word category (Match vs. Mismatch-<i>en</i> vs. Mismatch-<i>er</i>) was not significant: $F < 1$. The interaction between Home Language and Category was not significant $F < 1$. Taken together, these results show that participants perform similarly on all words and across all home language groups.

<em>Age Group</em>

Looking at high-input participants at three age points tests whether participants show improved plural comprehension with cumulatively high input. Table 5-12 presents descriptive results for each age group and category.

**Table 5-12. Acceptability accuracy by Age Group and Category**

<table>
<thead>
<tr>
<th>Age Group</th>
<th>All Items</th>
<th>Match</th>
<th>Mismatch-&lt;i&gt;en&lt;/i&gt;</th>
<th>Mismatch-&lt;i&gt;er&lt;/i&gt;</th>
<th>Irregular</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M (SD)$</td>
<td>$M (SD)$</td>
<td>$M (SD)$</td>
<td>$M (SD)$</td>
<td>$M (SD)$</td>
</tr>
<tr>
<td>High-input</td>
<td>0.59 (0.18)</td>
<td>0.58 (0.22)</td>
<td>0.55 (0.26)</td>
<td>0.60 (0.20)</td>
<td>0.63 (0.28)</td>
</tr>
<tr>
<td>Elementary</td>
<td>N = 13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High School</td>
<td>0.77 (0.18)</td>
<td>0.78 (0.27)</td>
<td>0.83 (0.17)</td>
<td>0.67 (0.18)</td>
<td>0.89 (0.22)</td>
</tr>
<tr>
<td>N = 9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adult</td>
<td>0.86 (0.15)</td>
<td>0.85 (0.21)</td>
<td>0.83 (0.19)</td>
<td>0.82 (0.16)</td>
<td>0.94 (0.15)</td>
</tr>
<tr>
<td>N = 13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A mixed 3 X 3 ANOVA was conducted to compare the accuracy for each group and category. Age Group (Elementary vs. High School vs. Adult) was entered as a between
subjects variable and Category (Match vs. Mismatch-\textit{en} vs. Mismatch-\textit{er}) was entered as a within subjects variables. As all participants were used in the ANOVA analysis, the data presented in Table 5-12 represent the data entered into the ANOVA. Irregular items were not included in this analysis because there were fewer items in this category.

There was no main effect of Category: $F(2, 64) = 1.038, p = .360$, partial $\eta^2 = .031$. There was a main effect of Age Group: $F(2, 32) = 7.597, p = .002$, partial $\eta^2 = .322$. Post-hoc comparisons showed that the main effect of Age Group is driven by significantly lower accuracy among elementary school students compared to adults ($p = .002$), and marginally significantly lower accuracy compared to high school students ($p = .053$). There was not a significant difference between high school students and adults ($p = .567$). The interaction between Category and Age Group was not significant: $F(4, 64) = 1.540, p = .201$.

To summarize both comparisons: Elementary school students as a whole struggle with the acceptability task and perform at chance level across all Home Language groups and Categories. A comparison between high-input bilinguals in three age groups shows that while elementary school students are not yet adult-like in their comprehension of plural markings, high school students are not significantly different from adult speakers.

\textbf{Response Bias}

As the elementary school students as a whole struggled with the acceptability task, a follow-up analysis was conducted to see whether participants favored one type of response. The proportion of each response was calculated for each group in order to see whether participants favored either positive or negative responses. As seen in Figure 5-6,
participants overwhelmingly choose acceptable judgments over either intermediate responses or negative responses. Elementary school students show a larger bias toward marking items as more acceptable than either high school or adult speakers. The bias persists in all three groups of elementary school students. Taken together, all groups of participants favor acceptable responses, but the trend is more prominent among elementary school students, regardless of the quantity of Frisian input.

The strong response-bias may indicate that the participants did not fully understand the task methodology. To test this possibility, a follow-up analysis of two filler items was conducted. Two of the filler items included an erroneous stem change, in combination with an incorrect plural suffix, as in *kleeker, ‘clocks’, whereas the correct plural form is klook-klooken, ‘clock-clocks’. Overall, participants generally rejected these filler items. There were three elementary school students who did not rate both of these filler items as a 3 (an indeterminate rating) or below. One minimal-input student categorically marked the most acceptable face for all items, and one minimal-input and one high-input student only rejected one of these two filler items. The descriptive statistics for responses to these two double negative filler items are presented in Table 5-13.

Overall, the participants with less Frisian input are more accepting of these double negative filler items. However, the generally low ratings, and few participants who accepted these items suggests that most of the students understood how to use the scale and would use the unhappy face when they felt it was required.
Table 5-13. Average ratings for double negative filler items

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adult N = 13</td>
<td>1.19</td>
<td>.38</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>High School N = 9</td>
<td>1.22</td>
<td>.67</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>High-Input Elementary N = 13</td>
<td>1.96</td>
<td>.97</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Middle-Input Elementary N = 12</td>
<td>2.25</td>
<td>.86</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Minimal-Input Elementary N = 7</td>
<td>2.57</td>
<td>1.37</td>
<td>1</td>
<td>5</td>
</tr>
</tbody>
</table>

Figure 5-6. Proportion of responses by age and home language

**Difference Scores**

Given the overall response bias, it is not possible to determine whether the tendency to mark all stimuli as correct masks a pattern of sensitivity to differences between correct and incorrect responses. Consider a child who sits down and selects the
number 5 happy face for the first item, then the number 4 happy face for the second item, followed by the number 5 face again, and then the number 4 face, and so forth. This pattern of results could represent two scenarios. It is possible that she is reliably choosing the number 5 face for the correct items, and the number 4 face for the incorrect items, and simply not using the entire range of the scale to identify correct versus incorrect items, or she could be unaware of the differences between words, and pointing randomly to each face.

Both the 4 and 5 faces are considered “correct judgments”, meaning that the child marked correct for all items, and hence, half are accurate and half are inaccurate. Thus, in both cases, the child’s accuracy score would be at chance level, but for different reasons. In each case, it is true that the child shows a strong response bias, but the question of whether the child is sensitive to the differences between correct and incorrect stimuli or simply responding randomly remains. This is the question addressed in the analyses in this section.

The difference between the ratings of correct and incorrect versions was calculated for each item in order to determine whether participants are sensitive to plural marking violations. These difference scores were calculated as follows. First responses were converted into a five point Likert scale: happiest blue face = 5; the semi-happy blue face = 4; the half red-half blue face = 3; the semi unhappy red face = 2; the unhappiest red face = 1. Then for each item, a group’s average rating of the incorrect version was subtracted from the correct version. This means that the minimum score is -4\(^1\) meaning the average rating for the correct version was 1 and the average rating for the incorrect

\(^1\) Negative difference scores were rare.
version was 5. A difference score of 0 indicates that participants marked both correct and
incorrect versions of the item similarly and thus did not differentiate between correct and
incorrect items. For example, for the item tuun ‘fence’, seven high-input students heard
the correct plural, tuun–er ‘fences’ and rated it on average as 4.86. Six high-input
students heard the incorrect plural * tuun–en ‘fences’ and rated it on average 2.33.
Therefore the difference score for this item is 4.86 - 2.33 or 2.53. Larger positive
difference scores indicate that the participants made a greater distinction between
incorrect and correct versions; smaller difference scores indicate less differentiation
between the correct and incorrect version of a particular item.

Participants were randomly assigned to one of the two lists. The difference scores
represent the ratings by participants who completed List 1 compared to the ratings by
participants who completed List 2. Participants were split fairly evenly between lists
(High-input: List 1= 7, List 2= 6; Middle-input: List 1= 7, List 2= 5; Minimal-input: List
1 = 5; List 2 = 2; High School: List 1= 2, List 2= 7 Adult List 1= 6, List 2= 7). The
difference scores were calculated for each item and each home language group.

As a baseline for comparison, Ambridge et al. (2008) report that when adult
native speakers of English were asked to rate verbs from different semantic categories,
their difference scores ranged between 0.64 and 3.10. Younger participants (ages 5-6)
showed difference scores between 0 and 2.19, while older participants (9-10) were
similar to adult speakers, ranging between 0.75 and 3.17. Although the task in Ambridge
et al. tested different features, these results provide a range of typical performance using
the acceptability scale among children.
Home Language

The difference scores for elementary school students with different quantities of input were compared in order to test for the role of quantity of input. Table 5-14 summarizes the difference scores by category and home language group. High-input students have larger difference scores than the middle-input students and minimal-input students across all categories, suggesting that they differentiate more between correct versions of items and incorrect versions of items.

Table 5-14. Difference scores by Home Language and Category

<table>
<thead>
<tr>
<th>Home Language</th>
<th>Match</th>
<th>Mismatch-en</th>
<th>Mismatch-er</th>
<th>Irregular</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-input</td>
<td>N = 9</td>
<td>N = 6</td>
<td>N = 9</td>
<td>N = 4</td>
</tr>
<tr>
<td>M(SD)</td>
<td>M(SD)</td>
<td>M(SD)</td>
<td>M(SD)</td>
<td></td>
</tr>
<tr>
<td>High-input</td>
<td>1.29 (0.60)</td>
<td>1.35 (0.84)</td>
<td>1.41 (1.18)</td>
<td>2.46 (1.52)</td>
</tr>
<tr>
<td>Middle-input</td>
<td>0.83 (0.49)</td>
<td>1.04 (0.54)</td>
<td>0.74 (1.35)</td>
<td>1.60 (1.46)</td>
</tr>
<tr>
<td>Minimal-input</td>
<td>0.81 (0.93)</td>
<td>0.00 (0.89)</td>
<td>0.53 (1.22)</td>
<td>1.95 (1.67)</td>
</tr>
</tbody>
</table>

A mixed 3 X 3 ANOVA was used to compare the difference scores between groups of elementary school students from different Home Language Groups. As all participants were included in the analysis, the data presented in Table 5-14 reflects the data entered into the ANOVA. As with other analyses, irregular items were removed due to the differences in the number of items. Category (Match vs. Mismatch-en vs. Mismatch-er) was entered as a between-items variable, and Home Language was entered as a within-items variable.
The analysis reveals a main effect of Home Language: $F(2, 42) = 8.764, p = .001$, partial $\eta^2 = .294$, but no effect of Category ($F < 1$) or a significant interaction between Category and Home Language: $F(4, 42) = 1.136, p = .353$. Follow-up comparisons reveals that the high-input students showed significantly larger difference scores than either the middle-input ($p = .023$) or minimal-input students ($p = .004$). The middle-input students and minimal-input students showed similar difference scores ($p = .239$).

The results of the Home Language comparison suggest that the amount of Frisian input elementary school students have at home influences how much participants differentiate between correct and incorrect versions of plural markings. High-input students differentiate between correct and incorrect versions more than either middle-input or minimal-input students. The category of plural marking, however, did not influence elementary school students’ difference scores.

**Age Group**

The difference scores were then compared for high-input bilinguals in three Age Groups: Elementary school students, High School students, and Adults. This tests the role of cumulative input on acceptability difference scores. As with the Home Language comparison, for each group the average response to correct versions and the average response to incorrect versions was calculated for each item, then averages were computed for the items in each category. Table 5-15 presents the descriptive statistics for differences for each Age Group and Category.
Table 5-15. Difference scores by Age Group and Category

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Match</th>
<th>Mismatch-en</th>
<th>Mismatch-er</th>
<th>Irregular</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N = 9</td>
<td>N = 6</td>
<td>N = 9</td>
<td>N = 4</td>
</tr>
<tr>
<td>High-input Elementary</td>
<td>1.29 (0.60)</td>
<td>1.35 (0.84)</td>
<td>1.41 (1.18)</td>
<td>2.46 (1.52)</td>
</tr>
<tr>
<td>High School</td>
<td>3.25 (0.68)</td>
<td>3.08 (0.97)</td>
<td>2.56 (1.59)</td>
<td>3.38 (0.48)</td>
</tr>
<tr>
<td>Adult</td>
<td>3.08 (0.50)</td>
<td>3.13 (0.76)</td>
<td>2.82 (0.85)</td>
<td>3.67 (0.05)</td>
</tr>
</tbody>
</table>

A mixed 3 X 3 ANOVA was conducted to compare the difference scores across age groups and categories using difference scores as the dependent variable. As all participants are included in the analysis, the data in Table 5-15 reflects the data entered into the ANOVA analysis. Age Group was entered as a within-items variable and Category was entered as a between-items variable. The analysis reveals a main effect of Age Group: $F(2, 42) = 66.399, p < .001$, partial $\eta^2 = .760$. Follow-up pairwise comparisons revealed that this main effect is driven by the significantly lower difference scores among the elementary school students compared to both the high school students ($p < .001$) and adults ($p < .001$). The high school students and adults produced similar difference scores ($p = 1.00$). There was no main effect of Category ($F < 1$) and no significant interaction between Age Group and Category: $F(4, 42) = 1.278, p = .291$, partial $\eta^2 = 109$, indicating that all three age groups exhibited similar difference scores across categories.

To summarize both comparisons: Elementary school students with more Frisian input differentiate correct and incorrect versions of plural marked nouns more than
elementary school students with less Frisian input. When comparing high-input bilinguals from three age groups, elementary school students differentiate significantly less between correct and incorrect versions of plural marked nouns than either high school students or adults, but adults and high school students show similar difference scores. The category of plural marking did not influence the difference scores for either comparison, meaning that participants differentiate between correct versions and incorrect versions of plural markings similarly across all three plural marking categories. Overall, the results suggest that across all categories of items, participants are sensitive to violations of plural markings, but that the amount of Frisian input participants receive influences the degree of differentiation between correct and incorrect versions.

5.3.5 Discussion

The acceptability task addressed the first broad research question with two specific questions: 1a. How does the amount of Frisian input influence participants’ ability to identify correct and incorrect plural markings? 1b. How does the amount of Frisian input influence how much groups of participants differentiate between correct and incorrect plural markings? I will discuss the results as a means of answering these two questions, and then discuss some of the limitations of the acceptability task.

The accuracy analysis shows that adult speakers and high school students are able to identify both correctly and incorrectly marked plural forms in Fering. Elementary school students, however, struggled with this task and were less accurate than both the high school and adult participants. Elementary school students performed at chance level regardless of the quantity of Frisian they hear at home and regardless of the plural
category: Match versus Mismatch-\textit{en} versus Mismatch-\textit{er}. At the same time, the difference scores showed that the amount of Frisian input influences children’s acquisition of plural markings, paralleling the results from the production task.

\textit{Limitations}

The elementary school students as a group struggled with this task. The accuracy scores suggest that the task was difficult for elementary school students, regardless of the amount of Frisian input they had access to, and regardless of the type of item tested. Students showed a strong bias toward marking all items as acceptable. Most participants were able to correctly identify the two very wrong items, which demonstrates that they understood the task, and were willing to judge some items as incorrect. Similarly, Ambridge et al. (2008) and Ambridge (2010) showed that children at this age should be able to successfully complete this task.

There are several possibilities for why the elementary school students struggled with the task. In contrast to previous studies, the audio recordings of the stimuli were made by the same native speaker as the research assistant who conducted the task. If students noticed that the audio recordings were the same as the research assistant, they may have been reluctant to say that the speaker made an error. Another possibility is that if children were overall unsure of their responses they may have opted to mark correct, a tendency which has been shown in adult second language learners (e.g., Sabourin, Stowe, & de Haan, 2006).

