This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.
The Developmental Trajectory of
Grammatical Gender: Evidence from Arabic

Hawra Alali

A thesis submitted in fulfilment of requirements for the
degree of
Doctor of Philosophy
to Linguistics and English Language
School of Philosophy, Psychology & Language Sciences
University of Edinburgh

2022
This thesis is dedicated to the memory of my late little sister,

Aliyah

(1989-2020)

whose presence I constantly miss!

Without you Aliyah, no achievement would be complete!
Publications and conference presentations


Transliteration symbols

Consonants:

<table>
<thead>
<tr>
<th>Arabic symbol</th>
<th>Transliteration symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>ب́</td>
<td>b</td>
</tr>
<tr>
<td>ت́</td>
<td>t</td>
</tr>
<tr>
<td>ٔ</td>
<td>t</td>
</tr>
<tr>
<td>ج́</td>
<td>J</td>
</tr>
<tr>
<td>ح́</td>
<td>h</td>
</tr>
<tr>
<td>خ́</td>
<td>x</td>
</tr>
<tr>
<td>د́</td>
<td>d</td>
</tr>
<tr>
<td>ذ́</td>
<td>ḍ</td>
</tr>
<tr>
<td>ر́</td>
<td>r</td>
</tr>
<tr>
<td>ز́</td>
<td>z</td>
</tr>
<tr>
<td>س́</td>
<td>s</td>
</tr>
<tr>
<td>ش́</td>
<td>š</td>
</tr>
<tr>
<td>ص́</td>
<td>s</td>
</tr>
<tr>
<td>ض́</td>
<td>ḍ</td>
</tr>
<tr>
<td>ط́</td>
<td>ṭ</td>
</tr>
<tr>
<td>ظ́</td>
<td>ẓ</td>
</tr>
<tr>
<td>ع́</td>
<td>ʿ</td>
</tr>
<tr>
<td>غ́</td>
<td>š</td>
</tr>
<tr>
<td>ف́</td>
<td>f</td>
</tr>
<tr>
<td>ق́</td>
<td>q</td>
</tr>
<tr>
<td>ك́</td>
<td>k</td>
</tr>
<tr>
<td>ل́</td>
<td>l</td>
</tr>
<tr>
<td>م́</td>
<td>m</td>
</tr>
<tr>
<td>ن́</td>
<td>n</td>
</tr>
<tr>
<td>ه́</td>
<td>h</td>
</tr>
<tr>
<td>و́</td>
<td>w</td>
</tr>
<tr>
<td>ي́</td>
<td>y</td>
</tr>
<tr>
<td>ء́</td>
<td>ʾ</td>
</tr>
</tbody>
</table>

Vowels

<table>
<thead>
<tr>
<th>Arabic symbol</th>
<th>Transliteration symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>ا́</td>
<td>a</td>
</tr>
<tr>
<td>ا̀</td>
<td>ā</td>
</tr>
<tr>
<td>َ</td>
<td>u</td>
</tr>
<tr>
<td>ُ</td>
<td>ū</td>
</tr>
<tr>
<td>ِ</td>
<td>i</td>
</tr>
<tr>
<td>ٍ</td>
<td>e</td>
</tr>
<tr>
<td>ی́</td>
<td>ī</td>
</tr>
<tr>
<td>ئ́</td>
<td>ei</td>
</tr>
</tbody>
</table>
## List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>First person</td>
</tr>
<tr>
<td>2</td>
<td>Second person</td>
</tr>
<tr>
<td>3</td>
<td>Third person</td>
</tr>
<tr>
<td>ACC</td>
<td>Accusative</td>
</tr>
<tr>
<td>CA</td>
<td>Classical Arabic</td>
</tr>
<tr>
<td>CDI</td>
<td>Communicative Development Inventory</td>
</tr>
<tr>
<td>CDS</td>
<td>Child-Directed Speech</td>
</tr>
<tr>
<td>CL</td>
<td>Numeral classifier</td>
</tr>
<tr>
<td>DAT</td>
<td>Dative</td>
</tr>
<tr>
<td>F</td>
<td>Feminine</td>
</tr>
<tr>
<td>GEN</td>
<td>Genitive</td>
</tr>
<tr>
<td>HA</td>
<td>Hasawi Arabic</td>
</tr>
<tr>
<td>INS</td>
<td>Instrumental</td>
</tr>
<tr>
<td>L1</td>
<td>First language</td>
</tr>
<tr>
<td>L2</td>
<td>Second language</td>
</tr>
<tr>
<td>LOC</td>
<td>Locative</td>
</tr>
<tr>
<td>M</td>
<td>Masculine</td>
</tr>
<tr>
<td>MGIT</td>
<td>Michigan Gender Identity Test</td>
</tr>
<tr>
<td>MSA</td>
<td>Modern Standard Arabic</td>
</tr>
<tr>
<td>N</td>
<td>Neuter</td>
</tr>
<tr>
<td>NOM</td>
<td>Nominative</td>
</tr>
<tr>
<td>NP</td>
<td>Noun Phrase</td>
</tr>
<tr>
<td>Pl</td>
<td>Plural</td>
</tr>
<tr>
<td>POSS</td>
<td>Possessive</td>
</tr>
<tr>
<td>PST</td>
<td>Past tense</td>
</tr>
<tr>
<td>S</td>
<td>Singular</td>
</tr>
<tr>
<td>SA</td>
<td>Standard Arabic</td>
</tr>
<tr>
<td>SV</td>
<td>Subject-Verb</td>
</tr>
<tr>
<td>TP</td>
<td>Tolerance Principle</td>
</tr>
<tr>
<td>VS</td>
<td>Verb-Subject</td>
</tr>
</tbody>
</table>
Abstract and lay summary

There is a well-documented bias among children to disproportionately rely on morphophonological cues to determine noun gender classes (Culbertson et al., 2019; Gagliardi & Lidz, 2014; Karmiloff-Smith, 1979; Levy, 1983; Pérez-Pereira, 1991). However, it is still unclear when and how children overcome this bias and approximate an adult-like weighting of morphophonological, semantic and syntactic cues for grammatical gender classes. The current study set out to uncover this developmental trajectory by studying the acquisition of grammatical gender in Arabic. Arabic has a two-gender system, nouns being either masculine or feminine. There is a fairly strong morphophonological cue for gender: more than 99% of the non-human feminine nouns end in -a while only 1% of the non-human masculine nouns have that ending. Crucially, Arabic also has a highly reliable association between natural and grammatical gender in human nouns (e.g., walad (M) ‘boy’/bint (F) ‘girl’, ʾab (M) ‘father’/ʾumm (F) ‘mother’).

The current study tried to find out when the association between natural gender and grammatical gender is acquired by language learners. This association in Arabic has no exceptions. L1 Arabic learners do not have to learn only this association but also that this association always takes precedence over the morphophonological association.

Two elicited production experiments were conducted with 202 native Hasawi Arabic-speaking participants of five age groups: 2;6-3;11, 4;0-5;11, 6;0-
10;11, 11;0-15;11, and adults. The first experiment investigated the interaction between morphophonological (ending), semantic (natural gender), and syntactic (agreement) cues using novel nouns. The second experiment investigated the effect of natural gender further by providing no syntactic cues.

The results of this investigation revealed that the interaction between morphophonological information and semantic information follows a gradual developmental path. A number of factors contribute to the discovery of the cues to grammatical gender classes and using them in attributing gender to new nouns. Among the most important factors are, the degree of reliability of gender markers, the overt-coding of gender markers, the uniformity of those markers, and the consistency of semantic cues. Overall, this data from Arabic demonstrates that, despite early bias for morphophonological cues, sensitivity to semantic and syntactic cues exists as early as 2 to 3 years. The development of grammatical gender from that stage therefore involves adjustments in cue weighting rather than discovery of the cues. It is not until adolescence that Arabic learners reach adult-like weighting of semantic cues in assigning gender in the language.
Acknowledgements

Indeed, all praise and thanks are due to Allah (SWT) for guiding me throughout this project.

It would not have been possible to complete this thesis without the help and support of so many wonderful people.

Firstly, I would like to thank my inspiring supervisors, Prof. Peter Ackema and Prof. Mits Ota, for guiding me throughout this journey. Thank you for the continuous support, advice, encouragement, understanding, and valuable comments. As I have always said, and will continue to say, this project would not have been successful without you both by my side. Your special touch in directing my research to the right destination leaves me deeply indebted to you both. Without you, this thesis would never have been completed. Thank you from the bottom of my heart.

I would also like to thank my internal examiner, Prof. Jennifer Culbertson. Thank you, Jenny, for your willingness to attend my annual reviews and for your invaluable contributions, comments, and suggestions. Thank you for giving me the courage and confidence to submit abstracts and present my work to esteemed colleagues at educational conferences. I could not have done so without your encouragement. Thank you for your constructive comments on the first version of this thesis.
Thank you also to my external examiner, Dr. Alex Bellem. Thank you for your valuable comments and suggestions.