While the difference score analysis allows for testing acceptability independently of response bias, it is not, however, without fault. Since participants did not hear both
correct and incorrect versions of each item, it was always necessary to compare a sub-group of participants’ ratings to a second sub-group of participants. This left few ratings, and created arbitrary categories of participants. However, since participants were randomly assigned to each experimental list, it is unlikely that participants in each sub-group differed systematically from each other.

Despite these limitations, the difference scores analysis supports the production results in underscoring the importance of quantity of input in plural acquisition. Both within the elementary school participants and across three age groups of high-input bilinguals, participants with more language input are able to make a larger distinction between correct and incorrect versions of plural forms. Thus, testing comprehension in addition to production of plurals provides converging evidence that the quantity of input influences children’s acquisition of plural markings.

5.3.6 Summary

The acceptability task tested participants’ comprehension of plural markings in Fering by probing their ability to judge correctly and incorrectly marked plural forms. Participants heard words in the singular form and then with a plural inflection and were asked to rate the plural form on a five point scale. Overall, participants favored marking forms as acceptable rather than unacceptable, and this trend was especially clear among elementary school students. The overall bias toward acceptable responses obscured the accuracy results, where elementary school students performed at chance level regardless of their home language or the word type tested.
A post-hoc analysis of the difference scores—the average ratings of correct versions of items minus the average ratings of incorrect versions—showed that participants with more Frisian input differentiate more between correct and incorrect versions of plural forms. This was true within elementary school, where the high-input and middle-input students had larger difference scores than minimal-input students, and across three age groups of high-input bilinguals, where the high school students and adults had larger difference scores.

5.4 Summary and Discussion

This chapter tested participants’ acquisition of plural markings in Fering through two tasks: an elicitation task and an acceptability task. These two tasks answered the following two broader research questions:

1. How does the amount of Frisian language input influence participants’ acquisition of plural markings?

2. Do children regularization plural markings in Fering, as they do grammatical gender?

The plural elicitation task addressed both of the broader research questions. The adult responses support the pilot studies showing that adult native speakers produce plural markings consistently, as the adult participants in the elicitation task agreed on plural endings 96% of the time, and there were no words for which fewer than 75% of the adults agreed on the plural marking. From this, one can conclude that the input children hear vis-à-vis plural markings in Fering is reliable and consistent.

Overall, across both the home language comparison and age group comparison, the results show a strong influence of the amount of Frisian input on the production of
Frisian vocabulary and plural markings. High-input bilinguals from all three age groups are similarly accurate on plural markings, although the high school students and elementary school students show differential accuracy across categories. The elementary school students were less accurate on the Mismatch-\textit{er} category than on the Match or Mismatch-\textit{en} category, while the high school students were more accurate on the Match category than either of the Mismatch categories. Taken together, the results support the general prediction that participants with more Frisian input will be more adult-like in their production of plural markings.

In terms of regularization, the production task shows that when children are exposed to consistent input they do not fit a narrow definition of regularization, meaning that they do not produce exclusively one plural marking. Elementary school students show a tendency to favor the more frequent plural marking \textit{–en}, including overextending the \textit{–en} ending on words that should be marked with \textit{–er}. However, this tendency does not approach regularization, at least as defined by Hudson Kam and Newport (2005, 2009), with even the minimal-input students using \textit{–er} marking for an average of 20\% of items with a Frisian plural marking. Thus, narrowly defined regularization appears to be limited to cases where children are exposed to inconsistent input, as seen in grammatical gender in Fering.

The acceptability task tested student’s comprehension of plural markings by asking them to rate plural forms with correct and incorrect suffixes using a five-point smiley acceptability scale (Ambridge, 2010; Ambridge et al., 2008). The acceptability task addresses the first broader research question. The results of the acceptability task show that adult participants are highly accurate, meaning that they identify correctly-
marked plural forms, and rejected incorrectly-marked forms. In contrast, elementary school students are overall more accepting of both correct and incorrect responses, leading to chance level performance across all categories and home language groups. Follow-up analyses of the differences between judgments of correct and incorrect versions, however, show that high-input students and middle-input students rate correct plural forms more acceptable than incorrect forms. Taken together, the acceptability task shows that elementary school students show a strong “yes-bias”, which obscures the overall accuracy results, but that difference scores suggest that high-input elementary school students are sensitive to errors in plural marking on some level. This suggests that while high-input elementary school students have not yet fully mastered the plural system, they are sensitive to violations of plural markings.

These results support Gathercole and Thomas (2009) in highlighting the importance of quantity of input in minority language contexts. Even when children are exposed to consistent input they require a substantial amount of input to acquire morphological marking. The middle-input students are bilingual from birth and most often come from “one-parent one-language” families. These results suggest, then, that even roughly balanced exposure at home may not be sufficient for acquiring certain features of a minority language.

The differences between the vocabulary accuracy, plural accuracy, and the acceptability task show that it is important to consider how successful acquisition will be defined and measured by the community. This is especially important for language planners and community members to discuss when designing language programs. Defining exactly what the goal of the language promotion program is ensures that
planners are able to determine how much and what type of input is needed and to determine whether the language program is successful. In the case of Fering, if the goal of Frisian classes is to promote the use of native Frisian vocabulary, as teachers say it is, then the middle and minimal input kids may need more input to approximate their high-input peers. However, if the goal is to accurately use plural markings when children do produce a Frisian word, then students in the middle-input group appear to have sufficient input, but the minimal-input students require additional input. While the elementary-aged children in the present study were separated into groups according to the quantity of Frisian at home, additional input from school or community contexts may also be able to boost acquisition among children from lower-input families.

Broadly speaking, the results of the two plural tasks show that children are able to successfully acquire plural markings in Fering, but they require substantial input to master the noun-to-plural marking associations. Elementary school students, especially those with less input, tend to over-rely on the –en plural marking. However, even those with minimal input do not use the –en marking exclusively. This suggests that consistently produced plural markings may be acquired differently than inconsistently produced grammatical gender markings.
Chapter 6. Results: General Frisian Proficiency

6.1 Introduction

While the previous two chapters investigated how the amount of input influences the acquisition of specific morphological features — grammatical gender in Chapter 4 and plural markings in Chapter 5— this chapter investigates whether children with different amounts of input achieve proficiency in Frisian more generally. Language revitalization and promotion efforts often focus on whether the children speak the language generally, rather than focusing on specific grammatical forms (Grenoble & Whaley, 2006; Hinton & Hale, 2001). This is also true within the Fering community, where Frisian language teachers report that they focus on teaching general language skills and native Frisian vocabulary (rather than German borrowings). With this in mind, this chapter focuses on how the quantity of language input influences children’s general Frisian proficiency.

In this chapter I describe the results of a story-telling task, which tests participants’ Frisian proficiency with a more naturalistic task than the other experimental tasks. Throughout the chapter, I use the term proficiency to refer to speaker’s general language skills. This means that the speaker is able to both understand and speak in Frisian. Although reading and writing are taught in Frisian classes, proficiency does not necessarily entail being literate in Frisian. In fact, many adult Frisian native speakers are highly proficient, but do not use the written form of Frisian. This chapter uses three measures of Frisian proficiency: perceived fluency, utterance fluency, and vocabulary. Native speaker ratings were used to test perceived fluency. Utterance fluency includes
acoustic measurements of the speech sample, including the total speaking time, number of pauses, and the mean pause duration. Overall vocabulary and Frisian vocabulary were investigated through an analysis of the total number of words, types of words, and the ratio of types to tokens (TTR). Together the perceived fluency, utterance fluency, and vocabulary analyses provide an indication of whether participants are proficient speakers of Frisian.

I predict that participants with more Frisian input will demonstrate overall higher language proficiency and that this will be reflected in higher perceived fluency, utterance fluency, and larger Frisian vocabulary. This chapter is organized as follows. I continue in Section 6.1 with an overview of the relevant literature; then I present the participants, materials, and procedure of the story-telling task. Section 6.2 presents the results of the native speaker ratings. Sections 6.3 and 6.4 present the results of the utterance fluency analysis and vocabulary analysis respectively. Sections 6.5 and 6.6 discuss the results and summarize the chapter.

6.1.1 Background

Within minority language communities, the goal of language revitalization is often general proficiency. That is, the goal is that children learn the language generally, rather than focusing on specific grammatical structures. Previous studies have investigated broad language use as a proxy for general proficiency. For example, in a large survey of minority language transmission, de Houwer (2007) asked immigrant parents in Belgium whether their children “use” the home language or only speak the majority community language, without reference to the acquisition of specific features.
Similarly, Nolan (2008) asked students and parents to report on where they “learned” Gallo, a language spoken in Brittany, but he did not ask for further details about their proficiency in specific areas. Specifically turning to North Frisian, Steensen (2003) asked students about their use of North Frisian, and the contexts in which they use Frisian, rather than focusing on specific language skills.

Similar opinions are common among language teachers on Föhr. Elementary school teachers report that their goal in Frisian class is to encourage students to use the language and focus on general language proficiency, rather than teaching or testing specific grammatical forms, such as grammatical gender or plural markings. Taken together, general language proficiency has been the focus of minority language researchers and teachers alike. With this in mind, this chapter uses a story-telling task to investigate children’s general Frisian proficiency. The goal of this chapter is to test whether the amount of Frisian children hear at home influences their general Frisian proficiency.

6.1.2 Participants

Forty-six elementary school students in grades one through four completed the story-telling task. All participants also completed the grammatical gender tasks (See Chapter 4) and the plural marking tasks (See Chapter 5). Elementary school students completed the story-telling task individually with the primary investigator and a native speaking research assistant during school hours, most commonly during Frisian classes.

Eight elementary school participants were unable to complete the task in Frisian and spoke exclusively in German. These participants were removed from the
acoustically-based utterance and vocabulary analyses to separate Frisian proficiency from German proficiency. However, they remained in the native speaker rating analyses in order to reflect the groups as a whole. The native speaker ratings were used to determine whether the student told the story in Frisian or German. That is, all of the participants who were rated as zero by the native speakers were removed from the analysis. One of these eight stories contained three Frisian words, but this was still considered “exclusively German” by the native speaker raters, and was therefore excluded from the fluency analysis.

Of the 38 elementary school students who completed this task in Frisian, 21 were high-input students, nine were middle-input students, and eight were minimal-input students (See Chapters 4 and 5 for a full description of the methods for classifying participants). The results from the nine high-input high school students and 13 adult native speakers and was used to test for the effect of cumulative language exposure. The assumption is that because all of the high school and adult participants come from high-input homes, adults will have had cumulatively the most Frisian input over their lifetimes, followed by the high school students, and then the high-input elementary school students. Throughout the chapter, the influence of quantity of input will be tested through comparisons of these groups. For the perceived fluency analysis, the elementary school students from three home language groups were compared. For the utterance fluency and vocabulary analysis high-input bilinguals from three different age groups were compared in addition to the comparison of the three groups of elementary school students.
6.1.3 Materials and Procedure

General proficiency was tested through a story-telling task where participants were asked to tell a story about a cartoon line-drawing consisting of multiple panels, which was created for this task. The cartoon drawing can be seen in Figure 6-1. Participants were told that the title of the story was “The Living Book”, and the story was designed to be funny and enjoyable for students ages six through ten. The instructions for the task are presented in Appendix D.

Figure 6-1. Pictures for Story-telling Task

Participants were shown the pictures and had a chance to look them over before telling a story in Frisian. In order to not conflate vocabulary and utterance fluency, all participants were told that they could ask the research assistant for specific vocabulary items. Participants were told that they would tell the story twice: once as practice and
once for the recording. With the elementary school students it was sometimes necessary to go through the pictures with the students prior to telling the story the first time. In accordance with what they were told, only the second version of the story was used for analysis. All participants completed the story-telling task as the third task in the experiment as a whole. For elementary school students it was the last task in session one, and for high school and adult participants it was the third of six tasks completed in one session. Participants were presented with a paper version of the cartoon, and recorded using Audacity 2.01 and an external microphone.

The cartoon depicted a classroom scene, in order to ensure the content of the story was familiar to all students. Previous research has shown that bilinguals’ vocabulary is influenced by the context in which they learn the language (Bialystok et al., 2010). All participants learn Frisian in school, but only the high-input and middle-input students learn Frisian at home. Thus, a story which included non-school vocabulary may have put the minimal-input students at a disadvantage relative to the other groups.

6.2 Perceived Fluency

6.2.1 Background

The perceived fluency analysis tests whether elementary school students with different amounts of Frisian input are rated as more or less proficient by native speakers. The prediction is that students with more Frisian input will be rated as more proficient than lower-input students.
6.2.2 Analysis

A native speaker listened to audio recordings of the elementary school students’ stories and rated each child’s Frisian proficiency. About half \( N = 26, 58\% \) of the stories were also rated by a second native speaker, and the inter-rater reliability was high: 96% (Crombach’s \( \alpha = .963 \)). The high schools students’ and adult stories were not rated. The stories were presented to raters by participant number; consequently the order of home language of the participant was pseudo-randomized. The raters completed the ratings independently, but they were instructed to separate the ratings into several sessions to avoid fatigue. Stories were rated on a scale from 0-4, presented in Figure 6-2. Following Peter and Hirata-Edds (2008), the native speaker raters were instructed to give a “holistic rating” which reflected how well the student spoke Fering. The raters, who were unaware of the goal of the project, the students’ home language, and age, were asked to rate the students’ overall Fering proficiency, including the speaker’s utterance fluency, sentence and story complexity, and vocabulary size. Because only a sub-set of stories were rated by two native speakers, for consistency, the analyses presented here are the results of only the first rater’s scores.
0: The child was not able to complete the task in Frisian; Speaks exclusively German

1: The story is limited in scope and may be only several sentences long; child may be only able to use simple sentences or relies heavily on German vocabulary or structures

2: The story is more developed than a level 1 score, but the child speaks less fluently than a level 3 story. The story is halting and hesitates frequently.

3: The child speaks freely in Frisian, but hesitates sometimes or uses repetitive or limited vocabulary

4: The child speaks freely and confidently in Frisian, showing a wide range of vocabulary and sentence structures. Child may use some German vocabulary, but is not overly reliant on German borrowings.

Figure 6-2. Scale for native speaker ratings of story-telling task

6.2.3 Results

Overall, the results indicate that the elementary school students on Föhr were perceived as relatively proficient in Fering, and, as predicted, confirmed that the amount of Frisian spoken at home influenced students’ general Frisian proficiency.

Table 6-1 presents the average rating and ranges of scores for each Home Language Group. As predicted, the high-input students were rated as the most proficient, followed by the middle-input students and then minimal-input students. Students who
told stories in German were rated as zero on the scale, and were averaged into the group averages.\textsuperscript{16}

Table 6-1. Native speaker ratings of story-telling task by Home Language

<table>
<thead>
<tr>
<th>Home Language</th>
<th>Mean</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-input (N = 20)</td>
<td>3.50</td>
<td>0.95</td>
<td>0-4</td>
</tr>
<tr>
<td>Middle-input (N = 12)</td>
<td>2.50</td>
<td>1.38</td>
<td>0-4</td>
</tr>
<tr>
<td>Minimal-input (N = 13)</td>
<td>1.54</td>
<td>1.56</td>
<td>0-4</td>
</tr>
</tbody>
</table>

The high-input students were rated on average as 3.5, meaning that the students in this group spoke freely and confidently in Frisian, demonstrated a wide range of vocabulary and sentence structures, but may have hesitated or been slightly repetitive at times. On average, the middle-input students were rated by native speakers as 2.5, which means that the stories were more hesitant or halting than the high-input students, but many students were able to tell stories in Frisian. In contrast, the minimal-input students were rated on average as 1.5, which means their stories were limited to several sentences, drew heavily on German vocabulary, and participants hesitated or spoke in a halting manner.