Thank you to my parents. If it were not for your constant prayers, unwavering belief, and amazing support, I could never have achieved the completion of my thesis. The desire to reunite with you motivated me to continue working hard every day to finish this project.

Thank you to my husband, Hussain. Thank you for leaving everything behind, believing in me, and accompanying me in chasing my dream. Thank you for being a shoulder to cry on. Thank you for being the best father for our little ones, and for being a mother too when needed. Thank you for helping in every possible way. Thank you for giving me space when I needed it, and for being the calm in the storm when I was ill-temper ed. Thank you for carrying more than you can bear.

Thank you to my sister, Basmah. It is almost impossible for me to put into words how grateful I am for you. Thank you for being part of my best and worst moments. Thank you for your unconditional support. I know that nobody else would have done all that you have for me, only YOU. Thank you for everything – from the massive, crazy, purposeless voice messages and endless phone calls, to taking care of me and my little ones and trying to make us happy every day. Thank you for always being there for us.

Thank you to my sister, Wedad. In addition to the daily phone calls and messages, and all the support, love, and care you have shown me, thank you for turning the experimental line drawings I created into real 3D characters. Thank you
for all your effort and for every minute you spent creating them and for finishing
them on time. I knew that my ideas were safe in your hands. I am grateful to have
an incredibly talented artist like you for a sister. Thank you also for helping me with
the administrative work throughout my scholarship. Thank you for getting what is
in my mind before uttering it.

Thank you to my two little angels, Fatimah and Hala. Thank you, Fatimah,
for all the lovely cards you made for me and for every time you told me “Mummy
you’re the best mother in the whole world”. Thank you for being my mini-me in
your own way. Thank you, Hala, for giving me the motivation I needed to keep
moving forward and to be the best version of myself that I could possibly be. Thank
you both for the comforting hugs and for all the smiles and joy you brought to even
the hardest of days. Thank you for understanding about all the times I could have
spent playing and having fun with you but I spent working instead. I owe you every
moment.

Thank you to my sister, Ameerah. Thank you for taking care of my little
Hala when we were separated by the pandemic. Thank you for being a mother to
her when I was not there. With you, I knew my daughter would be safe, nurtured,
and loved.

Thank you to my two sisters, Fathiyah and Maryam. And thank you to my
sister-in-law, Layla and my mother-in-law. Thank you for all the contacts you made
to help me recruit participants, and for regularly turning your homes into
experimental laboratories.
Mustafa, my brother, and the person I look up to. Thank you for all the smiles you put on mine and my little one’s faces. Thank you for always being there for us.

Abdulaziz, my little brother and instant IT. Thank you for everything you did to help me when I got stuck, and for all the things you have taught me and continue to teach me in life.

Abeer, my friend and officemate. Thank you for being part of my journey. Thank you for every minute we wasted together, for every cup of coffee we enjoyed together, and for all the simple things we learned from one another, and of course, for all the generative and non-generative discussions we had.

Thank you to all the parents and participants who took part in this study. Without you, this research would not have been possible.

Finally, I wish to thank everybody in the Dugald Stewart Building, especially Katie Keltie, the first person I met at the University of Edinburgh and the first person I contacted whenever I needed any kind of help, and to Shirley Yuill, for the cheery morning greetings and friendly chats.
# Table of contents

PUBLICATIONS AND CONFERENCE PRESENTATIONS .......................... III

TRANSLITERATION SYMBOLS ........................................ IV

LIST OF ABBREVIATIONS ............................................. V

ABSTRACT AND LAY SUMMARY ...................................... VII

ACKNOWLEDGEMENTS ................................................ IX

TABLE OF CONTENTS ................................................ XIII

LIST OF TABLES ...................................................... XVII

LIST OF FIGURES ..................................................... XVIII

1. CHAPTER 1: INTRODUCTION ........................................ 1

   1.1 INTRODUCTION .................................................. 1
   1.2 WHY GRAMMATICAL GENDER? ................................ 7
   1.3 SETTING THE SCENE .......................................... 13
   1.3.1 THE SCOPE OF THE THESIS AND RESEARCH QUESTIONS 14
   1.3.2 KEYWORDS ................................................ 17
   1.3.3 RESEARCH QUESTIONS ................................... 18
   1.4 THE STRUCTURE OF THE THESIS ............................ 19

2. CHAPTER 2: LITERATURE REVIEW .................................. 21

   2.1. INTRODUCTION ............................................... 21
   2.2. CUES FOR GENDER CLASS .................................. 21
   2.3. RULE HIERARCHY/RANKING ................................. 31
   2.4. THE NOTION OF DEFAULT .................................. 42
   2.5. CHAPTER SUMMARY ........................................ 51

3. CHAPTER 3: THE ARABIC GENDER SYSTEM ......................... 53
6. CHAPTER 6: EXPERIMENT III: THE EFFECT OF NATURAL GENDER 166

6.1. INTRODUCTION 166
6.2. METHOD 169
6.3. RESULTS 171
6.4. DISCUSSION 183

7. CHAPTER 7: GENERAL DISCUSSION 188

7.1. INTRODUCTION 188
7.2. THE DEVELOPMENTAL TRAJECTORY OF GRAMMATICAL GENDER 189
7.3. MORPHOPHONOLOGICAL AND SYNTACTIC ASYMMETRY 214
  7.3.1. CUE UNIFORMITY AND OVERT CODING 216
  7.3.2. MASculine OR LACK OF GENDER 224
  7.3.3. PRODUCTIVE VS. UNPRODUCTIVE: THE TOLERANCE PRINCIPLE 228
  7.3.4. RULE HIERARCHY/ RANKING 239

CHAPTER 8: CONCLUSION AND RECOMMENDATIONS 253

8.1. CONCLUSION AND MAIN CONTRIBUTION 253
8.2. LIMITATIONS AND FUTURE DIRECTIONS 259

REFERENCES 263

APPENDIX A: THE ARABIC SCRIPT FOR EXPERIMENT I WITH SOME EXAMPLE ANSWERS 277

APPENDIX B: EXAMPLE ARABIC SCRIPT FOR EXPERIMENT II WITH EXAMPLE ANSWERS 278

APPENDIX C: EXAMPLE ARABIC SCRIPT FOR EXPERIMENT III WITH EXAMPLE ANSWERS 279

APPENDIX D: PICTURES USED FOR INTRODUCING THE WORDS ‘MOTHER’ AND ‘FATHER’ IN EXPERIMENT I 280

APPENDIX E: PICTURES OF THE TOYS USED IN EXPERIMENT I 281
List of tables

Table 2.1. Declension-gender correlation in Russian. .................................................. 24
Table 2.2. Hierarchies of physical objects investigated by Rosch et al. (1976). .. 33
Table 3.1. Example of root and pattern morphology in Standard Arabic .......... 65
Table 3.2. The correlation between the natural gender information and the morphophonological information in the four corpora........................................... 82
Table 3.3. The distribution of the endings found in the no -a words. ............... 84
Table 4.1. List of the familiar items used in experiment I. ................................. 97
Table 4.2. The distribution of the agreeing elements by age group.................. 99
Table 4.3. Mean correct grammatical gender attributed to the different items by each age group................................................................. 103
Table 4.4. The types of errors produced by the children in experiment I ........ 114
Table 5.1. List of novel nouns used in Experiment II........................................... 129
Table 5.2. The mean rate of the feminine responses attributed by the five age groups in each condition. ................................................................. 134
Table 5.3. The effect size of each type of information in each age group (ηp2 and η2). ................................................................. 143
Table 5.4. The distribution of the no -a endings in the four corpora vs. in the experimental items. ................................................................. 149
Table 5.5. Summary of the one-way within-subjects ANOVA for the effect of noun ending in each age group......................................................... 152
Table 6.1. List of the items used in Experiment III. .............................................. 170
Table 7.1. Mappings between gender and nominative singular suffixes in Icelandic. ......................................................................................... 231
Table 7.2. Numerical distribution of the singular nouns produced by the Egyptian children in Salama Corpus (Salama & Alansary, 2016) by ending and gender. 233
Table 7.3. Numerical distribution of the singular nouns produced by the children in Salama Corpus (Salama & Alansary, 2016) by ending and gender .......... 236
List of figures

Fig. 2.1. Assignment of gender to Norwegian bonde 'peasant'. (Nesset, 2006, p. 1374). ............................................................................................................................... 38
Fig. 2.2. Assignment of gender to Norwegian bonde 'peasant'. (Nesset, 2006, p. 1374). ............................................................................................................................... 51
Fig. 3.1. The distribution of the gender classes (masculine & feminine) in Arabic. ............................................................................................................................................ 79
Fig. 3.2. The distribution of the correlation between morphophonological ending and grammatical gender. ............................................................................................................. 81
Fig. 4.1. Examples of the toys used in experiment I ................................................................................................................................. 96
Fig. 4.2. Mean rate of feminine responses by age and correct grammatical gender. Error bars represent standard errors of the mean (SEM). ........................................... 104
Fig. 4.3. Mean rate of feminine responses by age and noun ending. Error bars represent standard errors of the mean (SEM). ................................................................. 105
Fig. 4.4. Mean rate of feminine responses by age and natural gender. Error bars represent standard errors of the mean (SEM). ........................................................................... 106
Fig. 4.5. Mean rate of feminine responses by age and noun ending given by the participants who passed both screening criteria. Error bars represent standard errors of the mean (SEM). ......................................................................................................................... 109
Fig. 4.6. Mean rate of feminine responses by age and natural gender given by the participants who passed both screening criteria. Error bars represent standard errors of the mean (SEM). ....................................................................................................................................... 110
Fig. 5.1. Examples of the 3D characters used in experiment II ................................................................................................................................. 131
Fig. 5.2. Mean rate of feminine responses by age × word ending × agreement × natural gender ........................................................................................................................... 135
Fig. 5.3. Mean rate of feminine responses by age and syntactic agreement in prompt. Error bars represent standard errors of the mean (SEM). ...................... 137
Fig. 5.4. Mean rate of feminine responses by age and noun ending. Error bars represent standard errors of the mean (SEM). ................................................................. 138
Fig. 5.5. Mean rate of feminine responses by age and natural gender. Error bars represent standard errors of the mean (SEM). ........................................................................... 139
Fig. 5.6. Mean rate of feminine responses by age × word ending × agreement × natural gender (re-analysed). ....................................................................................................... 145
Fig. 5.7. Mean rate of feminine responses by age and syntactic agreement in prompt (re-analysed). Error bars represent standard errors of the mean (SEM). 146
Fig. 5.8. Mean rate of feminine responses by age and noun ending (re-analysed). Error bars represent standard errors of the mean (SEM). .............................................. 147
Fig. 5.9. Mean rate of feminine responses by age and natural gender (re-analysed). Error bars represent standard errors of the mean (SEM). .............................................. 148
Fig. 5.10. Mean rate of feminine responses summarised by the type of ending in the no –a experimental items. Error bars represent standard errors of the mean (SEM).

Fig. 5.11. Mean rate of feminine responses attributed to each type of ending by the different age groups. Error bars represent standard errors of the mean (SEM).

Fig. 5.12. Mean rate of feminine responses by age × word ending × agreement × natural gender.

Fig. 6.1. Mean rate of feminine responses by age and natural gender. Error bars represent standard errors of the mean (SEM).

Fig. 6.2. Mean rate of feminine responses by age, natural gender and item familiarity. Error bars represent standard errors of the mean (SEM).

Fig. 6.3. Mean rate of feminine responses by natural gender and noun ending. Error bars represent standard errors of the mean (SEM).

Fig. 6.4. Mean rate of feminine responses attributed to the different endings by the different age groups. Error bars represent standard errors of the mean (SEM).

Fig. 6.5. Mean rate of feminine responses attributed by the five age groups to the three nouns ending in -a. Error bars represent standard errors of the mean (SEM).

Fig. 6.6. Mean rate of feminine responses attributed to novel items by age and natural gender (re-analysed). Error bars represent standard errors of the mean (SEM).

Fig. 7.1. The mean rate of feminine responses by age × noun ending × syntactic agreement × natural gender (circles highlight the pre-schoolers group responses).

Fig. 7.2. The mean rate of feminine responses by age × noun ending × syntactic agreement × natural gender (circles highlight the three older groups responses).

Fig. 7.3. The mean rate of feminine responses by age × noun ending × syntactic agreement × natural gender (circles highlight the youngest age group responses).

Fig. 7.4. Mean rate of feminine response attributed by the youngest group to the different endings. Error bars represent standard errors of the mean (SEM).

Fig. 7.5. Gender assignment to bint.

Fig. 7.6. Gender assignment to walad.

Fig. 7.7. Gender assignment to dukttur-a.

Fig. 7.8. Gender assignment to xalifa (wrong assignment).

Fig. 7.9. Gender assignment to xalifa (correct assignment).

Fig. 7.10. Gender assignment to tawila.

Fig. 7.11. Gender assignment to kursi.

Fig. 7.12. Gender assignment to shams.
Chapter 1: Introduction

1.1 Introduction

Children need to learn the grammatical gender\footnote{Grammatical gender and gender will be used interchangeably throughout this thesis.} of every noun in the language they are acquiring if it happened to be one of the 40% of the world’s languages that have grammatical gender systems (Corbett, 2013a). To acquire Hasawi Arabic (HA), children need to learn, among other things, that the words denoting their bodies, heads, mouths, noses, hair, and backs are masculine whereas the words denoting their eyes, ears, hands, and feet are feminine. Nonetheless, learning every noun’s gender is not the ultimate destination children are heading to. They need to learn in addition to that that the gender of the noun affects or more precisely controls the forms of other syntactically related elements. That is, they need to learn gender agreement. Gender agreement can take place between nouns and many other elements such as determiners, adjectives and verbs. The sentences (1 - 4) are representative examples from Arabic, French, Russian, and Spanish.\footnote{Throughout the thesis, the gender of the noun will be given in parentheses. The gender of the agreeing target will also be given in parentheses or after/before a hyphen if it has an affixal nature.}

1

Grammatical gender and gender will be used interchangeably throughout this thesis.

2 Throughout the thesis, the gender of the noun will be given in parentheses. The gender of the agreeing target will also be given in parentheses or after/before a hyphen if it has an affixal nature.
1) Standard Arabic

a. \( \text{hādā}^{3} \ 'al-qamar \quad sāṭi' \)
   
   this(M) the-moon(M) bright(M)
   
   ‘This moon is bright.’

b. \( \text{hādīhi} \ 'al-šams}^{4} \quad sāṭi-'a \)
   
   this(F) the-sun(F) bright-F

   ‘This sun is bright.’

c. \( \text{jā 'a} \ 'al-walad \quad 'al-jamīl \)
   
   came the-boy(M) the-beautiful(M)

   ‘The beautiful boy came.’

d. \( \text{jū 'a-t} \ 'al-bint \quad 'al-jamīl-a \)
   
   came-F the-girl(F) the-beautiful-F

   ‘The beautiful girl came.’

2) French (Seigneuric et al., 2007)

a. \( \text{Le} \quad \text{verre} \quad \text{est} \quad \text{sur} \quad \text{la} \quad \text{table}. \)

   The(M) glass(M) is on the(F) table(F)

   ‘The glass is on the table.’

---

3 Henceforth, the agreeing targets will be given in bold.

4 The lateral \( l \) in the Arabic definite article \( 'al \) assimilates to the following coronal consonant. This will not be shown in this thesis because of its irrelevance and because of the recommended practice in Arabic transcription (Reichmuth, 2011).
b. Est-ce que tu as un frère ou une soeur?

Is-it that you have a(M) brother(M) or a(F) sister(F)?

‘Do you have a brother or a sister?’

3) Russian (Rodina & Westergaard, 2012)

a. Strann-yj malčik-Ø ležal-Ø na skamejke.

Strange-M boy-M lie.PST-M on bench

‘A strange boy was lying on the bench.’

b. Strann-ača devočk-a ležal-a na skamejke.

Strange-F girl-F lie.PST-F on bench

‘A strange girl was lying on the bench.’

c. Strann-oe piskm-o ležal-o na skamejke.

Strange-N letter-N lie.PST-N on bench

‘A strange letter was lying on the bench.’

4) Spanish (Pérez-Pereira, 1991)

a. Un italiano sincer-o

a(M) Italian-M sincere-M

‘A sincere Italian.’
The examples in (1) – (4) show that syntactic agreement is the defining property of the grammatical gender of a particular noun. In the Arabic examples (1), qamar (moon) and walad (boy) are masculine because they cooccur with masculine agreeing forms whereas šams (sun) and bint (girl) are feminine because they cooccur with feminine agreeing forms. Similarly in French (2), table (table) and soeur (sister) are feminine while verre (glass) and frère (brother) are masculine for the same reasons. In Russian (3), malkčik-Ø (boy) is masculine, devočk-a (girl) is feminine, and piskm-o (letter) is Neuter and they cooccur with respectively masculine, feminine, and neuter agreeing verbs and adjectives. The same applies for the Spanish examples in (4). Thus, gender is an inherent feature of the noun itself that is expressed through the forms of other elements such as articles, adjectives, verbs, and relative pronouns, i.e. through agreement.

The child may approach the task of learning gender in different ways. They may simply memorize the cooccurrences of the noun and the determiner or the noun and the adjective. There is converging evidence that infants seem to start learning their languages in an item-based or lexically based pattern (see Dąbrowska, 2004 for a review). Some infants like Travis, Tomasello’s daughter, acquire verb-specific combinations (Tomasello, 2000). Other infants first
acquire the base form of the noun, which could be a bare or an inflected form depending on the language (Stephany & Voeikova, 2009). My two-and-three-month-old daughter started using the first-person pronoun I in a fixed context “I like this” but she used her name in all other contexts where she should use I. Rote learning can manifest itself in units as small as single words, bigger units such as determiner + noun, or even bigger as “clap your hands.” or “what colour is this?” Language learning does not always progress from small to big or from one word to two words until it reaches a complex sentence. Children seem to start big (Arnon & Ramscar, 2012; Peters, 1977; Skarabela et al., 2021; inter alia) and they learn big units as unanalysable chunks. A French infant for example may learn la main (hand) as lamain or le mari (husband) as lemari (Tucker et al., 1968). My daughter used *go awayed instead of went away and instead of *goed away which suggests that she was learning verb + particle as an unanalysable chunk.