To test for differences in perceived fluency across groups, a one-way between-subjects ANOVA was conducted using the native speaker rating as the dependent variable and Home Language (High-input vs. Middle-input vs. Minimal-input) as a

\textsuperscript{16} The results were similar when only students who completed the task in Frisian were included. The zero score ratings are presented here because these ratings reflect the groups’ performance as a whole.
between subjects variable. The analysis revealed a main effect of Home Language: $F(2, 44) = 9.611, p < .001$. Follow-up comparisons\textsuperscript{17} revealed that the high-input group was rated significantly higher than the minimal-input group ($p < .001$). However, there was no significant difference between the high-input group and middle-input group ($p = .089$) or between the middle-input group and the minimal-input group ($p = .152$). Thus, while the high-input group was considered by the native speaker to be more proficient overall in Frisian than the minimal-input group, there was not a significant difference between the high-input group and the middle-input group.

**Case Studies**

It is important to keep in mind that while I separated participants into groups based on the amount of Frisian spoken by their parents at home, there are a number of factors, such as the language spoken with friends and during extracurricular activities, which may influence participants’ Frisian proficiency. A comparison between two minimal-input students with drastically different language outcomes highlights the importance of such factors.

The first student, Lena\textsuperscript{18}, produced a level four story, meaning that she was rated as similar to many high-input students. Lena is in the fourth grade and neither of her parents speaks Frisian. She attends the rural elementary school in western Föhr, where about two-thirds of students speak Frisian at home, and she lives in a small village in the western part of the island, where Frisian is commonly spoken in the community. She reports playing with friends out of school for about 20 hours a week, where she speaks

\textsuperscript{17} All follow-up comparisons were computed using Tukey HSD.
\textsuperscript{18} Both of the names have been changed.
exclusively Frisian. While other speakers would usually speak German with her because they know she does not speak Frisian at home, she insists that they speak Frisian with her and responds in only Frisian, and seems highly motivated to speak Frisian.

The second student, Imke, produced a level zero story, as she was not able to complete the task in Frisian and spoke exclusively German. Imke is also a fourth grade student and neither of her parents speaks Frisian. She lives in a village in central Föhr and attends the rural elementary school in eastern Föhr, where about a third of students speak Frisian at home. Like Lena, she reports playing with friends for about 20 hours a week, but she estimates that she speaks Frisian rarely, only about 10% of the time, while playing with friends.

These two students have the same home language background and amount of Frisian instruction. However, at least three additional factors differentiate them. First, they live in different areas of the island and attend different elementary schools, both of which influence the level of community bilingualism. Secondly, while they spend the same amount of time playing with friends, they use different languages with their friends. Finally, Lena appears to be more highly motivated to learn Frisian. While the larger applicability of these case studies must remain tentative given the limited scope, the results suggest a network of factors influences minority language acquisition above and beyond whether Frisian is spoken at home.

6.2.4 Discussion and Summary

Language revitalization research often focuses on whether children become proficient speakers of the language, rather than focusing on the acquisition of specific
grammatical structures. The results of the native speaker ratings suggest that there is not a significant difference between the high-input students and the middle-input students, and are similar in terms of perceived fluency. This would suggest that students with Frisian input roughly half of the time at home are able to be generally fluent in a minority language. The results of the case studies provide preliminary evidence that even minimal-input students can become proficient speakers of the minority with enough input from additional sources, including other community members and friends.

While the native speaker ratings reflect holistic differences, it is unclear which factors influence the native speaker’s perception of general proficiency. Raters were trained to rate “how well the child speaks Frisian”, and they were told that they may take the speaker’s fluency, vocabulary, grammatical accuracy and grammatical complexity into account. However, both raters reported struggling with what it means to speak Frisian well. From the native speaker ratings it is difficult to determine which features influenced their perception of Frisian proficiency. Thus, in the following sections I investigate two features in detail: utterance fluency and vocabulary.

### 6.3 Utterance Fluency

#### 6.3.1 Background

Utterance fluency refers to how smoothly a speaker produces the language, and is measured through analyses, such as how quickly someone speaks or how often she hesitates (de Jong, Steinel, Florijn, Schoonen, Hulstijn, 2013). In contrast to perceived fluency, which focuses on how native speakers perceive the speaker, utterance fluency is a measurement of the speech sample itself. Utterance fluency can be further broken down
into three types: speed (e.g., syllables per second), breakdown (e.g., the number and duration of pauses), and repair (e.g., the number of repetitions).

In this section I analyze the story-telling tasks based on four measures of utterance fluency, focusing primarily on breakdown. The measurements are: total speaking time, total pausing time, number of pauses, and mean pause duration. The prediction is that speakers with higher input will produce longer stories with fewer pauses, and that their pauses will be shorter in duration. Specifically, I predict that within elementary school participants, high-input and middle-input students will produce more fluent stories than minimal-input students. Comparing participants with different cumulative language exposure, I predict that high school students and adults will produce more fluent stories than high-input elementary school students.

6.3.2 Analysis

Prior to analysis, all recorded stories were trimmed to include only the participants’ speech. If the research assistant was included in the recording, for example, encouraging the participant, this was removed. All recordings were analyzed using a Praat script used to mark the boundaries between speaking time and silence (de Jong & Wempe, 2009). All pauses longer than 300 ms were marked as silence. The threshold for marking pauses was set at -25 dB. The spectrograms were visually inspected by the primary investigator, and if needed, corrections were made by hand. The boundaries were adjusted to include unvoiced word final consonants, or especially quietly spoken sections that were not included as speaking time by the Praat script. Silences at the beginning and end of the story were not counted as pauses. One adult participant’s audio file contained
background noise and could not be coded by the Praat script, so the pausing time and speaking time were coded by hand. After the manual adjustments were made, a second Praat script was used to calculate the total time marked as speaking time, the total pausing time, and the number of pauses (Lennes, 2011). The mean pause duration was calculated by dividing the adjusted total pausing time by the adjusted number of pauses.

6.3.3 Results

To investigate how the amount of language input influences utterance fluency, participants with different amounts of language input were compared on four measures—total speaking time, total pausing time, the number of pauses, and the mean pause duration. First, within elementary school students, high-input, middle-input, and minimal-input students were compared. Second, to test cumulative exposure, high-input bilinguals in three age groups, elementary school, high school, and adults were compared. The results of each comparison are presented separately below.

Home Language

The descriptive statistics for the utterance fluency from the three groups of elementary school students with different amounts of Frisian input are presented in Table 6-2. The eight participants who told stories exclusively in German were not included in this analysis. This included one high-input student, two middle-input students, and five minimal-input students. Overall, the high-input students appear to tell longer stories than the other groups, but there is a wide range of story length in the high-input group.19

---

19 There is one outlier high-input participant whose story was almost five minutes long. The analyses were also conducted with this participant removed, but the general patterns of outcomes remained, suggesting
Table 6-2. Utterance fluency by Home Language

<table>
<thead>
<tr>
<th>Home Language</th>
<th>Speaking Time</th>
<th>Pause Time</th>
<th>Number of Pauses</th>
<th>Mean Pause Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(sec)</td>
<td>(sec)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High-input</td>
<td>51.78 (57.27)</td>
<td>25.50 (24.77)</td>
<td>26.52 (31.68)</td>
<td>1.03 (0.38)</td>
</tr>
<tr>
<td>N = 21</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle-input</td>
<td>31.34 (15.31)</td>
<td>15.17 (9.38)</td>
<td>17.56 (10.71)</td>
<td>0.97 (0.33)</td>
</tr>
<tr>
<td>N = 9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minimal-input</td>
<td>32.05 (11.72)</td>
<td>20.68 (14.68)</td>
<td>18.13 (8.46)</td>
<td>1.11 (0.36)</td>
</tr>
<tr>
<td>N = 8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To test for differences in utterance fluency across groups with different amounts of Frisian input at home, four one-way between-subjects ANOVAs were conducted with the total speaking time, total pause time, number of pauses, and the mean pause duration as dependent variables, with Home Language as a between-subjects variable.20

The analysis revealed no significant differences based on Home Language:
Speaking Time: $F(2, 37) < 1$; Pause Time: $F(2, 37) < 1$; Number of Pauses: $F(2, 37) < 1$; Mean Pause Duration: $F(2, 37) < 1$. Therefore, in terms of overall utterance fluency, the elementary school students told similar stories regardless of the amount of Frisian input they receive at home. However, it is important to keep in mind that these results include

---

20 All participants who told stories in Frisian were included in this analysis. Once the participants who told the stories in German were removed, the overall small number of participants meant that matching participants would have resulted in a substantial loss of statistical power.
both Frisian words and German borrowings and code-switches. While the utterance fluency results suggest similar overall utterance fluency, they may not indicate similar levels of Frisian proficiency.

Age group

To test the effect of cumulative language input, a comparison of high-input bilinguals in three age groups was conducted. High-input elementary school students, high-input high school students, and adult native speakers were compared on four fluency measures: total speaking time, total pausing time, number of pauses, and mean pause duration.

Table 6-3. Utterance fluency by Age Group

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Speaking Time (sec)</th>
<th>Pause Time (sec)</th>
<th>Number of Pauses</th>
<th>Mean Pause Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M(SD)</td>
<td>M(SD)</td>
<td>M(SD)</td>
<td>M(SD)</td>
</tr>
<tr>
<td>Adult</td>
<td>35.99 (25.72)</td>
<td>17.10 (12.54)</td>
<td>19.67 (16.81)</td>
<td>0.93 (0.26)</td>
</tr>
<tr>
<td>N = 12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High School</td>
<td>30.26 (12.92)</td>
<td>12.92 (15.18)</td>
<td>14.11 (10.75)</td>
<td>0.76 (0.24)</td>
</tr>
<tr>
<td>N = 9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High-input Elementary</td>
<td>51.78 (57.27)</td>
<td>25.50 (24.77)</td>
<td>26.52 (31.68)</td>
<td>1.03 (0.38)</td>
</tr>
<tr>
<td>N = 21</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
As with the elementary school students, one-way between-subjects ANOVAs were conducted for each utterance fluency measure to test the effect of Age Group on utterance fluency.

The analysis revealed no significant differences between Age Groups based on any utterance fluency measure: Speaking Time: $F(2, 41) < 1$; Pause Time: $F(2, 41) = 1.325, p = .277$; Number of Pauses: $F(2, 41) < 1$; Mean Pause Duration: $F(2, 41) = 2.064, p = .141$. Thus, the high-input bilinguals in elementary school, high school, and adults produced similar stories in terms of fluency. Across all three age groups, participants told stories with similar total speaking time and total pausing time, and they did not differ in the numbers of pauses or mean pause duration.

6.3.4 Discussion and Summary

The utterance fluency analysis showed no group differences between elementary school student students with different levels of input, nor across different age groups. It is important to remember that this analysis only reflects the stories by students who completed the task in Frisian. Students were encouraged to tell the story in Frisian, but told they were told that they could tell the story in German, if they were not able to do so in Frisian. One high-input student, two middle-input, and five minimal-input students were unable to tell the story in Frisian, and told the story exclusively in German. The higher number of middle and minimal-input students who were not able to complete the task in Frisian suggests that fewer of these students were highly proficient. Thus, these results may reflect that there are no differences between higher proficiency students regardless of home language group. This result should be interpreted with caution.
because the lack of a difference in this sub-set may not reflect the differences in the groups as a whole, since they were excluded from this analysis because they were not able to tell stories in Frisian.

Further, while the participants who spoke only German were removed, participants who mixed Frisian and German were still included in the analysis. Therefore, the results of the utterance fluency analysis may reflect participants’ overall fluency, rather than Frisian fluency per se. For example, a participant who relied heavily on borrowings and code-switches from German and a participant who spoke only Frisian may have similar utterance fluency, but this does not mean that they have similar Frisian proficiency.

These results suggest that Frisian proficiency should be separated from overall fluency. It is difficult to isolate Frisian fluency from overall fluency, as measured by acoustical analyses of utterance fluency. It would be difficult, for example, to measure only the speaking time for Frisian words or pauses surrounding Frisian words. However, it is possible to separate Frisian and German fluency by considering vocabulary, as will be pursued in Section 6.4.

6.4 Vocabulary

6.4.1 Background

While the amount of Frisian language input did not significantly affect participants’ overall utterance fluency, as analyzed in the previous section, it is possible that the group differences that were reflected in the native speaker ratings are rooted in differences in vocabulary across the different participant groups. To test for differences in
vocabulary across participants with different amounts of Frisian input, the story-telling task results were analyzed for the overall number of different words (Types), the total number of words (Tokens), and the Type-Token Ration (TTR). The vocabulary analysis additionally allows for isolating a speaker’s Frisian vocabulary from her German vocabulary, which was not possible in the utterance fluency analysis. In the analyses presented below, I first present the results of all words, and then I compare speakers on only Frisian words.

Following the general prediction that more Frisian input will lead to higher Frisian proficiency, I predict that students with more Frisian input—high-input and middle-input elementary school students—will produce stories with a larger range of vocabulary than lower input elementary school students. Similarly, participants with more cumulative Frisian input—adults and high school students—will produce stories with a wider range of vocabulary than elementary school students, who have had cumulatively less Frisian input.

6.4.2 Analysis

All recordings were transcribed by a native speaking research assistant using a modified version of the CHAT transcription method (MacWhinney, 2000). The research assistant was unaware of the age or home language of the participants. The CLAN program was used to determine total number of words (Tokens), the number of different types of words (Types) and the ratio between Types and Tokens (TTR). Each different inflection of a base word was counted as a different word type. For example, singular and
plural forms were counted as different types (e.g., *buk* ‘book’ and *buken* ‘books’) were counted as two types.

First, all words were analyzed and then a second analysis was conducted on the subset of words produced in Frisian. For the Frisian vocabulary analysis, the primary investigator coded each word as either German or Frisian and all German words were removed from the analysis. The transcripts of the stories from the native Frisian speaker were used to determine which words were German and which were Frisian. All words that conformed to German orthography were considered German, and all others were considered Frisian. The few words that are spelled the same in Frisian and German, for example, *noch* ‘still’ were considered German words, and removed from the analysis. The Frisian vocabulary analysis tests whether speakers with different amounts of input differ in their Frisian vocabulary, rather than relying on German borrowed words.