Nevertheless, it has been found that learning arbitrary classes that have syntactic dependencies such as gender seems to be difficult if not impossible by relying only on the distributional evidence or co-occurrences (Braine, 1987; Braine et al., 1990; Brooks et al., 1993; Frigo & McDonald, 1998; Smith, 1969). This learning however becomes possible when at least a subset of members in each class shares some intrinsic markers. A great body of literature has tried to describe those markers or properties in order to find out the principles according to which nouns are classified into gender classes in different languages (see Corbett, 1991). Simply put, Corbett found that those markers could be of two types: semantic markers
represented by the meaning or the referent of the noun and morphophonological markers represented by the form of the noun. One of the most common semantic properties with which gender tends to correlate is natural gender (males vs. females) and such a correlation is manifested by finding the nouns denoting males in one gender class, i.e. masculine, and the nouns denoting females in another gender class, i.e. feminine, in most of the world’s languages (75% of the languages that have gender systems according to Corbett, 2013b). Morphophonological properties are of different types such as word beginnings, word endings, suffixes, and prefixes. However, when children start acquiring gender, they are often found to be sensitive to the second (i.e. morphophonological) but not to the first (i.e. semantic) type of those properties (Cain et al., 1987; Gagliardi & Lidz, 2014; Karmiloff-Smith, 1979; Pérez-Pereira, 1991; Seigneuric et al., 2007; and Szagun et al., 2007, inter alia). Adult learners on the other hand, are found to be more sensitive to the semantic properties shared by the nouns in each gender class (Carroll, 1999; Culbertson et al., 2019; Gagliardi & Lidz, 2014). It is still however not clear when a child learner arrives at the adult-like gender system and when semantic

---

5 I will give more details on this, where relevant, in the next sections and in the subsequent chapters.

6 There are other semantic properties with which gender correlates in languages such as animacy, shape, size, edibility. Also, the correlation between grammatical gender and natural gender is not without exceptions. In addition to that, there are languages that have all human beings or living things in one class; i.e., those languages do not show a gender distinction between males and females. All these issues will be discussed in more detail (where relevant) in the next chapter and in the discussion chapter.
information is integrated and entertained by the language learner/user as part of the
gender assignment system.

1.2 Why grammatical gender?

Gender is an inherent feature of the noun unlike other nominal features like
number or case. Considering number for example, a noun can be singular, plural,
or dual if the language has a dual number. Gender on the other hand violates
Corbett’s (2012, p. 45) exhaustiveness criteria: “Every lexical item of every part of
speech has available to it all values of all features. (Alternatively: every feature
value applies to all lexical items)”. Generally speaking, a noun has only one gender
value and this value has to be learned as part of the noun in order for the noun to
function syntactically as the controller in agreement. Agreement is the key function
of gender (Hockett, 1958). Though for instance one can study the acquisition of
number (morphology) and number agreement as separate phenomena, the
acquisition of gender cannot be studied away from gender agreement because
agreement is the defining element of gender. Thus, gender is related to both
morphology and syntax; a morphosyntactic feature.

Furthermore, the meaning of gender is, in most cases, arbitrary. This is
obvious in the sentence stated at the beginning of this thesis. Though head, mouth,
nose, hair, back, eye, ear, hand, and foot are all body parts, some of the words that

7 ‘Pluralia tantum’ is a special case. See Corbett (2019) for a discussion.
denote those body parts are masculine while some are feminine in Arabic. Gender contributes no additional meaning to the utterance especially in the case of inanimate nouns. This arbitrariness is illustrated in the two examples from English (5) and Italian (6) (cited from Audring, 2014). Nothing is contributed to the meaning of the Italian sentence by marking ‘story’ as feminine.

5) A long story.
6) *Un-a lung-a storia.*

   a-F long-F story(F)

   ‘A long story.’

Nonetheless, as I mentioned above, in three quarters of the languages which have semantic assignment systems (Corbett, 2013b), grammatical gender correlates with natural gender or biological sex, i.e. the distinction between males and females. The gender assignment principles that are based on biological sex distinctions are not only part of the gender system of those languages but in most cases, with very few exceptions, those principles must override the morphophonological assignment principles in order for the nouns to be classified into the correct gender classes (see the discussion in sections 2.2 and 2.3, chapter 2).

It has been argued that gender can be acquired by children as a formal feature (meaning-less feature) early and easily (Levy, 1983; and Schlesinger, 1976).
Despite this, in most of the investigated systems such as Russian, German, French, and Spanish, semantic assignment principles must be learned in order for the acquisition process to reach its end state and in order for each noun to function properly.

Children’s ability to use natural gender in linguistic classification requires that they represent the conceptual distinction between males and females just as the conceptual distinction between one and more than one is a prerequisite for the linguistic distinction between singular and plural (Barner et al., 2007). Ferry et al. (2020) claimed that Italian infants were able to map meaning to morphophonological regularities (number and gender) by their first birthday except for masculine number distinction. They presented the infants with images of faces of young children, young adults, and elderly adults. They varied the representation of the images to test four different conditions (feminine singular vs masculine singular, feminine plural vs masculine plural, feminine singular vs feminine plural, and masculine singular vs masculine plural). The definite article and the final vowel of the noun in the labelling phrases represented the morphological distinctions that the authors were interested in. They found that children looked longer to the labelled image than to the distractor image. Their claim is that children were able to map the natural gender (and number) distinctions in the images to the morphological distinctions in the articles and noun-final vowels by their first birthday. Additionally, they argued that children succeeded in this despite that they have not yet learned the meaning of some of the labelling nouns they used such as
ragazzo/ragazza (man/woman). Ferry et al.’s finding is inconsistent with the findings that infants as old as 20 months failed to conceptually distinguish between one and more than one (Barner et al., 2007; Feigenson & Carey, 2005) or the fact that infants do not make the distinction between males and females before their third birthday (Dull et al., 1975). In other words, it is very unlikely that twelve-month old infants have the conceptual distinction between males and females, let alone mapping this to the grammatical distinction between feminine and masculine. It is very likely however that children were sensitive to the morphological regularities and that they learn grammatical gender before natural gender. Ferry et al.’s finding, though claimed to document the first evidence of young infants’ abilities to map meaning to morphological regularities, should not be overstated because the design they used has some limitations. As pointed out by the authors themselves, “we can conclude if infants know something about the distinctions, but not specifically what they know” (p. 44). Therefore, investigating gender from a developmental perspective contributes to our understanding of form-meaning mapping and how it develops with respect to different grammatical categories. If gender is acquired as a formal category, when does the meaningful correlation between natural gender and grammatical gender become an integrated part of the gender system? Investigating the acquisition of gender also contributes to the long-lasting debate between single route and dual route mechanisms for language acquisition.

---

8 The results of MGIT test (chapter 4) also support this finding.

9 I will not go through this here because it is beyond the scope of my thesis.
learning (Albright & Hayes, 2003; Clahsen, 1999; Marcus et al., 1995; McClelland and Patterson, 2002; Pinker and Prince, 1988; Yang 2002). Despite its seeming arbitrariness, a number of regularities are found in the gender systems of most languages. The language learner therefore may acquire the gender of nouns item by item following a single route mechanism, or may make some generalizations and use them productively and learn the exceptions item by item; i.e. the learner may follow a dual route.

Moreover, gender constitutes a perfect case for investigating the order in which a particular morphological paradigm is acquired. Neeleman and Weerman (1999) argued that the default form in a morphological paradigm is acquired earlier by children. The default form according to them is the form that lacks feature specifications or that is underspecified for features\(^\text{10}\). Their argument is that children overgeneralize the default form because they have not yet acquired the feature specifications required for using the other specified forms. In other words, the feature specifications required to produce the distinctions in a morphological paradigm are not yet part of the children’s system. In, almost all languages, there is a default linguistic gender which is used in cases of feature under-specification, as the Greek sentence in (7) (from Tsimpli & Hulk, 2013) illustrates. In this sentence, the neuter definite determiner to, rather than the masculine o or feminine i, is used because the complement phrase has no gender feature specifications. This default

\(^{10}\) The notion of default has been used in the linguistic literature to mean different things as I will discuss in detail in Chapter 2.
gender is expected to be learned first, and presumably overgeneralized, by language learners. Rethinking the non-canonicity of gender as a nominal feature, it seems that what is meant by default gender in gender systems is in essence default agreement. A noun has, almost always, one inherent gender and it is either that a learner learns the correct gender of the noun or not. What is meant by assigning a noun the default gender is that the default agreement is used with this noun. learners are therefore predicted to overgeneralize the default agreement until they learn the feature specifications necessary for learning the other non-default agreement patterns. So, in a two-gender system where the distinction is between masculine and feminine, and where masculine is the default gender as in Arabic and Spanish for example, learners are predicted to overgeneralize the masculine agreement until they learn the feature specification [FEM] required for assigning feminine agreement.

7) To pote tha erthi ine agnosto

The(N) when will come.3SG is unknown
‘When he will come is unknown.’

It is essential to adopt this definition for default gender, i.e that default gender is nothing but default gender agreement. For, in many contexts, the term ‘default gender’ is not used to refer to the gender with which the default agreement is used. Rather, it is used in reference to the most frequent gender or to the gender
that corresponds to the base form, i.e. the form that has no morphophonological marking. High frequency sometimes, but not necessarily always, corresponds to the default status of a gender class. I will discuss in more detail in Chapter 2 how the notion of default has been used to mean different things in the linguistic literature. For now, suffice it to say, that what I mean by default gender in this thesis is default agreement.

In sum, gender acquisition is multidimensional: it is a morphosyntactic feature that has to be learned as an intrinsic feature of the noun which itself correlates with the form and/or the meaning of the noun and which controls the morphophonology of syntactically related elements. Thus, the main aim of this thesis is to find out when the correlation between natural gender and grammatical gender becomes a productive pattern in a gender assignment system whose gender correlates with this concept and how the productivity of this pattern interacts with type frequency\(^\text{11}\). In the way to achieve this aim, several related issues will be investigated.

### 1.3 Setting the scene

The majority of the studies that investigated the morphosyntactic feature ‘gender’ were cross-sectional. Snapshots of the acquisition of gender by first and

\(^{11}\) Type frequency and productivity are not the same. A suffix could be very frequent in a language but not productive. For example, the suffix (-ment) is found in hundreds of words in English, but only four neologisms were attested with this suffix in the twentieth century (Haspelmath & Sims, 2013).
second language learners of different languages are found. This provides the main impetus for the current investigation. This thesis tries to fill the gap in the literature, by looking at the developmental trajectory of grammatical gender. To put it differently, this thesis is interested in tracking and tracing the development of the generalizations and patterns language learners are sensitive to as they progress from the two-word utterance stage to adulthood. I look at gender acquisition as a continuous process of updating the patterns and generalizations as more information is integrated into the system along the course of development.

1.3.1 The scope of the thesis and research questions

To make this contribution, I will look at the development of grammatical gender by L1 Hasawi Arabic learners/speakers, which is a contribution in itself as it provides data on the development of the gender system of a language understudied in this respect. Though Arabic gender has been an area of interest for many researchers (Alamry, 2019; Albirini et al., 2013; Moawad, 2006; inter alia), to the best of my knowledge, no study has looked at the developmental trajectory of the gender assignment patterns and generalizations entertained by the Arabic learners/speakers at the different stages throughout the course of development. Arabic is a perfect language for such investigation for many reasons.

➢ The Arabic gender system constitutes a perfect case for investigating the effect of the regularity and the predictivity of the system on the ease of discovery and the learnability of this system. The Arabic gender system is very
regular and predictive. Assessing the complexity of the Arabic gender system according to the three dimensions suggested in Audring (2014), namely the number of gender values, the number and nature of assignment rules, and the amount of formal marking, Arabic gender system appears to be among the very simple systems. To explain, Arabic has two gender classes, which is the minimum number and the most common number found in the world’s languages (Corbett, 2013a). Second, gender assignment principles in Arabic predict the correct gender of more than 90% of the lexicon. Third, gender is formally marked in a fair number of targets. Adjectives, demonstratives, personal pronouns, relative pronouns (though only in Standard Arabic), and verbs express gender agreement in Arabic. All these features could contribute to the ease of discovery and fast learning and mastering of the (Hasawi) Arabic gender system.

➢ In Arabic, unlike many other investigated languages, nouns are not obligatorily accompanied by agreeing determiners, but nouns are mostly used on their own because Arabic does not have an indefinite article and the definite article (when used) does not change according to the gender of the noun. This is important because it rules out the possibility of learning determiner + noun combinations as chunks or learning the gender of the noun from the determiner.

---

12 This is determined on the basis of the corpora analysis that will be presented in Chapter 3.

13 The existing discrepancies between Standard Arabic and the spoken dialectal variety under investigation, i.e. Hasawi Arabic will be presented and discussed in the background chapter on the Arabic gender system (Chapter 3).
Learning determiner + noun combinations could reflect a high degree of rote learning and item-specific learning at the early stages of language acquisition which might affect the reliability of the patterns and cues that the children learning languages with obligatory determiners are claimed to be sensitive to especially when studying familiar nouns. This is indeed not the case in many studies where novel nouns, not familiar nouns were used (Culbertson et al., 2019; Gagliardi & Lidz, 2014; Karmiloff-Smith, 1979; Pérez-Pereira, 1991, inter alia). Notwithstanding, it is always the case that investigating a different language with its special system illuminates issues and helps in providing a better understanding of the phenomenon under investigation, i.e. the acquisition of grammatical gender.

This thesis will also try to look into the order in which the Arabic gender paradigm is acquired and the role that default agreement plays in this order of acquisition. The results must be interpreted, however, with the caveat that Arabic is not the perfect case for such investigation. As I will show in Chapter 3, masculine agreement is the default agreement in Arabic, i.e. the agreement that lacks feature specifications. Yet, it is also the agreement that lacks morphophonological marking. In other words, the morphophonological marking of the Arabic feminine gender corresponds to the feature specification of this gender. Therefore, the early acquisition of masculine gender in Arabic could, but does not necessarily, indicate that the default gender is acquired earlier than the non-default gender. If the results
revealed an early acquisition of masculine gender compared to feminine gender, this could be due to the masculine agreement being the morphophonologically base, i.e. morphophonologically unmarked form which is reported to be acquired earlier than marked/inflected forms in many languages (Lust, 2006; Stephany & Voeikova, 2009).

Moreover, this thesis focuses on the development of the gender of singular nouns but not plural nouns because only singular nouns provide reliable evidence to the gender of the noun in Arabic. Gender and number interact in a special way in Arabic as will be discussed in chapter 3. These interactions between number and gender will not be considered in this thesis as the focus is on the acquisition of gender away from its interaction with number, so on the assignment of gender to singular nouns.

1.3.2 Keywords

- Gender assignment/attribution:

There are two ways in which the term gender assignment is used in the linguistic literature (Thornton, 2009). The first has to do with gender assignment in acquisition. It refers to the process through which children learn which gender value to attribute to a given noun in order for that noun to function syntactically as the controller of the agreement relation with items syntactically related to it. The second has to do with gender assignment in grammar. It refers to the rules that have been proposed to account for the distribution of gender across nouns in language systems.
Both interpretations of the term *gender assignment* are sometimes used interchangeably in the literature (see Corbett, 1991 for an overview).

The focus of this thesis is on the first process mentioned above. That is, it focuses on the question how children acquiring their mother tongue learn what gender they should attribute to a given noun. Some reference is also made to the second interpretation of *gender assignment*, however, in particular when presenting some proposals of gender assignment in grammar and discussing the potential implications of these proposals for the acquisition of gender (in chapter 2 and chapter 7). The reader will be alerted to this switch in the use of *gender assignment/attribution* where necessary.

- **Pattern/ generalization/ rule/ schema:**

  Patterns, generalizations, rules, and schemas are used interchangeably in this thesis, in an atheoretical way, to mean abstract representations that could be used productively in new contexts. Rule as an algebraic representation is only used when presenting Yang’s Tolerance Principle (chapter 7).

### 1.3.3 Research questions

This thesis sets out to answer two main questions. The first question is when do children learn the association between natural gender and grammatical gender as part of their representation of grammatical gender? The second question is: what
is the developmental trajectory that grammatical gender undertakes to reach its end state?

Attempting to answer these two general questions, a number of other important questions will be addressed throughout the thesis.

1.4 The structure of the thesis

In chapter 2, I present a review of gender acquisition studies. I highlight the role of the different cues to grammatical gender and how they are used by learners/speakers of different languages. I also present different proposals for rule hierarchy and rule ordering. I finish by discussing the notion of default in gender.

In chapter 3, I outline the Arabic gender system. I present the assignment principles, how gender is manifested in gender agreement in Arabic, and the distribution of the gender classes and the distribution of the noun-internal gender correlating information based on the analysis of four corpora. This chapter is intended to provide the background knowledge necessary for understanding the Arabic gender system. It also highlights the similarities and discrepancies between the spoken Hasawi Arabic dialect and Standard Arabic.

In chapter 4, I present the two screening criteria I used to filter the participants for Experiments II and III. The first part of this chapter presents the Michigan Gender Identity Test [MGIT] and its results. The second part presents Experiment I which is a semi-spontaneous production task used to check children’s
knowledge of gender agreement which is used as a filter for participation in the following experiments.

In chapter 5, I present Experiment II which tested the interaction between noun-internal (morphophonological and semantic) and noun-external (syntactic) information in assigning gender to novel nouns. This experiment aimed at tracking the weighting of those cues and how it changes as the language user develops from a child language learner into an adult language speaker.