6.4.3 Results

In order to test the effect of the amount of input on vocabulary, the number of types, tokens, and TTR was compared for groups of participants with different amounts of Frisian input. Within elementary school students, high-input, middle-input, and low-input students were compared. To test for the effect of cumulative language input, high-input bilinguals from three age groups—elementary school, high school, and adults—were compared. First I present the results of all words, and then I present the second analysis of only Frisian vocabulary.
**Home Language**

*All words*

The descriptive statistics for the vocabulary analysis including all words are presented in Table 6-4. These results show descriptively that high-input students appear to use more words and greater variety of vocabulary, but a similar TTR compared to the middle- and minimal-input students. As with the utterance fluency analysis, these averages reflect a high level of variability between speakers in the high-input group.²¹

Table 6-4. All Types, Tokens, and TTR by Home Language

<table>
<thead>
<tr>
<th>Home Language</th>
<th>Types $M(SD)$</th>
<th>Tokens $M(SD)$</th>
<th>TTR $M(SD)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-input $N = 21$</td>
<td>55.57 (22.90)</td>
<td>132.19 (138.92)</td>
<td>0.52 (0.13)</td>
</tr>
<tr>
<td>Middle-Input $N = 9$</td>
<td>39.89 (11.65)</td>
<td>83.67 (34.97)</td>
<td>0.52 (0.14)</td>
</tr>
<tr>
<td>Minimal-input $N = 8$</td>
<td>42.25 (13.59)</td>
<td>70.75 (32.95)</td>
<td>0.63 (0.11)</td>
</tr>
</tbody>
</table>

A series of one-way between-subjects ANOVAs was conducted for each of the three dependent variables: Types (total number of word types), Tokens (total number of words), and TTR (ratio between types and tokens) with Home Language group as the between-subjects variable. The analyses revealed a marginally significant effect of the number of different types of words: $F(2, 37) = 2.745, p = .078$. However, there was not a significant difference in the total number of words used by groups of participants with different amounts of Frisian exposure: $F(2, 37) = 1.246, p = .300$, or the TTR: $F(2, 37) =$

²¹ As with the utterance fluency analysis, both of the vocabulary analyses were conducted with the one outlier high-input student removed. The overall analysis is the same when this participant is removed. The results presented here include the outlier participant.
2.448, \( p = .101 \). While there was a trend toward differences in the number of different types, overall, when considering all words, it appears that elementary school students produced stories that have similar numbers of overall words, and ratios of types to tokens.

**Frisian Words**

The previous analysis included all words, including German words. As such, the results do not reflect the speakers’ Frisian vocabulary. Therefore, as a follow-up, all of the German words were removed and the number of Frisian types, Frisian tokens, and Frisian TTR were calculated. The descriptive results of the types, tokens, and TTR for only Frisian words are presented in Table 6-5.

Table 6-5. Frisian Types, Tokens, and TTR by Home Language

<table>
<thead>
<tr>
<th>Home Language</th>
<th>Types</th>
<th>Tokens</th>
<th>TTR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( M(SD) )</td>
<td>( M(SD) )</td>
<td>( M(SD) )</td>
</tr>
<tr>
<td>High-input ( N = 21 )</td>
<td>48.05 (20.15)</td>
<td>118.67 (126.29)</td>
<td>0.50 (0.14)</td>
</tr>
<tr>
<td>Middle-input ( N = 9 )</td>
<td>33.89 (10.35)</td>
<td>74.33 (31.58)</td>
<td>0.50 (0.14)</td>
</tr>
<tr>
<td>Minimal-input ( N = 8 )</td>
<td>32.13 (12.03)</td>
<td>58.50 (31.67)</td>
<td>0.59 (0.11)</td>
</tr>
</tbody>
</table>

Similar to the analysis with all words, a series of one-way between-subjects ANOVAs was conducted to test for differences among speakers with different amounts of Frisian input. Home Language (High-input vs. Middle-input vs. Minimal-input) was entered as the between-subjects variable. ANOVAs were conducted for each of the three dependent variables: Frisian Types, Frisian Tokens, and Frisian TTR. The analyses showed a
significant difference in the number of Frisian Types: $F(2, 37) = 3.722, p = .034$, but not in the number of Frisian Tokens: $F(2, 37) = 1.376, p = .266$ or Frisian TTR: $F(2, 37) = 1.396, p = .261$. As predicted, follow-up comparisons showed that the significant difference in Frisian types comes from marginally more Frisian types being produced by the high-input students than the minimal-input group ($p = .074$), but there is not a significant difference between high-input students and the middle-input students ($p = .104$), or between middle-input students and minimal-input students ($p = .975$). Thus, when considering only Frisian vocabulary, both the high-input and middle-input students use a similar range of Frisian vocabulary, but the minimal-input students lag behind their higher input peers.

**Age group**

**All Words**

A comparison of high-input bilinguals in three different age groups was conducted to investigate whether cumulative language input influences the participants’ vocabulary. The descriptive statistics for all high-input bilinguals in the three age groups are presented in Table 6-6.

Table 6-6. Types, Tokens, and TTR by Age Group

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Types $M(SD)$</th>
<th>Tokens $M(SD)$</th>
<th>TTR $M(SD)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adult $N = 12$</td>
<td>70.50 (24.67)</td>
<td>130.58 (80.84)</td>
<td>0.56 (0.19)</td>
</tr>
<tr>
<td>High School $N = 9$</td>
<td>56.22 (20.46)</td>
<td>110.33 (60.29)</td>
<td>0.56 (0.13)</td>
</tr>
<tr>
<td>High-input $N = 21$</td>
<td>55.57 (22.90)</td>
<td>132.19 (138.92)</td>
<td>0.52 (0.13)</td>
</tr>
</tbody>
</table>
To test for differences in vocabulary, a series of one-way between subjects ANOVAs was conducted. Age Group (Elementary School vs. High School vs. Adult) was entered as the between-subjects variable. ANOVAs were conducted for each of three dependent variables: Types, Tokens, and TTR. The analyses reveal no significant effects: Types-\(F(2, 41) = 1.770, p = .184\); Tokens- \(F < 1\); TTR- \(F < 1\). That is, across all age groups, high-input bilinguals produce stories with similar numbers of overall total word types, tokens, and TTR.

**Frisian Words**

As with the Home Language comparison, a follow-up analysis of only the Frisian words was conducted in order to separate the acquisition of overall vocabulary from Frisian vocabulary. The descriptive results are presented in Table 6-7.

### Table 6-7. Only Frisian Types, Tokens, and TTR by Age Group

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Types (M(SD))</th>
<th>Tokens (M(SD))</th>
<th>TTR (M(SD))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adult (N = 12)</td>
<td>64.58 (23.05)</td>
<td>122.92 (78.07)</td>
<td>0.59 (0.11)</td>
</tr>
<tr>
<td>High-input (N = 21)</td>
<td>48.05 (20.15)</td>
<td>118.67 (126.29)</td>
<td>0.50 (0.14)</td>
</tr>
<tr>
<td>High School (N = 9)</td>
<td>53.00 (20.55)</td>
<td>103.67 (61.09)</td>
<td>0.56 (0.12)</td>
</tr>
</tbody>
</table>

A series of three one-way between subjects ANOVAs for each of three vocabulary variables—Frisian Types, Frisian Tokens, and Frisian TTR—was conducted. The Age Group (Elementary school vs. High School vs. Adult) was entered as the between subjects variable. Similar to the analyses with all words, the analyses revealed
no significant differences between speakers from different age groups: Frisian Types-$F(2, 41) = 2.357, p = .108$; Frisian Tokens- $F < 1$; Frisian TTR- $F < 1$. This suggests that high-input bilinguals from all age groups produced stories with similar Frisian vocabulary, seen in the number of Frisian words, Frisian types and Frisian TTR.

To summarize the vocabulary analyses: The three groups of elementary school students produced stories that were similar in the overall number of words, and TTR, but the minimal-input students used fewer Frisian word types than the middle-input and high-input students. The high-input participants from the three age groups produced stories that were similar with regard to vocabulary use, regardless of whether all words were compared or only Frisian words, and regardless of the specific measure of vocabulary use investigated. Taken together, these results suggest that the high-input elementary school students produced similar stories to both adults and high school students.

6.4.4 Discussion and Summary

The results of the vocabulary analysis suggest that it is important to separate the analysis of Frisian proficiency from overall fluency. When considering only Frisian words, the high-input students and middle-input students produced more Frisian word types, but the minimal-input students lagged behind their higher-input peers in Frisian vocabulary. The comparison between high-input bilinguals in three age groups (Elementary vs. High school vs. Adult) also showed similar vocabulary sizes across groups. That is, for the analysis of only Frisian words, the high-input elementary school students, high school students, and adults produced stories which were similar in terms of the number of different words they used, the total number of words, and the TTR.
These vocabulary analyses address the broader research question of whether the amount of Frisian input, both currently and cumulatively over time, influences Frisian proficiency. As predicted, the participants with more Frisian input, both the high-input and middle-input participants, told stories with more diverse Frisian vocabulary than minimal-input students. Together with the higher native speaker ratings, the vocabulary analysis points to higher Frisian proficiency among participants with more Frisian input compared to those with less Frisian input.

There are several considerations, however, which may have contributed to the similarity between elementary school students. I will discuss three of these factors here: self-selection of participants, the story content, and the task methodology. First, as with utterance fluency, the vocabulary analyses only reflect the results of participants who were able to complete the task in Frisian. Thus, the results may show the vocabulary similarities among a subset of more proficient participants, rather than reflecting the Frisian proficiency of the high- versus middle- versus minimal-input students as a whole.

Secondly, the story was designed to include somewhat repetitive vocabulary, and several items, including pencil, book, paper, and student, occurred several times throughout the cartoon. This resulted in a story-telling task that was overall relatively undemanding. Thus, the story may not have been able to effectively distinguish high-input students from middle-input students, or the adults from the high-input elementary school students. While the high-input students may be able to produce more complex stories if they were given a more demanding task, it is possible that the middle-input

---

22 This allows for testing intra-speaker regularization, but the stories were not analyzed for regularization, due to the low numbers of definite articles in the stories.
students could not. Similarly, I would expect that the adults and high-input high school students would be able to produce more complex stories than high-input elementary school students, but the current story-telling task was not able to differentiate these groups. In the future, researchers should use a more complex story-telling task to more effectively test for differences between high-input participants.

Moreover, school vocabulary was selected in order to ensure that students who only learn Frisian in a school context could participate, as previous studies have revealed vocabulary differences based on the context in which children learn their languages (Bialystok et al., 2010). It is possible that the lower-input students have mastered school vocabulary, and while the high-input bilinguals have additional home-language vocabulary, middle or minimal-input bilinguals may not. The vocabulary accuracy analyses from the grammatical gender and plural marking production tasks, which tested a wider range of vocabulary, also point to differences in vocabulary across home language groups.

Despite these limitations, the results suggest that students who hear Frisian at home, both the high-input and middle-input students, use a larger range of Frisian vocabulary than the minimal-input students. The Frisian vocabulary results parallel the perceived fluency results in showing that children who hear Frisian at home, even roughly half of the time, are able to become fluent in Frisian.

It is encouraging that middle-input students are able to produce stories similar in vocabulary to their high-input peers, even though they differed in terms of their success in acquiring specific features, such as grammatical gender. The results underscore the importance for language communities to define exactly what their goals are, for example
language revitalization, as different features are differentially influenced by the amount of language input. For example, if the goal is grammatical accuracy on complex or inconsistent features, such as grammatical gender in Fering, substantial language input will be required. However, if the goal is general proficiency, perhaps less input is still sufficient.

6.5 General Discussion

This chapter investigated how the amount of Frisian input influence speakers’ overall Frisian proficiency. Proficiency was tested through three analyses: perceived fluency, utterance fluency, and vocabulary fluency. This section summarizes the results of each analysis, and then discusses how the results relate to the broader research question.

Native speaker ratings were used to test for differences in perceived fluency across elementary school students with different amounts of input. The results show that, as predicted, the participants with more Frisian input (including both the high-input and middle-input students) were perceived as more fluent than the minimal-input students. In terms of utterance fluency, the stories were analyzed for four features: total speaking time, total pausing time, the number of pauses, and the mean pause duration. There were no significant differences across groups according to Home Language for any of these measures, suggesting that at least the participants who were able to complete the story-telling task in Frisian, produced similarly fluent stories. Moreover, high-input elementary school students are similarly fluent to high-input high school students and adult native
speakers. However, this analysis includes both Frisian words and German words, and therefore reflects overall fluency, rather than specifically Frisian proficiency.

The Frisian vocabulary analysis attempts to separate Frisian proficiency from German proficiency. Paralleling the native speaker ratings, the vocabulary analysis showed that when considering only Frisian vocabulary, the higher-input students (both high-input and middle-input) produced stories with more diverse vocabulary than minimal-input students. However, I suggest that the story-telling task was not complex enough to fully differentiate between groups of higher-input bilinguals.

In terms of how the amount of input influences participants’ Frisian proficiency, the results suggest that students who hear Frisian at home, either from high-input homes or middle-input homes, are more proficient in Frisian than students who hear minimal amounts of Frisian at home. This was seen in the perceived fluency, as rated by an adult native speaker of Frisian, where the high-input students were rated as significantly more fluent than the minimal-input students. Conversely, the high-input students and middle-input students did not differ significantly from each other. This pattern held true for the vocabulary analysis when considering only Frisian words. The vocabulary analysis showed that high-input students used significantly more Frisian word types than minimal-input students. The high-input and middle-input elementary school students produced stories that were similar, which suggests that, for this type of story-telling task, the high-input and middle-input elementary school students have similar proficiency.

One of the overall limitations of the task is the simple content of the story. It is possible that the story was not complicated enough to reveal differences, especially between the high-input and middle-input elementary school students and between the
high-input elementary school students and the high school and adult participants. Even adults told fairly short stories (on average around 45 seconds total), suggesting that the task may not have challenged speakers enough to see differences between higher-input speakers. In the future, studies on minority languages should include several proficiency tasks, including stories with a range of topics and varying complexity, could be used together to determine participants’ highest level of proficiency.

Importantly, in terms of minority language acquisition, the results of this task show that even when students struggle with certain grammatical forms, such as grammatical gender, this limitation does not imply overall low Frisian proficiency. While the high-input and middle-input students were similarly proficient on the story-telling task, they differed in their production of grammatical gender. Recall that the definite article elicitation task in Chapter 4 showed that the high-input students produced adult-like inconsistency, while the middle-input students regularized grammatical gender. Moreover, eight minimal-input students felt confident enough to complete the task in Frisian and produce stories which were similar to peers from homes where exclusively Frisian is spoken. Taken together with the results from Chapters 4 and 5, the results suggest that overall proficiency and grammatical accuracy are differentially affected by the amount of input. Thus, it is important for language communities to clearly articulate their goals when designing minority language revitalization programs.

These results suggest that language planners, teachers, and community members should be careful in assuming that students’ regularization implies lower overall proficiency. Moreover, as with the differences between acquisition of vocabulary and the acquisition of morphology seen in Chapters 4 and 5, the results of the story-telling task
demonstrate the importance of defining the goals of language revitalization. While Chapter 4 showed substantially different outcomes as a function of home language input for inconsistent morphology, the general proficiency results suggest that students from both high-input and middle-input homes are similarly proficient in Frisian.

6.6 Summary

A story-telling task was used to test for differences in general Frisian proficiency among participants with different amounts of Frisian input. The amount of Frisian input was tested through a comparison of elementary school students from different home language groups and high-input bilinguals from three age groups. When isolating Frisian vocabulary from German vocabulary, the overall finding was that among participants who hear Frisian at home, both high-input and middle-input show diverse Frisian vocabulary, while minimal-input students as a group lag behind their higher-input peers. Some minimal-input students were able to complete the task in Frisian, suggesting that they, too, are proficient in Frisian. Moreover, the results suggest that the regularization seen in the grammatical gender production task does not reflect overall low Frisian proficiency. While the middle-input and high-input students differ in the rate of grammatical gender regularization, they are rated similarly by a native speaker on overall proficiency, and show a similar range of Frisian vocabulary.