In chapter 6, I present Experiment III which further investigated the role of natural gender in assigning gender to novel nouns while controlling for the effect of morphophonological cues and providing no syntactic information.

In chapter 7, I discuss the results of the three experiments and explain the implications they potentially have for the study of gender in particular and morphosyntactic categories in general.

In chapter 8, I conclude the thesis by summarizing the main contributions and discussing the limitations of the current investigations as well as providing some recommendations for future studies.
Chapter 2: Literature review

2.1. Introduction

Grammatical gender has elicited a large body of literature. I will focus in this review on the studies that looked at the acquisition of grammatical gender mainly by first language learners. In section 2.2, I will discuss the different types of information that might facilitate the learner’s acquisition of gender. I will summarise the findings in relation to the use of those types of information and how learners were found to behave in case of conflicting information. In section 2.3, I will present different accounts that were proposed to explain the interaction between gender assignment principles in language systems and the implications those accounts have for the acquisition of gender. In section 2.4, I will present the different notions of default in addition to some studies that investigated the notion of default in the acquisition of gender. Section 2.4 will summarise the discussion in this chapter.

2.2. Cues for gender class

As discussed in the previous chapter, the defining element of grammatical gender is syntactic agreement. Agreement is the relationship that defines gender by relating the noun syntactically to other elements such as determiners, adjectives, verbs, relative pronouns among others. Masculine nouns are nothing but nouns that co-occur with masculine agreeing forms and feminine nouns are nothing but nouns
that co-occur with feminine agreeing forms and the same applies for other gender classes. The agreeing forms vary cross-linguistically (Corbett, 1991). Some languages have gender marked articles and adjectives such as French and Spanish, other languages mark gender only on vowel-initial verbs, adjectives and adverbs such as Tsez, and others may have gender neutral articles but have gender marked adjectives and verbs such as Arabic. Examples 1-4 in the previous chapter show how articles, adjectives, and verbs take different forms according to the gender of the noun in different languages.

Thus, syntactic dependencies are what leaners ultimately need to acquire. Despite this, it has been found that learning arbitrary classes that have syntactic dependencies such as gender seems to be difficult if not impossible by relying only on the distributional evidence or cooccurrences (Braine, 1987; Braine et al., 1990; Brooks et al., 1993; Frigo & McDonald, 1998; Smith, 1969). In Braine et al.'s (1990) experiment for example, both children and adults failed to learn the arbitrary subclasses of the artificial language they tried to teach their subjects. This learning however becomes possible when at least a subset of members in each class shares some intrinsic markers. Those markers do not need to be of a particular nature (semantic or morphophonological). “The use of the features ‘male’ and ‘female’ is almost certainly not crucial to the learning. No doubt we could have made half the objects of one class striped, and half of the objects of the other class spotted, and achieved the same results.” (Braine, 1987, p.80).
The noun-internal properties with which gender tends to correlate hence make learning gender classes an accomplishable mission. Those properties can be of two sorts: semantic properties represented by the meaning of the noun and morphophonological properties represented by the form of the noun (Corbett, 1991). Several semantic properties with which gender tends to correlate are found in the world’s languages. Among these, we find natural gender or biological sex. In Tamil\textsuperscript{14} (a Dravidian language) for example, male humans are masculine, female humans are feminine and everything else is neuter. Similarly, some Indo-European languages such as Spanish, French, and Italian exhibit a correlation between natural gender and grammatical gender though different from Tamil. That is, in those languages, masculine gender has male humans in addition to many other things, and in the same way feminine gender has female humans in addition to many other things. Animacy is another semantic property with which gender tends to correlate in some languages as in Ojibwa (an Algonquian language). In Ojibwa, animate nouns belong to one gender whereas inanimate nouns belong to another. Other languages may classify nouns according to other semantic properties such as shape, size, and edibility.

Considering the formal properties with which gender tends to correlate, we find morphological and phonological properties. As there is no clear-cut distinction between phonological and morphological information (Corbett, 1991), I will use

\textsuperscript{14} All the example languages are taken from Corbett (1991) unless otherwise indicated.
the term *morphophonological*\(^{15}\) to refer to the information provided by the shape or the form of the noun. A well-documented language where the gender of the noun correlates with its shape is Russian. In Russian, masculine nouns follow declensional class I, feminine nouns follow declensional class II and III and neuter nouns follow declensional class IV. As presented in table 2.1 (from Rodina & Westergaard, 2012), nouns in declensional class I end in a non-palatalized consonant and have zero ending (-Ø), nouns in declensional class II end in -\(a\), nouns in declensional class III end in a palatalized consonant and have zero ending (-Ø), and nouns in declensional class IV end in -\(o\). In Spanish and Italian, a great proportion of the nouns ending in -\(o\) are masculine whereas a great proportion of the nouns ending in -\(a\) are feminine. In Arabic (which is the focus of this thesis), nouns ending in -\(a\) are feminine and nouns that lack this ending are masculine (see chapter 3 for a discussion).

Table 2.1. Declension-gender correlation in Russian.

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOM</td>
<td>mal’čik -Ø</td>
<td>devočka -(a)</td>
<td>sol’-Ø</td>
<td>pis’mo -(o)</td>
</tr>
<tr>
<td>ACC</td>
<td>mal’čik -(a)</td>
<td>devočka -(u)</td>
<td>sol’-Ø</td>
<td>pis’mo -(o)</td>
</tr>
<tr>
<td>GEN</td>
<td>mal’čik -(a)</td>
<td>devočka -(i)</td>
<td>sol’-i</td>
<td>pis’mo -(a)</td>
</tr>
<tr>
<td>DAT</td>
<td>mal’čik -(u)</td>
<td>devočka -(e)</td>
<td>sol’-(i)</td>
<td>pis’m-u</td>
</tr>
<tr>
<td>INS</td>
<td>mal’čik -(om)</td>
<td>devočka -(oj)</td>
<td>sol’-(ju)</td>
<td>pis’m-(om)</td>
</tr>
<tr>
<td>LOC</td>
<td>mal’čik -(e)</td>
<td>devočka -(e)</td>
<td>sol’-(i)</td>
<td>pis’m-(e)</td>
</tr>
</tbody>
</table>

\(^{15}\) I will use this term throughout this thesis also because I am not assuming any difference in the way phonological compared to morphological properties affect gender acquisition.
Yet, the noun itself sometimes provides conflicting cues for its gender, i.e. the form of the noun correlates with one gender whereas its meaning correlates with another. To give some examples, in Russian, *djadjja* (uncle) is masculine even though it has a form typical for feminine nouns. In Arabic, *bint* (girl) is feminine even though it has a form typical of a masculine noun. The opposite direction is also found, i.e. where morphophonology takes precedence over semantics. In German, *mädchen* (girl) is neuter even though it denotes a female human, which is typically feminine. Similarly, in French *sentinelle* (sentry) is feminine even though it denotes a male human, which is typically masculine.

Thus, in the endeavour to learn their mother language the language learner/user is equipped with the deterministic information provided by the gender-defining agreeing forms and the probabilistic information provided by the noun-internal (semantic and morphophonological) gender-correlating information. Children were found to be sensitive to the co-occurrences between nouns and articles at a very early age. Using a preferential-looking paradigm, Cyr and Shi (2013) found that French-speaking 20-month-old infants can categorize novel nouns into gender classes relying on their co-occurrence (distribution) with the determiners. They used nouns whose endings do not correlate with a particular gender in French to eliminate any potential phonological cues. Similarly, Van Heugten and Christophe (2015) found that 18-month-old French infants were sensitive to the co-occurrences between nouns and articles using the Visual Fixation
Procedure. Also, Correa and Name (2003) and Correa et al. (2011) found that 2-year-old Portuguese children relied on the gender of the determiners in assigning gender to novel nouns.

In contrast, Gerken et al. (2005) argued through a series of three experiments using a head-turn preference procedure that infants as young as 1;5 seemed to be able to distinguish syntactic category structures by relying on the distributional information. Yet, they could not succeed in the task without the help of the correlated phonological cues. In their experiments, they tried to teach American English-learning infants a small set from the Russian gender paradigm. Similarly, Finley and Wiemers (2015) found that both phonological and semantic consistencies within morphological paradigms facilitate the process of discovering the regularities and eventually learning them by adults. These findings fall in line with the earlier findings by Braine et al. (1990), Brooks et al. (1993), and Frigo and McDonald (1998). As I mentioned above, they all found that learning gender-like classes in artificial language settings was not possible without the help of some intrinsic features that (at least) a subset of the nouns in each class share.

The debate between those who agree on the importance of noun-internal intrinsic features for learning gender-like classes is on what type of features is more important or is predominant, semantic or morphophonological. Both semantic and morphophonological properties were found to be learnable (Culbertson et al., 2019) and were found to be used by children as well as adults in classifying familiar nouns (Gagliardi & Lidz, 2014). It is reported that children become sensitive to the
morphophonological regularities that characterise noun classes at a very early age, i.e. around three (Walter et al., 2021) and that a lack of transparency in those regularities results in error production (Schwartz et al., 2015). On the other hand, sensitivity to semantic cues is reported for six-year-olds only when the semantic cues are deterministic or fully consistent, i.e. without exceptions (Brown et al., 2018; Culbertson et al., 2019).

When those two types of properties (semantic and morphophonological) are in conflict, the general consensus of opinion is that morphophonological properties take precedence or predominate semantic properties. Such behaviour was found to predominate children’s responses up to the age of twelve. Adults on the other hand were found to rely more on the meaning of the nouns when attributing gender (see Karmiloff-Smith, 1979 and Seigneuric et al., 2007 for French; Cain et al., 1987 and Pérez-Pereira, 1991 for Spanish; Szagun et al., 2007 for German; Gagliardi & Lidz, 2014 for Tsez; Culbertson et al., 2019 for artificial language, inter alia). Cain et al. (1987) found that at 12 years, Spanish children used the semantic information of natural gender when assigning gender to novel nouns that have a mismatch between form and meaning.

Contrary to the well-documented morphophonological bias up to age twelve, Rodina and Westergaard (2012) reported an early sensitivity to semantic cues by Russian children when the noun-internal cues are in conflict. However, this sensitivity was not constant among all the different exceptional subclasses of nouns they investigated. They looked at four types of nouns where there is a mismatch
between semantics and morphology in Russian. The first type is *papa*-type nouns which are masculine but have a form typical for feminine nouns. The second type is double gender nouns which have a shape typical for feminine but a mismatch occurs when they refer to males, *e.g.* *plaska* (cry-baby). The third type is female names in *-ok/-ik*, *e.g.* *Lenok* and *Dusik* which exhibit the inflectional paradigm of declensional class I but are feminine because they refer to female humans. The fourth type is hybrid nouns referring to females (hybrid nouns have morphology typical for masculine nouns and a mismatch occurs when they refer to females, *e.g.* *vrac* (doctor)). Their elicited production data showed that with *papa*-type nouns and double gender nouns referring to males, gender was assigned according to the natural gender of the referent not according to the form of the noun at a very early age (2;6 – 4;0). However, in the case of female names in *-ok/-ik*, older children (5 – 6 year olds), similar to adults, followed the natural gender cues but not younger children (2;6 – 4;0). In the case of hybrid nouns referring to females, neither younger children nor older children were sensitive to the semantic cues of the nouns but adults were. Rodina and Westergaard argued that the variation in the patterns of responses they found in their study is due to the variation in the agreement patterns the child encounters in the input language. With the first and the second type of nouns (*i.e.* *papa*-type nouns and double gender nouns referring to males), agreement is consistently semantic as illustrated by the sentences in (8) and (9) respectively. With the third and the fourth types (*i.e.* female names in *-ok/-ik* and hybrid nouns referring to females), there is variation in the agreement patterns the
child is exposed to as the sentences in (10) and (11) illustrate. That is, gender agreement with those two types of nouns either follows the gender suggested by the form of those nouns or the gender suggested by the sex of the referent. In fact, Rodina and Westergaard’s results do not necessarily show that young Russian children have an early sensitivity to semantic cues compared to other languages. It is possible that Russian children learn the gender of the first two types by rote through the consistent agreement patterns they encounter in the input.

8) *naš-Ø*  *pap-a*  *prišel-Ø*
   our(M)  daddy(II)  come.PST.M
   ‘Our daddy came.’

9) a. *naš-Ø*  *plaksa*  (male referent)
   our(M)  cry-baby(II)
   ‘Our cry-baby.’

   b. *naš-a*  *plaksa*  (female referent)
   our-F  cry-baby(II)
   ‘Our cry-baby’

10) a. *Pomniš, Svetik byl-Ø tak-oj malenʹk-ij.*
    remember Svetik(I) be-PST.M such-M little-M
    ‘Remember Svetik was so little!’
b. Pomniš, Svetik byl-a tak-aja malen'k-aja.

remember Svetik(I) be-PST.F such-F little-F

‘Remember, Svetik was so little!’

11) a. naš-Ø vrač prišl-a

our(M) doctor(I) come.PST-F

‘Our doctor came.’

b. naš-a vrač

our-F doctor(I)

‘Our doctor.’

c. naš-Ø vrač

our(M) doctor(I)

‘Our doctor.’

Children’s early sensitivity to morphophonological regularities and their late sensitivity to semantic regularities may have something to do with the interaction between gender assignment principles in gender systems. In the following section, I will consider this potential relation by presenting different proposals that tried to account for the interaction between gender assignment principles.
2.3. Rule hierarchy/ranking

According to Corbett and Fraser (2000), there is always a semantic core in the gender systems. This observation is supported by the existence of purely semantic systems in languages where the gender of a noun can be determined according to its meaning and in the same way, given the gender of the noun, part of its meaning (i.e., the part that determines its gender class) can be predicted. This type of assignment system is found in languages like Godoberi, Zande, Tamil and other Dravidian languages (see Corbett, 1991 for more examples). On the other hand, a gender system that is purely formal (i.e. determined by the phonological and/or the morphophonological properties of the nouns) is not attested in any language.

Thornton (2009) asks, if a language has two assignment systems (semantic and formal), how gender is assigned to the noun if there is a conflict in the gender assigned to it by the different systems. Would the assignment principles be ordered in a particular way and is that order universal or language-specific? She also asks if there exist any constraints on the possible semantic rules for gender assignment. Thornton suggested The Basic Level Hyperonym Constraint which is stated in (12) as a constraint on the possible semantic gender assignment rules. She shows how

---

16 As I mentioned in Chapter 1, gender assignment in this section refers to the rules that have been proposed to account for the distribution of gender across nouns in language systems; i.e., gender assignment in grammar. The implication of these proposals for gender assignment in the acquisition of gender will also be discussed.
this constraint explains the assignment of feminine gender to names of cars and names of cities in Italian and also to names of cars in Hausa.

12) The Basic Level Hyperonym Constraint

“To be able to assign gender to its hyponyms, a hyperonym must be a basic level term.” (p. 27)

To explain this constraint, we need to consider the three levels for classifying physical objects given in Rosch et al. (1976) (table 2.2). Physical objects comprise three levels, a superordinate level such as clothing, a basic level such as socks, and a subordinate level such as knee socks. According to Thornton’s Basic Level Hyperonym Constraint, a noun inherits the gender of its hyperonym if and only if the hyperonym is a basic level noun. Thus, names of cars in Italian such as cabrio and torpedo are assigned feminine gender because the basic level noun macchina (car) is feminine. Thornton offered the same explanation for the names of cars in Hausa and also for other categories of nouns in Italian. Thornton’s Basic Level Hyperonym Constraint also offers an explanation to the assignment of feminine gender to names of cities and the assignment of masculine gender to names of letters in Arabic. To explain, the noun madīna (city) which is a basic level noun is feminine in Arabic and therefore its hyponyms, such as ʿalriyāḍ (Riyadh) dimašq (Damascus), ʿalqāhira (Cairo) and baĝdād (Bagdad), are all feminine. The same applies for the names of the Arabic letters. All the names of the Arabic letters,
such as ‘alif (A), bā’ (B), and tā’ (T), are masculine because their basic level hyperonym harf (letter) is masculine.

Table 2.2. Hierarchies of physical objects investigated by Rosch et al. (1976).

<table>
<thead>
<tr>
<th>Superordinate level</th>
<th>Basic level</th>
<th>Subordinate level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Musical instrument</td>
<td>Guitar, piano, drum…</td>
<td>Classical guitar, upright piano…</td>
</tr>
<tr>
<td>Tool</td>
<td>Hammer, saw, Claw hammer, cross-cutting screwdriver…</td>
<td>hand saw…</td>
</tr>
<tr>
<td>Clothing</td>
<td>Pants, socks, shirts…</td>
<td>Levis, knee socks…</td>
</tr>
<tr>
<td>Furniture</td>
<td>Table, lamp, chair…</td>
<td>Kitchen table, desk lamp…</td>
</tr>
<tr>
<td>Vehicle</td>
<td>Car, bus, truck…</td>
<td>Sports, car, city bus…</td>
</tr>
</tbody>
</table>

As discussed in Thornton (2009), different proposals have been offered regarding the ordering of gender assignment rules and for dealing with cases of conflicts between two genders. Steinmetz (1986) proposed two principles to deal with gender assignment in German: Gender Tally (13) and Gender Eclipsis (14) in addition to three types of gender assignment rules (M-rules, S-rules, and SC-rules). M-rules assign gender on the basis of the shape of the nouns, S-rules assign gender on the basis of the meaning (semantics) of the noun, and SC-rules are “subcategorization rules, which apply within the domain of certain S-rules and
assign gender either on the basis of morphological/phonetic shape or on the basis of meaning” (p. 191).

13) **Gender Tally**

“This principle can be expressed as an instruction: Count the number of times each gender is assigned\(^1\) and assign the noun the gender with the highest value.