Taken together, the results of the story-telling task suggest that some elementary school students on Föhr are proficient in Frisian, and students with more Frisian input are more likely to be proficient Frisian speakers than those with less Frisian input. Additionally, the results of the case studies suggest that beyond home language input a
network of other factors, including the student’s motivation and the language spoken with friends and in extracurricular activities, contribute to students’ Frisian proficiency.
Chapter 7. General Discussion

7.1 Goals of the Dissertation

The purpose of this dissertation was to examine how the type and quantity of linguistic input influences child language acquisition. Specifically, it asked whether the amount of input that speakers receive influences how they acquire inconsistent and consistent morphological markings, and whether children regularize inconsistently produced morphology. In order to answer these questions, this dissertation compared the acquisition of inconsistently produced grammatical gender in Fering to consistently produced plural markings. Bilinguals with different amounts of Frisian exposure were compared to investigate the influence of the quantity of language input.

Overall, it was predicted that speakers with low levels of input would regularize inconsistent input (based on Hudson Kam & Newport 2005, 2009), but that speakers with more input might be able to overcome regularization and faithfully acquire the adult-like pattern of inconsistency. With respect to consistently produced plural markings, it was predicted that participants with more Frisian input would be more accurate in plural markings than those with less Frisian input, but that participants would be less likely to regularize plural markings.

7.2 Summary of results

In order to examine how the quantity of input influences the acquisition of morphology, three features were investigated: grammatical gender, which is produced inconsistently, plural marking, which is produced consistently, and general proficiency.
Grammatical gender acquisition was tested using three tasks: a definite article elicitation task, a gender-marked pronoun elicitation task, and a gender-marked pronoun comprehension task. Plural marking was tested using a plural marking elicitation task and an acceptability task. A story-telling task was used to test for general Frisian proficiency.

Throughout the dissertation, the amount of input was examined through two comparisons: a comparison of elementary school students with different amounts of Frisian exposure at home, and a comparison of high-input bilinguals from three age groups. The age group comparison reflects differences in the amount of cumulative language input.

7.2.1 Grammatical Gender

When controlling for differences in vocabulary accuracy across groups, the definite article task showed that the amount of Frisian input significantly influenced how participants produced grammatical gender markings, with participants with less Frisian input more likely to rely on masculine gender. Participants showed three production patterns: random inconsistency, adult-like inconsistency, and regularization. The high-school students showed the most adult-like pattern, with a strong correlation with adults between the percentage of masculine responses on an item-by-item basis. The high-input elementary school students also showed adult-like inconsistency. The middle-input students tended to regularize grammatical gender markings, while the minimal-input students produced randomly inconsistent grammatical gender markings. Taken together, the definite article elicitation task shows that while some elementary school students
regularized inconsistently produced grammatical gender markings, high-input students were able to faithfully reproduce adult-like patterns of inconsistency.

The pronoun production task shows that, regardless of the amount of Frisian input, participants do not use gender marked pronouns often, and when they do, they are not very accurate on gender agreement. One of the most striking results of the pronoun production task was the low rate of gender-marked pronouns across all participant groups. While some participants almost always provided a gender marked pronoun, some rarely did. The amount of Frisian input did not appear to influence whether participants used gender-marked pronouns or not. Roughly half of elementary school students never produced a gender-marked pronoun, and they were evenly distributed across home language groups. When only examining the gender-marked pronoun responses, all three groups of elementary school students were equally low in accuracy, and an age group comparison showed that there was not a significant difference across groups with different amounts of cumulative input. In terms of the types of responses, all groups favored det, which is both the common gender and the demonstrative pronoun, and is not marked for grammatical gender.

7.2.2 Plural Marking

Overall, the results from the plural tasks showed that across both the home language comparison and the age group comparison, the participants with more Frisian input were more accurate in producing and comprehending plural markings. On the production task, the high-input students were significantly more accurate than the minimal input students, but the difference between the high-input and middle-input
students was not significant, although the differences trended toward higher accuracy among the high-input students. The age group comparison showed that the high school students and adults were at ceiling in terms of plural accuracy, and the high-input elementary school students, who have cumulatively less Frisian input, were less accurate. Participants with less input tended to produce –en plural markings more often than higher-input students. This tendency, however, did not qualify as regularization, according to the strict definition used in Hudson Kam and Newport (2005, 2009). On the acceptability task, the difference score analysis showed that both in the home language comparison and the age groups comparison, participants with more Frisian input made greater distinctions between correct and incorrect versions of the plural markings. Taken together, as predicted, in production and in comprehension, participants with more Frisian input were more adult-like in plural markings and did not produce the same regularized pattern as in the grammatical gender production task.

7.2.3 General Proficiency

In terms of general proficiency, the groups of bilinguals did not differ in their utterance fluency, as measured by total speaking time, total pausing time, the number of pauses, and the mean pause duration. However, the Frisian vocabulary analysis and the perceived fluency analysis, as measured by native speaker ratings, showed that both the high-input and middle-input bilinguals were similarly proficient, but that the minimal input students lagged behind those students who hear Frisian at home.
7.3 Conclusions

7.3.1 Language Acquisition

The results of this dissertation provide insight about how the amount and type of language input influences child language acquisition. These experiments were not designed to test models of child language acquisition, but the results are consistent with both generativist and usage-based models of language acquisition.

Some of the findings are consistent with generative approaches to language acquisition, such as the Variational Model (Yang, 1999; 2002). The Variational Model argues that children use linguistic input to determine which features from Universal Grammar are present in their language through a process of altering the probabilities associated with each prospective grammar from UG. Each grammar is associated with a probability that it is consistent with the adult target grammar. As the child encounters linguistic input, if the input is consistent with the prospective grammar, it is rewarded and the probability that the grammar represents the adult grammar increases. If the input is inconsistent with the prospective grammar, it is punished and the probability that the grammar represents the adult grammar decreases. The Variational Model argues that when children determine whether their language is consistent with a certain grammar, this parameter setting should affect all items equally. Consequently, the Variational Model predicts that children should initially regularize their input.

In line with this prediction, the results of the middle-input students are consistent with the Variational Model, in that the middle-input students regularized inconsistent gender markings. However, the Variational Model is less useful in explaining how the
high-input children successfully acquire the lexically-determined adult-like pattern of inconsistency.

Further, the Variational Model predicts that children would show protracted acquisition if the input they receive is unreliable. That is, if children receive inconsistent input, they may require more input in order to determine which features their language encodes. The current findings are in line with this prediction of the Variational Model. For inconsistently-marked features, such as grammatical gender, only the students with high-input were able to successfully acquire the system in elementary school, while the middle-input and minimal-input students lagged behind. In contrast, when children received consistent input, as with plural markings, there was not a significant difference between children with high-input and middle-input.

One of the main limitations of the Variational Model for explaining this dissertation’s findings is that the variational model is primarily concerned with how children determine which features their language marks, and is not concerned with how children acquire individual words, as is required in a lexically determined system. The Variational Model is a powerful model for understanding how children determine which features her language marks and explaining how children do so even when they are faced with different types of input. The Variational Model focuses on the acquisition of features, and has yet to incorporate how children acquire item-by-item correspondences of which markings are associated with each noun. In other words, the Variational Model sets out to explain how children determine that their language marks grammatical gender, rather than how they determine that fask ‘fish’ is associated with masculine gender.
To explain these noun-to-marking correspondences I turn to usage-based models, which are concerned with these item-by-item patterns of acquisition. The results of the high-input and middle-input elementary school students are also consistent with usage-based approaches to child language acquisition. These usage-base approaches predict that children do not rely on innate features of Universal Grammar, but rather, acquire words in a piece-meal fashion, using skills of general cognition (Lieven & Tomasello, 2008; Paradis, Nicoladis, Crago, & Genesee, 2011; Tomasello, 2003, see also Bybee 2007, Goldberg, 2013). This approach can explain how the high-input children and high school students were able to determine which grammatical gender marking is associated with each noun. According to this model, children track the frequency of each noun-to-form correspondence and the frequency of each pairing determines how quickly it will be acquired. The high-input students were able to acquire the by-item, lexically determined pattern of gender marking, which suggests that their production can be best explained by usage-based approaches.

Interestingly, students did not regularize consistently produced plural markings, suggesting that regularization is tied to the input type. Although students’ production of plural markings did not meet the definition of regularization used in Hudson Kam and Newport (2005, 2009), there was a tendency for lower input students to over rely on one plural marking, namely, -en. This highlights the importance of considering the pattern of language use, beyond the narrow definition of regularization.

Perfors (2012) suggests that the regularization seen in Hudson Kam and Newport (2005, 2009) stems from how children approach artificial language learning tasks. Specifically, children do not feel the pressure to be accurate on the task and therefore
adopt a regularization strategy to best meet the task demands. In contrast, adults’ experience with language learning situations leads them to assume that there is an underlying systematic pattern, and they struggle to determine the pattern, which results in probability matching. This possibility seems an unlikely explanation for the current results. If the differences in how children and adults approach learning artificial languages carries over to how they learn natural languages, this theory may account for the differences in the age-group comparison, but it cannot account for differences in the home-language comparison. If age is main factor influencing how participants approach the task, the high-input, middle-input and minimal-input students would be expected to be equally affected by the task demands. Despite this, the elementary school students from different home language groups showed different outcomes.

Finally, the findings of this dissertation are particularly informative for research on the acquisition of minority languages. These findings provide converging evidence that the quantity and quality of language input interact in minority language contexts. The current results are in line with Gathercole and Thomas (2005, 2009) and Thomas et al., (2013), who argue that children in minority language contexts need a large quantity of input to acquire complex and opaque systems, such as Welsh grammatical gender, and Welsh plural marking. Overall, these studies indicate that children require a large amount of input to acquire minority languages, and that this is especially true for complex, opaque, or inconsistent features.
7.3.2 Language Change

Based on the current results, we can speculate how inconsistent input may influence language change. Specifically, the results suggest an on-going change in the gender agreement system in Fering, which is consistent with the predicted pattern of gender loss in Germanic languages (Trudgill, 2012; Wahring-Burfeind, 1989). More broadly, the results can inform the discussion of whether children drive language change through regularization.

This dissertation showed that when children have high levels of input they may be able to acquire lexically determined inconsistency, and this supports the previous literature that grammatical gender in Fering has been inconsistent for at least the last 100 years (Ebert, 1998). That is, children may be able to retain even a highly-inconsistent gender marking system. Further, a number of factors influence language change, and it is important to keep in mind that variation does not necessarily indicate language change (see Dorian, 1994). That being said, gender loss in Fering has been suggested and predicted in a number of previous studies, including Parker (1993), Ebert (1998), and Wahring-Burfeind (1989). The limited sample sizes in the previous studies make it difficult to determine whether gender loss has progressed over the last 30 years, but the current results are in line with the prediction that gender is being lost in Fering, similar to the ongoing process in Modern Dutch (De Vogelaer & De Sutter, 2011).

Even within the high-input elementary group, some students regularized grammatical gender. If these speakers transmit their regularized gender system to the next generation, their children will acquire a regularized gender system with only one grammatical gender. Complete gender loss occurs when the definite article system ceases
to distinguish between genders even for animate references. There is some evidence from the story-telling task that some students have begun this change. There were rare cases in which the default masculine article, *de*, was used for female animate nouns, such as teacher. If grammatical gender loss is on-going, I would expect that these rare cases would become more common in future generations.

I predict that, if gender loss is currently in progress, the speed at which this takes place will likely depend on the amount of input children receive. If more children receive higher levels of input, I would expect that inconsistency would persist longer than if fewer children receive high levels of Fering input. It is especially important to keep in mind that even if Fering loses grammatical gender, that this does not imply language death. The results of the plural marking task and the story-telling task suggest that many children are currently highly proficient in Fering, although they may have regularized grammatical gender markings. Gender syncretism is typical in Germanic languages and does not necessarily indicate language loss (Trudgill, 2012). Keep in mind that English has lost grammatical gender but there is no doubt that it continues to be a vital language.

The changes throughout the grammatical gender agreement system in Fering appear to follow the expected pattern of diachronic gender loss in Germanic languages (Wahring-Burfeind, 1989, see Trudgill, 2012 for a summary). The current system in Fering appears to be moving towards a system similar to modern Standard Swedish and modern Standard Danish, which distinguish between two grammatical genders in the definite article system– a neuter gender and a common gender– but three genders in the pronominal system. Masculine gender pronouns are used for male animates, feminine gender pronouns are used for female animates, and a new pronoun has emerged for
inanimates. In Danish and Swedish, this new pronoun for inanimate nouns originates from the historically masculine distal demonstrative: *den* ‘that’ (Trudgill, 2012).

Demonstratives being used as pronouns for inanimate references can also be seen in the use of *däi* ‘that’ in East Frisian Low German (Matras & Reerschemius, 2003). Similarly, the demonstrative, *det*, appears to be surfacing as a pronoun for inanimate references in Fering. It remains unclear whether speakers continue to use gender-marked pronouns for animate references.

While I did not directly test gender use for pronouns with animate nouns in this project, the results from several tasks suggest that Fering might be following this pattern. First, participants tended to over-use the demonstrative, *det* ‘that’, in the pronoun production task (see Chapter 4). Secondly, all groups of participants were at ceiling for the comprehension of gender-marked pronouns when the referents were animates, meaning that they have mastered pronouns with animates referents. This suggests that the change may be in the use of pronouns for inanimate references, while use of gender-marked pronouns for animate references has remained stable. Thirdly, some elementary school students use *det* to refer to inanimate references (e.g., an de dring skel nü <det> [//] en Arbeitsblatt maai, an hi fantj *det* ei so gud. ‘and the boy should do a worksheet now, and he does finds it not so good’), while continuing to use the gender-marked pronouns (hi ‘he’ and hat ‘she’) for animate references (e.g. diar as en dring. Hi sat bi a boosel ‘here is a boy. He sits at the table’). Unfortunately the story-telling has was not designed to elicit pronouns, and therefore there were not enough pronouns to be able to fully analyze these results. Together, it appears that Fering distinguishes between two grammatical genders—masculine and common gender—in the definite article system, but
which uses a three-way distinction in the pronominal system. This system appears to rely on natural gender and animacy, where animate males are marked with *hi* ‘he’, animate females are marked with *hat* ‘she’, and inanimates are marked with *det* ‘that’. This system would be similar to Danish, Swedish and East Frisian Low German.

Beyond grammatical gender, one interesting corollary to the potential loss of grammatical gender in Fering is the impact that this change could have on other areas of the morphological system. Historically, there has been a strong connection between plural markings and grammatical gender in Fering (Artsen & Tadsen, 2009). Masculine words were typically marked with *–er* plural marking, while common gender words were typically marked with *–en* plural marking. In this study, children favored the masculine gender as the default grammatical gender, but they favored the *–en* plural marking in the plural tasks, thus suggesting a decoupling of this previous connection between grammatical gender and plural marking in Fering.

Interestingly, the words that conform to the historical pattern tend to have higher-consensus for gender—masculine words are marked with *–er* plural marking and common gender are marked with *–en* plural marking—than those that do not conform to the historical pattern. In the items from the grammatical gender and plural marking tasks, the masculine words are still predominantly marked with *–er* plural marking and there were no words that have common gender but *–er* plural marking. When words deviate from the expected historical pattern, it is in the direction of masculine words being marked with *–en* plural marking, and these words tend to have lower consensus than the masculine words with *–er* marking. Taken together, these observations show that when there is
change in one morphological feature, such as gender syncretism, it may also have ramifications throughout the agreement system.

Beyond the Germanic context, these results can be used to test whether children drive language change over time. One of the purposes of studies by Hudson Kam and Newport (2005, 2009) was to simulate the learning environment that leads to the creation of creole languages. They argue that children regularize inconsistent input and drive language change by imposing additional systematicity to the inconsistent input they receive. In contrast, adults match the probability in their input and therefore are unlikely to drive language change. This dissertation finds partial support for these findings. More specifically, there is evidence that some children are able to faithfully acquire the adult-like pattern of inconsistency. From a diachronic perspective this means that it is possible for languages to maintain inconsistency, and provides further evidence of the extent and type of inconsistency that languages can maintain.

At the same time, some children did indeed regularize inconsistent input. While some students from each home language group regularized grammatical gender, regularization was most prevalent in the middle-input group. This suggests that children may drive language change through regularization, as Hudson Kam and Newport (2005, 2009) suggest. However, this occurs when children do not have enough input to be able to faithfully acquire the adult-like pattern. This may be especially true in cases of unbalanced bilingualism, such as in minority languages.

To summarize: the results from this dissertation are compatible with previous suggestions that Fering is in the process of grammatical gender loss (Ebert, 1994, Parker, 1993). The speed of this change will depend on the amount of input available to children.
Moreover, the change appears to follow the typical pattern of gender loss in Germanic languages and, further, changes to the grammatical gender system can lead to changes in other inflectional paradigms, like plural morphology. More generally, these results highlight that investigations of language change in progress can increase our understanding of how children influence language change, but also helps to separate language change from language death.

7.3.3 Application for Language Planning

The results of this dissertation also inform language planning. Much of the research on minority languages focuses on descriptions of the languages or describing patterns of intergenerational transmission. This project shows that experimental evaluations can add to these descriptive studies. Experimental evaluations of language outcomes can be useful in determining which specific features students struggle with, and where additional attention should be paid. Specifically, comparisons of groups of students with different amounts of Frisian input can help teachers and language planners to know how much minority language input is needed in order for the language to be successfully acquired. Without evaluations of language outcomes, there is no way to measure whether the community’s investment has been effective, or how the program may be altered to better meet the needs of the diverse group of students. As language planners and minority language teachers are well aware, there are rarely enough resources to be able to implement the program they ideally envision. Experimental studies on language outcomes can help teachers and language planners to focus on the
aspects of the language that require the most attention, and design more efficient
language programs, or better lobby for additional resources, if they are required.

One of the most important findings of this dissertation is that the acquisition of
specific morphological features, such as grammatical gender, should be seen as separate
from general proficiency. The results of the grammatical gender task, in conjunction with
the previous studies on Frisian grammatical gender (Ebert, 1998, Parker 1993), suggest
that Frisian may be in the process of losing grammatical gender. However, it is important
not to view this change as an indication of language death. Indeed, both of the groups of
elementary school students who hear Frisian at home perform well on the story-telling
task.

Finally, one other key finding was that different features may require different
amounts of language input to be successfully acquired. This means that communities
need to define their language revitalization and maintenance goals, as different features
will require different amounts of input. For example, if a community decides that their
goal is general proficiency, then the community can recommend that both high and
middle levels of input are sufficient for success. However, if the community focuses on
the acquisition of specific, complex morphological features, such as inconsistently
produced grammatical gender, or minority language vocabulary over borrowings, then
the recommendation will need to be that students need more input than middle-input
homes provide. The current research did not investigate whether this additional input
would need to be at home or whether classroom and community input would be equally
effective in encouraging language acquisition, but this would be an interesting avenue for
future research.
7.4 Future research

In this dissertation I showed that the quantity of input and the type of language input interact in language acquisition. I investigated the case of inconsistently produced grammatical gender marking in Fering and the consistently produced plural marking across participants with different amounts of input. The results suggest that while some children regularize inconsistently produced morphology, those with higher levels of input are able to faithfully re-produce the lexically determined adult-like pattern of inconsistency. However, a number of questions remain.

One of these unanswered questions is whether there is a complex underlying systematicity to the Fering grammatical gender system. While this investigation did not find a pattern of grammatical gender assignment, it is possible that the gender system includes cues to gender assignment, such as a mass/count distinction, or that alternations of gender are related to semantic differences, as Semplicini (2012) argues is the case with double-gender nouns in Dutch. To address this possibility, future research should include more free speech interviews, which would test grammatical gender in more varied contexts. If there is an underlying pattern, it is possible that the high-input children have acquired the pattern, meaning that grammatical gender acquisition does not reflect the acquisition of inconsistent input, but rather the acquisition of a complex, but potentially consistent system. However, even if this is the case, such a complex system may initially appear as inconsistent to children, and therefore the difference between a complex system and an inconsistent system may not significantly influence acquisition.
It is essential that future studies investigate the amount of input in more depth. This dissertation grouped elementary school students based on broad categories of input based on the language their parents spoke at home. Future studies should gather more specific information about the exact quantity of input. This study shows that the amount of language input must be considered in bilingual language acquisition, and future research should include more fine-grain measurements of the amount and source of such input.

This study provided an initial investigation into the type of input, but the exact nature of the input available to children should be investigated further in future studies. For example, I set aside the possibility that the type of input may have differed across groups of children. That is, it is possible that the lower input students differed not only in the quantity of language input but also in how often they encountered input from second language learners, or that the adults that provide the input for low- and middle-input children produce grammatical gender differently than those adults with whom the high-input students interact. For this study, more detailed information about the speech partners was collected for some students, but these data were not analyzed because they were not available for all students. A more thorough analysis of the language use questionnaires may be able to determine whether children’s speech partners were native or non-native speakers. Further, all of the parents of children in the study were asked to participate, but the connection between particular parents and children’s speech was not analyzed in this study. Future studies should focus more information about who children interact with and directly study the child directed speech patterns, as has been done in Miller (2013) and Smith et al. (2007, 2009).
Additionally, the source of language input should be taken into account. It remains unclear whether input from all sources is equally weighted in the total amount of input. This is especially important in combination with an investigation of the amount of input. If input from friends and school instruction can also be useful in language acquisition, then language planners can focus on increasing the amount of input in these areas. However, if future research shows that home language holds a privileged position in language acquisition, then this means that language planners should focus their attention on increasing language input inside the home. This could take the form of educating the community about the importance of language input at home, or adult education courses to increase the number of parents who feel confident speaking the minority language at home to their children. Studies on the source of language input would also deepen our understanding of the connection between input quantity and type.

Most generally, this dissertation shows that it is important to investigate minority languages, as they provide unique environments for testing theories of language acquisition and language change. The results of this study uncover only a fraction of the knowledge that endangered minority languages provide, and therefore future projects should consider minority languages for their ability to inform a diverse range of research questions in the field of linguistics.
## Appendix A. Results of adult online survey

Table A-1. Responses to adult survey for grammatical gender marking task

<table>
<thead>
<tr>
<th>Fering (dictionary gender)</th>
<th>English</th>
<th>German (gender)</th>
<th>Masculine (1 and 2)</th>
<th>Inter-changeable (3, 4, 5)</th>
<th>Common (6 and 7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>tuun (m)</td>
<td>fence</td>
<td>Zaun (m)</td>
<td>95%</td>
<td>0%</td>
<td>5%</td>
</tr>
<tr>
<td>strun (m)</td>
<td>beach</td>
<td>Strand (m)</td>
<td>93%</td>
<td>2%</td>
<td>4%</td>
</tr>
<tr>
<td>kai (m)</td>
<td>key</td>
<td>Schlüssel (m)</td>
<td>93%</td>
<td>0%</td>
<td>7%</td>
</tr>
<tr>
<td>boosel (m)</td>
<td>table</td>
<td>Tisch (m)</td>
<td>91%</td>
<td>4%</td>
<td>4%</td>
</tr>
<tr>
<td>dööbel (m)</td>
<td>dice</td>
<td>Würfel (m)</td>
<td>91%</td>
<td>0%</td>
<td>9%</td>
</tr>
<tr>
<td>tus (m)</td>
<td>tooth</td>
<td>Zahn (m)</td>
<td>91%</td>
<td>5%</td>
<td>5%</td>
</tr>
<tr>
<td>fask (m)</td>
<td>fish</td>
<td>Fisch (m)</td>
<td>89%</td>
<td>4%</td>
<td>7%</td>
</tr>
<tr>
<td>tusbasel (m)</td>
<td>toothbrush</td>
<td>Zahnbürste (f)</td>
<td>88%</td>
<td>5%</td>
<td>7%</td>
</tr>
<tr>
<td>spiker (m)</td>
<td>nail</td>
<td>Nagel (m)</td>
<td>87%</td>
<td>9%</td>
<td>4%</td>
</tr>
<tr>
<td>sedel (m)</td>
<td>kettle</td>
<td>Kessel (m)</td>
<td>86%</td>
<td>7%</td>
<td>7%</td>
</tr>
<tr>
<td>riam (m)</td>
<td>belt</td>
<td>Gürtel (m)</td>
<td>84%</td>
<td>7%</td>
<td>9%</td>
</tr>
<tr>
<td>basel (m)</td>
<td>brush</td>
<td>Bürste (f)</td>
<td>82%</td>
<td>7%</td>
<td>11%</td>
</tr>
<tr>
<td>skuch (m)</td>
<td>shoe</td>
<td>Schuh (m)</td>
<td>82%</td>
<td>4%</td>
<td>13%</td>
</tr>
<tr>
<td>skor (c)</td>
<td>skates</td>
<td>Schlittschuh (m)</td>
<td>75%</td>
<td>14%</td>
<td>11%</td>
</tr>
<tr>
<td>hole (m)</td>
<td>bull</td>
<td>Stier (m)</td>
<td>74%</td>
<td>12%</td>
<td>14%</td>
</tr>
<tr>
<td>hünjmots (m/c)</td>
<td>mushroom</td>
<td>Pilz (m)</td>
<td>73%</td>
<td>13%</td>
<td>13%</td>
</tr>
<tr>
<td>German</td>
<td>English</td>
<td>German</td>
<td>English</td>
<td>German</td>
<td>English</td>
</tr>
<tr>
<td>--------</td>
<td>---------</td>
<td>--------</td>
<td>---------</td>
<td>--------</td>
<td>---------</td>
</tr>
<tr>
<td>mös (m)</td>
<td>mouth</td>
<td>Mund (m)</td>
<td>73%</td>
<td>2%</td>
<td>25%</td>
</tr>
<tr>
<td>switter (m)</td>
<td>pullover</td>
<td>Pullover (m)</td>
<td>71%</td>
<td>24%</td>
<td>5%</td>
</tr>
<tr>
<td>skütel (m)</td>
<td>saucer</td>
<td>Untertasse (f)</td>
<td>71%</td>
<td>4%</td>
<td>24%</td>
</tr>
<tr>
<td>swaaelstook (m)</td>
<td>match</td>
<td>Streichholz (n)</td>
<td>71%</td>
<td>13%</td>
<td>16%</td>
</tr>
<tr>
<td>dask (m)</td>
<td>bowl</td>
<td>Schüssel (f)</td>
<td>69%</td>
<td>11%</td>
<td>20%</td>
</tr>
<tr>
<td>stääär (c)</td>
<td>star</td>
<td>Stern (m)</td>
<td>65%</td>
<td>19%</td>
<td>16%</td>
</tr>
<tr>
<td>brag (c)</td>
<td>bridge</td>
<td>Brücke (f)</td>
<td>61%</td>
<td>14%</td>
<td>25%</td>
</tr>
<tr>
<td>hunskreep (m)</td>
<td>purse</td>
<td>Handtasche (f)</td>
<td>60%</td>
<td>13%</td>
<td>27%</td>
</tr>
<tr>
<td>skreet (m)</td>
<td>pants pocket</td>
<td>Hosentasche (f)</td>
<td>58%</td>
<td>13%</td>
<td>29%</td>
</tr>
<tr>
<td>kastem (c)</td>
<td>TV</td>
<td>Fernseher (m)</td>
<td>56%</td>
<td>31%</td>
<td>13%</td>
</tr>
<tr>
<td>kwaast (m)</td>
<td>paintbrush</td>
<td>Pinsel (m)</td>
<td>56%</td>
<td>24%</td>
<td>20%</td>
</tr>
<tr>
<td>riap (m)</td>
<td>rope</td>
<td>Seil (n)</td>
<td>56%</td>
<td>22%</td>
<td>22%</td>
</tr>
<tr>
<td>kakinje (c)</td>
<td>candy</td>
<td>Bonbon (m)</td>
<td>51%</td>
<td>13%</td>
<td>36%</td>
</tr>
<tr>
<td>sees (m)</td>
<td>cheese</td>
<td>Käse (m)</td>
<td>49%</td>
<td>11%</td>
<td>40%</td>
</tr>
<tr>
<td>brek (m/c)</td>
<td>pants</td>
<td>Hose (f)</td>
<td>49%</td>
<td>9%</td>
<td>42%</td>
</tr>
<tr>
<td>laid (m/c)</td>
<td>lightening</td>
<td>Blitz (m)</td>
<td>47%</td>
<td>19%</td>
<td>35%</td>
</tr>
<tr>
<td>sink (c)</td>
<td>sink</td>
<td>Spülbecken (n)</td>
<td>43%</td>
<td>16%</td>
<td>41%</td>
</tr>
<tr>
<td>höös (c)</td>
<td>stocking</td>
<td>Strumpf (m)</td>
<td>40%</td>
<td>11%</td>
<td>49%</td>
</tr>
<tr>
<td>tau(c)</td>
<td>rope</td>
<td>Seil (n)</td>
<td>40%</td>
<td>25%</td>
<td>35%</td>
</tr>
<tr>
<td>blijant (m)</td>
<td>pencil</td>
<td>Bleistift (m)</td>
<td>36%</td>
<td>50%</td>
<td>14%</td>
</tr>
<tr>
<td>riw (c)</td>
<td>rake</td>
<td>Harke (f)</td>
<td>31%</td>
<td>7%</td>
<td>62%</td>
</tr>
<tr>
<td>jüg (c)</td>
<td>wing</td>
<td>Flügel (m)</td>
<td>31%</td>
<td>18%</td>
<td>51%</td>
</tr>
<tr>
<td>triiuk (m)</td>
<td>triangle</td>
<td>Dreieck (n)</td>
<td>29%</td>
<td>7%</td>
<td>64%</td>
</tr>
<tr>
<td>ruus (m/c)</td>
<td>flower</td>
<td>Blume (f)</td>
<td>29%</td>
<td>7%</td>
<td>64%</td>
</tr>
<tr>
<td>slöfk (c)</td>
<td>slipper</td>
<td>Hausschuh (m)</td>
<td>28%</td>
<td>50%</td>
<td>23%</td>
</tr>
<tr>
<td>reiluk (c)</td>
<td>curtain</td>
<td>Vorhang (m)</td>
<td>27%</td>
<td>11%</td>
<td>62%</td>
</tr>
<tr>
<td>Fering</td>
<td>English</td>
<td>German</td>
<td>Fering</td>
<td>German</td>
<td>–en</td>
</tr>
<tr>
<td>--------</td>
<td>--------------</td>
<td>-------------</td>
<td>--------</td>
<td>--------</td>
<td>-----</td>
</tr>
<tr>
<td>brasetae</td>
<td>fried egg</td>
<td>Spiegelei</td>
<td>-er</td>
<td>-er</td>
<td>33%</td>
</tr>
<tr>
<td>godeschenk</td>
<td>lady bug</td>
<td>Marienkäfer</td>
<td>-en</td>
<td>φ</td>
<td>51%</td>
</tr>
<tr>
<td>skor</td>
<td>skates</td>
<td>Schlittschuh</td>
<td>-en</td>
<td>-e</td>
<td>49%</td>
</tr>
<tr>
<td>tau</td>
<td>rope</td>
<td>Seil</td>
<td>-en</td>
<td>-e</td>
<td>49%</td>
</tr>
<tr>
<td>steulis</td>
<td>swing</td>
<td>Schaukel</td>
<td>-en</td>
<td>-n</td>
<td>67%</td>
</tr>
<tr>
<td>jiug</td>
<td>wing</td>
<td>Flügel</td>
<td>-en</td>
<td>φ</td>
<td>84%</td>
</tr>
<tr>
<td>skaab</td>
<td>closet</td>
<td>Schrank</td>
<td>-en</td>
<td>U-e</td>
<td>87%</td>
</tr>
</tbody>
</table>

Table A-2. Responses to adult survey for plural marking task
<table>
<thead>
<tr>
<th>bleech</th>
<th>sheet</th>
<th>Bettlaken</th>
<th>-en</th>
<th>ø</th>
<th>80%</th>
<th>7%</th>
<th>0%</th>
<th>7%</th>
<th>2%</th>
<th>4%</th>
</tr>
</thead>
<tbody>
<tr>
<td>triihuk</td>
<td>triangle</td>
<td>Dreieck</td>
<td>-er</td>
<td>-e</td>
<td>47%</td>
<td>47%</td>
<td>0%</td>
<td>0%</td>
<td>2%</td>
<td>4%</td>
</tr>
<tr>
<td>riiw</td>
<td>rake</td>
<td>Harke</td>
<td>-en</td>
<td>-e</td>
<td>67%</td>
<td>27%</td>
<td>0%</td>
<td>0%</td>
<td>2%</td>
<td>4%</td>
</tr>
<tr>
<td>skau</td>
<td>scarecrow</td>
<td>Vogelscheuche</td>
<td>-en</td>
<td>-e</td>
<td>47%</td>
<td>9%</td>
<td>0%</td>
<td>11%</td>
<td>31%</td>
<td>2%</td>
</tr>
<tr>
<td>slöfk</td>
<td>slipper</td>
<td>Hausschuhe</td>
<td>-en</td>
<td>-e</td>
<td>49%</td>
<td>20%</td>
<td>2%</td>
<td>2%</td>
<td>24%</td>
<td>2%</td>
</tr>
<tr>
<td>gaaw</td>
<td>present</td>
<td>Geschenk</td>
<td>-en</td>
<td>-e</td>
<td>78%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>20%</td>
<td>2%</td>
</tr>
<tr>
<td>sink</td>
<td>sink</td>
<td>Spülbecken</td>
<td>-en</td>
<td>ø</td>
<td>78%</td>
<td>9%</td>
<td>2%</td>
<td>0%</td>
<td>9%</td>
<td>2%</td>
</tr>
<tr>
<td>henk</td>
<td>chick</td>
<td>Küken</td>
<td>-en</td>
<td>ø</td>
<td>78%</td>
<td>7%</td>
<td>0%</td>
<td>7%</td>
<td>7%</td>
<td>2%</td>
</tr>
<tr>
<td>höös</td>
<td>stocking</td>
<td>Strumpf</td>
<td>-en</td>
<td>U-e</td>
<td>89%</td>
<td>0%</td>
<td>0%</td>
<td>4%</td>
<td>4%</td>
<td>2%</td>
</tr>
<tr>
<td>brag</td>
<td>bridge</td>
<td>Brücke</td>
<td>-en</td>
<td>-n</td>
<td>80%</td>
<td>13%</td>
<td>0%</td>
<td>0%</td>
<td>4%</td>
<td>2%</td>
</tr>
<tr>
<td>bleed</td>
<td>newspaper</td>
<td>Zeitung</td>
<td>-en</td>
<td>-n</td>
<td>89%</td>
<td>4%</td>
<td>0%</td>
<td>0%</td>
<td>4%</td>
<td>2%</td>
</tr>
<tr>
<td>kop</td>
<td>cup</td>
<td>Tasse</td>
<td>-en</td>
<td>-n</td>
<td>89%</td>
<td>4%</td>
<td>0%</td>
<td>0%</td>
<td>4%</td>
<td>2%</td>
</tr>
<tr>
<td>wel</td>
<td>bicycle</td>
<td>Fahrrad</td>
<td>-en</td>
<td>U-er</td>
<td>87%</td>
<td>2%</td>
<td>0%</td>
<td>7%</td>
<td>2%</td>
<td>2%</td>
</tr>
<tr>
<td>blijant</td>
<td>pencil</td>
<td>Bleistift</td>
<td>-er</td>
<td>-e</td>
<td>44%</td>
<td>27%</td>
<td>2%</td>
<td>4%</td>
<td>22%</td>
<td>0%</td>
</tr>
<tr>
<td>käärs</td>
<td>cherry</td>
<td>Kirsche</td>
<td>-en</td>
<td>-n</td>
<td>76%</td>
<td>7%</td>
<td>4%</td>
<td>0%</td>
<td>13%</td>
<td>0%</td>
</tr>
<tr>
<td>stääär</td>
<td>star</td>
<td>Stern</td>
<td>-en</td>
<td>-e</td>
<td>76%</td>
<td>2%</td>
<td>2%</td>
<td>9%</td>
<td>11%</td>
<td>0%</td>
</tr>
<tr>
<td>switter</td>
<td>pullover</td>
<td>Pullover</td>
<td>-n</td>
<td>ø</td>
<td>51%</td>
<td>27%</td>
<td>7%</td>
<td>4%</td>
<td>11%</td>
<td>0%</td>
</tr>
<tr>
<td>ruus</td>
<td>flower</td>
<td>Blume</td>
<td>-en</td>
<td>-n</td>
<td>89%</td>
<td>0%</td>
<td>0%</td>
<td>2%</td>
<td>9%</td>
<td>0%</td>
</tr>
<tr>
<td>spiker</td>
<td>nail</td>
<td>Nagel</td>
<td>-n</td>
<td>U</td>
<td>64%</td>
<td>0%</td>
<td>0%</td>
<td>31%</td>
<td>4%</td>
<td>0%</td>
</tr>
<tr>
<td>reiluk</td>
<td>curtain</td>
<td>Vorhang</td>
<td>-en</td>
<td>U-e</td>
<td>93%</td>
<td>2%</td>
<td>0%</td>
<td>0%</td>
<td>4%</td>
<td>0%</td>
</tr>
<tr>
<td>brek</td>
<td>pants</td>
<td>Hose</td>
<td>-en</td>
<td>-n</td>
<td>96%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>4%</td>
<td>0%</td>
</tr>
<tr>
<td>schört</td>
<td>shirt</td>
<td>Hemd</td>
<td>-en</td>
<td>-n</td>
<td>87%</td>
<td>11%</td>
<td>0%</td>
<td>0%</td>
<td>2%</td>
<td>0%</td>
</tr>
<tr>
<td>riam</td>
<td>belt</td>
<td>Gürtel</td>
<td>-er</td>
<td>ø</td>
<td>4%</td>
<td>87%</td>
<td>0%</td>
<td>0%</td>
<td>2%</td>
<td>7%</td>
</tr>
<tr>
<td>hüunjmot</td>
<td>mushroom</td>
<td>Pilz</td>
<td>-en</td>
<td>-e</td>
<td>40%</td>
<td>51%</td>
<td>0%</td>
<td>0%</td>
<td>2%</td>
<td>7%</td>
</tr>
<tr>
<td>kwaast</td>
<td>paintbrush</td>
<td>Pinsel</td>
<td>-er</td>
<td>ø</td>
<td>33%</td>
<td>56%</td>
<td>0%</td>
<td>0%</td>
<td>7%</td>
<td>4%</td>
</tr>
<tr>
<td>dööbel</td>
<td>dice</td>
<td>Würfel</td>
<td>-bler</td>
<td>ø</td>
<td>0%</td>
<td>82%</td>
<td>0%</td>
<td>9%</td>
<td>4%</td>
<td>4%</td>
</tr>
<tr>
<td>Word</td>
<td>Category</td>
<td>German</td>
<td>Gender</td>
<td>Case 1</td>
<td>Case 2</td>
<td>Case 3</td>
<td>Case 4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>----------</td>
<td>-------------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>basel</td>
<td>brush</td>
<td>Bürste</td>
<td>-sler</td>
<td>$n$</td>
<td>7%</td>
<td>78%</td>
<td>2%</td>
<td>4%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>boosel</td>
<td>table</td>
<td>Tisch</td>
<td>-sler</td>
<td>$-e$</td>
<td>2%</td>
<td>89%</td>
<td>0%</td>
<td>0%</td>
<td>4%</td>
<td>4%</td>
</tr>
<tr>
<td>dask</td>
<td>bowl</td>
<td>Schüssel</td>
<td>-er</td>
<td>$\phi$</td>
<td>4%</td>
<td>87%</td>
<td>0%</td>
<td>0%</td>
<td>4%</td>
<td>4%</td>
</tr>
<tr>
<td>sedel</td>
<td>kettle</td>
<td>Kessel</td>
<td>-dler</td>
<td>$\phi$</td>
<td>9%</td>
<td>80%</td>
<td>0%</td>
<td>4%</td>
<td>2%</td>
<td>4%</td>
</tr>
<tr>
<td>tuun</td>
<td>fence</td>
<td>Zaun</td>
<td>-er</td>
<td>$U$-$e$</td>
<td>2%</td>
<td>89%</td>
<td>0%</td>
<td>2%</td>
<td>2%</td>
<td>4%</td>
</tr>
<tr>
<td>fask</td>
<td>fish</td>
<td>Fisch</td>
<td>-er</td>
<td>$-e$</td>
<td>0%</td>
<td>93%</td>
<td>0%</td>
<td>0%</td>
<td>4%</td>
<td>2%</td>
</tr>
<tr>
<td>swaelstook</td>
<td>match</td>
<td>Streichholz</td>
<td>-er</td>
<td>$\phi$</td>
<td>4%</td>
<td>91%</td>
<td>0%</td>
<td>0%</td>
<td>2%</td>
<td>2%</td>
</tr>
<tr>
<td>kai</td>
<td>key</td>
<td>Schlüssel</td>
<td>-er</td>
<td>$\phi$</td>
<td>4%</td>
<td>84%</td>
<td>0%</td>
<td>7%</td>
<td>4%</td>
<td>0%</td>
</tr>
<tr>
<td>mös</td>
<td>mouth</td>
<td>Mund</td>
<td>-er</td>
<td>$U$-$er$</td>
<td>27%</td>
<td>27%</td>
<td>0%</td>
<td>31%</td>
<td>11%</td>
<td>4%</td>
</tr>
</tbody>
</table>
Appendix B. Materials for grammatical gender tasks

Each of the grammatical gender tasks tested each of the following items.

Table B-1. Stimuli for grammatical gender tasks

<table>
<thead>
<tr>
<th>Fering</th>
<th>English</th>
<th>German</th>
<th>German gender</th>
<th>Consensus</th>
<th>Survey Gender</th>
<th>Masculine</th>
<th>Interchangeable</th>
<th>Common</th>
</tr>
</thead>
<tbody>
<tr>
<td>basel</td>
<td>brush</td>
<td>Bürste</td>
<td>die</td>
<td>high</td>
<td>Masc</td>
<td>88%</td>
<td>0%</td>
<td>12%</td>
</tr>
<tr>
<td>boosel</td>
<td>table</td>
<td>Tisch</td>
<td>der</td>
<td>high</td>
<td>Masc</td>
<td>96%</td>
<td>0%</td>
<td>4%</td>
</tr>
<tr>
<td>dask</td>
<td>bowl</td>
<td>Schüssel</td>
<td>die</td>
<td>high</td>
<td>Masc</td>
<td>80%</td>
<td>4%</td>
<td>16%</td>
</tr>
<tr>
<td>fask</td>
<td>fish</td>
<td>Fisch</td>
<td>der</td>
<td>high</td>
<td>Masc</td>
<td>92%</td>
<td>4%</td>
<td>4%</td>
</tr>
<tr>
<td>riam</td>
<td>belt</td>
<td>Gürtel</td>
<td>der</td>
<td>high</td>
<td>Masc</td>
<td>96%</td>
<td>4%</td>
<td>0%</td>
</tr>
<tr>
<td>tus</td>
<td>tooth</td>
<td>Zahn</td>
<td>der</td>
<td>high</td>
<td>Masc</td>
<td>96%</td>
<td>4%</td>
<td>0%</td>
</tr>
<tr>
<td>brag</td>
<td>bridge</td>
<td>Brücke</td>
<td>die</td>
<td>low</td>
<td>Masc</td>
<td>68%</td>
<td>12%</td>
<td>20%</td>
</tr>
<tr>
<td>brek</td>
<td>pants</td>
<td>Hose</td>
<td>die</td>
<td>low</td>
<td>Masc</td>
<td>52%</td>
<td>4%</td>
<td>44%</td>
</tr>
<tr>
<td>kwaast</td>
<td>paintbrush</td>
<td>Pinsel</td>
<td>der</td>
<td>low</td>
<td>Masc</td>
<td>68%</td>
<td>20%</td>
<td>12%</td>
</tr>
<tr>
<td>sees</td>
<td>cheese</td>
<td>Käse</td>
<td>der</td>
<td>low</td>
<td>Masc</td>
<td>48%</td>
<td>12%</td>
<td>40%</td>
</tr>
<tr>
<td>skreep</td>
<td>purse</td>
<td>Handtasche</td>
<td>die</td>
<td>low</td>
<td>Masc</td>
<td>64%</td>
<td>12%</td>
<td>24%</td>
</tr>
<tr>
<td>stääär</td>
<td>star</td>
<td>Stern</td>
<td>der</td>
<td>low</td>
<td>Masc</td>
<td>67%</td>
<td>17%</td>
<td>17%</td>
</tr>
<tr>
<td>bleed</td>
<td>newspaper</td>
<td>Zeitung</td>
<td>die</td>
<td>high</td>
<td>Com</td>
<td>17%</td>
<td>4%</td>
<td>79%</td>
</tr>
<tr>
<td>knif</td>
<td>knife</td>
<td>Messer</td>
<td>das</td>
<td>high</td>
<td>Com</td>
<td>17%</td>
<td>4%</td>
<td>79%</td>
</tr>
<tr>
<td>kop</td>
<td>cup</td>
<td>Tasse</td>
<td>die</td>
<td>high</td>
<td>Com</td>
<td>24%</td>
<td>0%</td>
<td>76%</td>
</tr>
<tr>
<td>riw</td>
<td>rake</td>
<td>Harke</td>
<td>die</td>
<td>high</td>
<td>Com</td>
<td>32%</td>
<td>0%</td>
<td>68%</td>
</tr>
<tr>
<td>triihuk</td>
<td>triangle</td>
<td>Dreieck</td>
<td>das</td>
<td>high</td>
<td>Com</td>
<td>24%</td>
<td>8%</td>
<td>68%</td>
</tr>
<tr>
<td>---------</td>
<td>----------</td>
<td>---------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>wel</td>
<td>bicycle</td>
<td>Fahrrad</td>
<td>das</td>
<td>high</td>
<td>Com</td>
<td>20%</td>
<td>4%</td>
<td>76%</td>
</tr>
<tr>
<td>hōōs</td>
<td>stocking</td>
<td>Strumpf</td>
<td>der</td>
<td>low</td>
<td>Com</td>
<td>40%</td>
<td>8%</td>
<td>52%</td>
</tr>
<tr>
<td>pokluad</td>
<td>pencil</td>
<td>Bleistift</td>
<td>der</td>
<td>low</td>
<td>Com</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>reiluk</td>
<td>curtain</td>
<td>Vorhang</td>
<td>der</td>
<td>low</td>
<td>Com</td>
<td>38%</td>
<td>8%</td>
<td>54%</td>
</tr>
<tr>
<td>ruus</td>
<td>flower</td>
<td>Blume</td>
<td>die</td>
<td>low</td>
<td>Com</td>
<td>29%</td>
<td>8%</td>
<td>63%</td>
</tr>
<tr>
<td>schört</td>
<td>shirt</td>
<td>Hemd</td>
<td>das</td>
<td>low</td>
<td>Com</td>
<td>33%</td>
<td>0%</td>
<td>67%</td>
</tr>
<tr>
<td>skaab</td>
<td>closet</td>
<td>Schrank</td>
<td>der</td>
<td>low</td>
<td>Com</td>
<td>33%</td>
<td>17%</td>
<td>50%</td>
</tr>
</tbody>
</table>

23 The form *blijant* ‘pencil’ was included in the survey, but pokluad is more common in the dialect spoken in western Föhr
Definite article production task

The order of the definite article elicitation task and the plural marking production task was counter balanced across participants. If the definite article production task appeared first, the following instructions were given. If the definite article production task appeared second, the instructions for the plural marking task were given.

Instructions: In diesem Spiel versuchen wir Alison ein Bisschen Friesisch beizubringen. Sie hat einen Arbeitsblatt und sie braucht deine Hilfe ihn richtig auszufüllen. Du wirst zwei Bilder sehen, und ein davon wird umkreist. Alison muss schreiben was umkreist ist. Du hilfst ihr, indem du ihr sagt was sie auf ihrem Blatt umkreisen muss. Wir üben ein paar Mal bevor wir beginnen, damit du sehen kannst, was wir machen. Alles klar? Willst du damit helfen? Prima!

‘In this game we will try to teach Alison a little Frisian. She has a worksheet and she needs our help to fill it out correctly. You will see two pictures, and one of them is circle. Alison has to write what is circle. You will help in by telling her what she needs to circle. We will practice with a few before we begin so that you can see what we are doing. Make sense? Do you want to help? Great!’


‘We will begin with three examples so that you can see what you say. Make sense? When we are half-way through you will get a sticker for your notebook. Good! Let’s start. Now we will speak Frisian’

Figure B-1. Example item for definite article production task

Assistant: Wat schochst dü heer?
‘What do you see here?’
Child: een stääër an een fask
‘a star and a fish’
Assistant: Sai Alison, wat heer umkreist as.
‘Tell Alison what is circled.’
Child: Alison, de fask as umkreisert.
‘Alison, the.M fish is circled.’
Pronoun production task


‘Here is a puppet. His name is Kurt. Now we will see what people are doing. But Kurt can’t see any colors. Can you help me and tell Kurt what colors things are? We will see and example first, and then we will start. Does that sound good? Do you want to play with me? Great. Let’s practice’


‘here is another girl. What is she doing? She is buying a purse. What color is the.MASC purse? Tell Kurt, what color the.MASC purse is.’

Child:  

hi as gül

Figure B-2. Example item for grammatical gender pronoun production task
Pronoun comprehension task


‘Now we are going to play a guessing game. Here is Andreas. He is a magician. He waves his magic wand and he makes one of two things disappear. In this game, you will guess which thing disappeared. You have to pay good attention and listen carefully. We will practice a few times, so that you see how it works. Does that sound good? Are you ready to start? Good! Let’s go. Now we will speak in Frisian’

Lukens heer! Heer san en kop uun en brek. Lachs in bilj as det kop, rochs in bilj as de brek.
‘Look here! Here is a cup and pants. Left in the picture is the.COM cup and right in the picture is the.MASC pants.’

Zauber Zauber! Hi.MASC as ferswünjen! Hi.MASC as wech!
‘Hocus Pocus! It.MASC is gone! It.MASC disappeared!’

Wise mi wat as ferswünjen.
‘Show me what disappeared.’

Figure B-3. Example item for grammatical gender comprehension task
Appendix C. Experimental materials for plural marking tasks

Each of the plural marking tasks tested each of the following items.

Table C-1. Stimuli for plural marking tasks

<table>
<thead>
<tr>
<th>Fering</th>
<th>English</th>
<th>German Plural</th>
<th>German Plural</th>
<th>German Plural</th>
<th>Fering Plural</th>
<th>–er</th>
<th>–en</th>
<th>other</th>
<th>–s</th>
<th>ø</th>
<th>don't know</th>
</tr>
</thead>
<tbody>
<tr>
<td>brag</td>
<td>bridge</td>
<td>Brücke</td>
<td>-n</td>
<td>match</td>
<td>-en</td>
<td>17%</td>
<td>78%</td>
<td>4%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>brek</td>
<td>pants</td>
<td>Hose</td>
<td>-n</td>
<td>match</td>
<td>-en</td>
<td>0%</td>
<td>100%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>käärs</td>
<td>cherry</td>
<td>Kirsche</td>
<td>-n</td>
<td>match</td>
<td>-en</td>
<td>4%</td>
<td>87%</td>
<td>0%</td>
<td>4%</td>
<td>0%</td>
<td>4%</td>
</tr>
<tr>
<td>ruus</td>
<td>flower</td>
<td>Blume</td>
<td>-n</td>
<td>match</td>
<td>-en</td>
<td>0%</td>
<td>96%</td>
<td>4%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>schört</td>
<td>shirt</td>
<td>Hemd</td>
<td>-n</td>
<td>match</td>
<td>-en</td>
<td>13%</td>
<td>83%</td>
<td>0%</td>
<td>4%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>riuw</td>
<td>rake</td>
<td>Harke</td>
<td>-n</td>
<td>match</td>
<td>-en</td>
<td>25%</td>
<td>71%</td>
<td>4%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>bleed</td>
<td>newspaper</td>
<td>Zeitung</td>
<td>-n</td>
<td>match</td>
<td>-en</td>
<td>0%</td>
<td>96%</td>
<td>4%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>kop</td>
<td>cup</td>
<td>Tasse</td>
<td>-n</td>
<td>match</td>
<td>-en</td>
<td>0%</td>
<td>96%</td>
<td>4%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>swaaelstook</td>
<td>match</td>
<td>Streichholz</td>
<td>-er</td>
<td>match</td>
<td>-er</td>
<td>91%</td>
<td>4%</td>
<td>4%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>henk</td>
<td>chick</td>
<td>Küken</td>
<td>ø</td>
<td>mismatch</td>
<td>-en</td>
<td>0%</td>
<td>96%</td>
<td>4%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>skaab</td>
<td>closet</td>
<td>Schrank</td>
<td>U-e</td>
<td>mismatch</td>
<td>-en</td>
<td>4%</td>
<td>92%</td>
<td>4%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>wel</td>
<td>bicycle</td>
<td>Fahrrad</td>
<td>U-er</td>
<td>mismatch</td>
<td>-en</td>
<td>0%</td>
<td>96%</td>
<td>0%</td>
<td>0%</td>
<td>4%</td>
<td>0%</td>
</tr>
<tr>
<td>höös</td>
<td>stocking</td>
<td>Strumpf</td>
<td>U-e</td>
<td>mismatch</td>
<td>-en</td>
<td>0%</td>
<td>88%</td>
<td>4%</td>
<td>8%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>reiluk</td>
<td>curtain</td>
<td>Vorhang</td>
<td>U-e</td>
<td>mismatch</td>
<td>-en</td>
<td>0%</td>
<td>100%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>stääär</td>
<td>star</td>
<td>Stern</td>
<td>-e</td>
<td>mismatch</td>
<td>-en</td>
<td>0%</td>
<td>91%</td>
<td>0%</td>
<td>4%</td>
<td>4%</td>
<td>0%</td>
</tr>
<tr>
<td>dööbel</td>
<td>dice</td>
<td>Würfel</td>
<td>ø</td>
<td>mismatch</td>
<td>-bler</td>
<td>87%</td>
<td>0%</td>
<td>9%</td>
<td>0%</td>
<td>4%</td>
<td>0%</td>
</tr>
</tbody>
</table>
Kwaast ‘paintbrush’ was not included in the survey. This item was included after the survey was conducted in order to include more items in the Mismatch-er category.

Irregular plural items were added after the online survey was conducted.
Plural marking Tasks

Plural marking production task

The order of the plural marking production task and the definite article production task were counter balanced across participants. If the plural marking production task was administered as the first task, the instructions from the definite article production task were administered, if the plural task was administered second, the following instructions were given:

Instructions: Dieses Spiel ist sehr ähnlich zu dem letzten Spiel. Noch ein Mal versuchen wir Alison ein Bisschen Friesisch beizubringen. Sie hat noch einen Arbeitsblatt und sie braucht deine Hilfe ihn richtig auszufüllen. Du wirst zwei Bilder sehen, und ein davon wird umkreist. Alison sieht die gleiche Bilder, aber weißt nicht, was um kreist ist. Du hilfst ihr, indem du ihr sagt was sie auf ihrem Blatt umkreisen muss. Wir üben ein paar Mal bevor wir beginnen, damit du sehen kannst, was wir machen. Alles klar? Kannst du ihr helfen? Willst du damit helfen? Prima!

‘This game is very similar to the last game. Once again we will try to teach Alison a little Frisian. She has another worksheet and she needs your help to fill it out correctly. You will see two pictures, and one of them is circled. Alison will see the same pictures, but does not know which is circled. You will help her by telling her what she needs to circle on her sheet. We will practice a few times before we begin so that you see what we are doing. Make sense? Can you help her? Do you want to help with that? Great!’


‘We will begin with three examples, so that you can see what you will say. Make sense? When we are half-way through you will get a sticker for your notebook. Good! Let’s start. We will speak Frisian now.’
Figure C-1. Example item for plural marking production task

**Plural marking acceptability task**

Instructions: *Bei diesem Spiel sage ich ein Paar Wörter. Manchmal sage ich sie richtig, aber manchmal sage ich sie ein Bisschen komisch. Kannst du mir sagen, wenn ich sie richtig sage, und wenn ich sie falsch sage? Wenn ich das Wort richtig sage, dann zeigst du hier (Gesicht #5), und wenn ich das Wort falsch sage, zeigst du hier (Gesicht #1).*


‘In this game I will say a pair of words. Sometimes I say it correctly and sometimes I say it a little funny. Can you tell me when I say it correctly and when I say it wrong? When I say it correctly, you will point here (points to Face #5), and when I say it incorrectly, you will point here (points to face #1).

But sometimes I say the word correctly, but a little funny. When I do that, you will point here (points to face #4). Sometimes I say it wrong, but not totally wrong. When I do that, you will point here (points to face #2). And when I say a word that is a little wrong and a little right, you can point in the middle. Ok? Make sense? Do you want to play with me? Good!’

Assistant: *Wat schochst dü heer? ‘What do you see here?’*

Child: *trii stäären an trii dööbler ‘three stars and three die’*

Assistant: *Sai Alison, wat heer umkreist as. ‘Tell Alison what is circled.’*

Child: *Alison, trii dööbler san umkreisert. ‘Alison, three die are circled.’*
Appendix D. Instructions for general proficiency task

General Proficiency


‘Here is our last game for today, and it a lot of fun! Here you see some pictures. The story is called “The living book”. Can you tell a story about these pictures? First, look at the pictures, and then tell me a funny story! If you need a word, you can ask me and then I will help you. You will tell the story twice. Once as practice and then we will make a recording of it.’

| Heer as een gus. |
| Here is a goose. |

| And then here? Do you see now? Now there are three geese /*gooses. |

| Haa ik det rocht saad? |
| Did I say that right? |
REFERENCES


VITA
Alison Eisel Hendricks

EDUCATION:
2014  Ph.D., in German and Language Science, Pennsylvania State University
Passed Comprehensive Exams: December 2011
2008  B.A. cum laude, Philosophy; German minor, Occidental College

GRANTS AND FELLOWSHIPS:
National Science Foundation Dissertation Improvement Grant. Carrie N. Jackson, Alison
Eisel Hendricks, Karen Miller. 2013-2015. The Quality and Quantity of Language
Input: North Frisian Child Second Language Acquisition. [$12,000]
U.S. Fulbright Full Research Grant to Germany. 2012-2013. [$20,000]
Pennsylvania State University Center for Global Science Early Career Development Award.
2012. [$2,000]
Center for Language Science Scholar Fellowship: The Pennsylvania State University, 2008-
2009. [Teaching release and stipend increase equal to $8,000]

PUBLICATIONS:
on Föhr, Germany. Us Wurk, 63:2, 31-53.
Miller, K., Hendricks, A. E. (2014). What impact does ambiguous input have on bilingual
language acquisition? Linguistic Approaches to Bilingualism. 4:3, 362-367.
L1 and L2 German. Linguistic Approaches to Bilingualism. 3:4, 448-477.

PRESENTATIONS:
tasks in elementary school age children with Specific Language Impairment.
Poster Presentation. American Speech and Hearing Association Convention,
Orlando, FL
Hendricks, A. E.; Miller, K.; Jackson, C.N. (2014). Regularization or Probability-
Matching? Acquisition of Inconsistent Gender Marking in Fering-Speaking
Children. Poster Presentation. Boston University Conference on Child Language
Development, Boston, MA
Hendricks, A. E. (2013). Status of North Frisian Instruction. TABU Dag Workshop on
Minority Languages in a Multilingual Europe, Groningen, Netherlands
Extra-Linguistic Factors. Day of Frisian Linguistics, Leeuwarden, Netherlands
Germanic Linguistics Annual Conference; Bloomington, IN
immersion vs. non-immersion context. International Symposium on Bilingualism;
Oslo, Norway
Linguistics Annual Conference; Austin, TX
Eisel, A. (2010). German Grammatical Gender Acquisition: Previous Language
Experience and Early Sensitivity to Phonological Cues. Germanic Linguistics
Annual Conference; Milwaukee, WI