14) **Gender Eclipsis**

If two or more genders have the same value, gender is assigned by the principle $m > f > n$, i.e. masculine eclipses feminine and feminine eclipses neuter.” (Steinmetz, 1986, p. 193)

To explain how Steinmetz’s proposal works, a noun can be thought of as having a gender table with all the potential rules that could assign gender to this noun. The winning gender is the gender assigned by the highest number of the rules as the principle of **Gender Tally** states. In case two or more genders have the same value, the **Gender Eclipsis** principle comes to work and assigns the noun the least marked gender. For example, the Frucht (fruit) gender table has the two assignment rules in (15), with the first assigning feminine gender and the second assigning

\(^{17}\)“The number of times each gender is assigned” here means the number of rules that assign a particular gender to the noun.
neuter gender. *Frucht* is assigned feminine gender according to *Gender Eclipsis* because neuter gender is more marked in German than feminine gender.

15) *Frucht*

\[-ucht = f\]

\[\text{Superordinate} = n\]

\[0m, 1f, 1n = f\]

Rice’s (2006) *optimal gender assignment theory* maintains that, universally, there is no ranking of gender assignment rules. Yet, according to Rice, the gender classes themselves are ranked according to their markedness in a particular language, and in cases of conflict, nouns are assigned the least marked gender which is very similar to Steinmetz’s *Gender Eclipsis* (14). Rice goes further by suggesting a distinction between *balanced* and *imbalanced conflicts*. A *balanced conflict* occurs when two rules, each suggesting a different gender, are in conflict with each other. For example, in German *Gemeinde* (congregation) begins with *Ge-* which suggests neuter gender but ends in *-e* which suggests feminine gender. In this case, *Gemeinde* is assigned feminine gender because in German neuter is more marked than feminine and therefore the noun is assigned the least marked gender. *Imbalanced conflict* on the other hand occurs when more than two rules are in conflict with each other, with two rules suggesting one gender while the third rule suggesting another gender. In this case, the noun is assigned the gender suggested
by the majority of the rules even if the other gender is less marked, which is very similar to Steinmetz’s *Gender Tally* (13). For example, *Gemüse* (vegetable) starts with *Ge-* , ends in *-e*, and denotes a superordinate (which correlates with neuter in German). In this case, two rules suggest neuter gender against one rule that suggests feminine gender. *Gemüse* is neuter because the majority of the rules win. Thus, Rice claims that different assignment rules whether semantic or formal are not ranked with respect to each other, “semantically-based gender rules as a group do not consistently beat shape based-rules. Shape-based rules do not consistently beat semantically-based rules. Ergo, there is no group-wise hierarchical ranking of these kinds of rules or constraints” (Rice, 2006, p. 1413).

Nesset (2006) discussed gender assignment from a usage-based perspective (Langacker, 1991, 1999) and he argued that looking at gender from this perspective makes the principles proposed by others, such as Steinmetz’s (1986) *Gender Tally* (13), follow naturally and for free. In his proposal two important factors suggested by Langacker are of crucial role, namely inherent ease of activation, and conceptual overlap. Conceptual overlap means that if more than one candidate is competing in the grammar, the higher the degree of conceptual overlap the candidate has with a schema, the higher the chance for this candidate to survive in the competition. Inherent ease of activation means that a schema that covers a large class of items is more easily activated than a schema that covers a small class of items “therefore, a candidate that is connected to the schema of a large class has an advantage in the competition with other schemas” (Nesset, 2006, p. 1373). This is related to type
frequency. Nesset suggests Fig. 2.1\textsuperscript{18} to explain how gender is assigned to the Norwegian word ‘\textit{bonde}’. This example is of particular interest because it illustrates Nesset’s proposal in the case of competing schemas. In Norwegian, nouns that end with umlaut (represented by Ö) in the plural are feminine (schema c). Similarly, nouns that end in -\textit{e} are also feminine in Norwegian (schema d). However, those two schemas are not without exceptions as represented in schemas a and b. Schema b specifies that when a noun ends in -\textit{e} and refers to a person, it is masculine in Norwegian. Though inherent ease of activation would assign feminine gender to \textit{bonde} because schemas c and d are more entrenched in the grammar because they represent nouns of high type frequency, conceptual overlap would prefer masculine gender because schema b is more specific and contributes more information than schema d. According to Nesset “This suggests that the factor of conceptual overlap takes precedence in rule interaction when it conflicts with inherent ease of activation” (p. 1374). “Inherent ease of activation only comes into play when degree of conceptual overlap does not facilitate the selection of a unique candidate” (p. 1376).

\textsuperscript{18} Nesset used full-drawn boxes to represent masculine gender and dashed boxes to represent feminine gender. He also used rectangular boxes to represent a fragment of the grammar of the language in question. Nesset also used boxes with thicker lines to represent strong schemas (ones associated with high type frequency) and thinner lines to represent schemas associated with lower type frequency.
Another advantage of looking at gender assignment from a usage-based model according to Nesset is that not only Gender Tally follows without stipulations but also Kiparsky’s (1973) Elsewhere Condition. Looking back at the example in Fig. 2.1, masculine gender is assigned to bonde because schema (b) is more specific, i.e. because it has a higher degree of conceptual overlap than schema (d) which is the same result that the elsewhere condition would predict. According to Nesset, the elsewhere condition follows naturally from the conceptual overlap factor.

Nesset further suggests that the principles that account for the interaction between different schemas (formal principles) must be supplemented by what he calls substantial ordering principles, “i.e. principles referring to particular types of
information, not to the relationships between rules or schemas” (p.1385). As an example of this latter type of principles, he presented his *Core Semantic Override Principle* (16). The Core Semantic Override Principle says that if a schema contains a particular type of information, i.e. reference to biological sex, it takes precedence over other schemas.

16) *Core Semantic Override Principle*

“Rules referring to biological sex take precedence in gender assignment.”
(Nesset, 2006, p. 1386)

If those proposals were to account for how the gender systems of different languages work, we would expect that the end state of the gender system as represented in the system of the adult speakers/users of the language would be sensitive to the proposed principles and constraints when attributing gender to nouns in order for the noun to function as the controller of the syntactic agreement with other elements. It does not seem a difficult task for someone with some linguistic background to come up with contradicting examples for each of the aforementioned accounts19. To give some examples, German *mädchen* (girl) is against Nesset’s *Core Semantic Override Principle* because the biological sex information of this noun does not take precedence in assigning gender to it. Similarly, Arabic *bint* (girl) is a counter example for Rice’s *optimal gender*

19 See Enger (2009), Thornton (2009) for such examples.
assignment theory because it provides a case of “balanced conflict with a mismatch between feature type (meaning or shape) and category markedness…… in which the shape correlates with a less marked category while the meaning correlates with a more marked category.” (Rice, 2006, p. 1402). To explain, bint denotes a female human which suggests feminine gender but it does not end in -a which suggests masculine gender. According to Rice, bint should be (wrongly) assigned masculine gender because in Arabic, as will be discussed in the next chapter, masculine is the least marked gender.

The Core Semantic Override Principle does not seem to make as strong claims about the predominance of semantic assignment rules over morphophonological rules as claimed in Corbett and Fraser (2000), “as is universally the case, the formal gender assignment rules […] are dominated by the semantic gender assignment rules” (p. 321). Though it might be tempting to say that this principle can be rejected on the basis of some counter examples such as German mädchen (girl), the cases where sex-based gender assignment is overridden “are limited and well-defined” (Nesset, 2006, p. 1385). Dahl (2000) discussed those cases and how they are associated with upgrading or downgrading purposes in most situations.

By and large, the discussed proposals make different predictions about the end state of the grammatical gender system as executed by the adult speakers/users of the language. It can be predicted from Rice’s account, which is very similar to Steinmetz’s account, that adults would not show preference for a particular type of
cues to gender, whether semantic or morphophonological when the cues provided by the noun itself conflict. That is, in the cases of balanced conflicts, the assignment of the least marked gender, which is indeed language-specific, is expected to prevail no matter what the noun refers to. Thus, in a language where masculine gender is considered the least marked gender, it is expected that masculine is assigned to nouns with balanced conflicting cues whether they refer to animates or inanimates or correlate with other semantic assignment rules in the language. It is clear from the literature discussed in the previous section that neither the child system nor the adult system could work according to Rice’s proposal. In the cases of conflict, children were found to assign gender according to the morphophonological properties of the nouns and adults were found to assign gender according to the semantic properties not according to the markedness hierarchy suggested by Rice.

Nesset’s account, on the other hand, makes a different prediction about the way adult users of the language are expected to behave in the cases of contradicting cues. It predicts that different behaviours would be manifested depending on the semantic properties of the noun in question. That is, when the semantic properties of a noun are sex-based, i.e. refer to the distinction between males and females, it is anticipated that adult speakers will predominantly assign gender according to the meaning of the noun, not according to its form, which is supported by the findings from different languages. But, when the meaning of the noun is not sex-based, gender assignment by adult speakers would depend on other criteria such as the inherent ease of activation and the conceptual overlap. Moreover, it can be
concluded from the review of the different studies in the previous section that child language learners were found not to be sensitive to Nesset’s principle. Yet, such a principle is very important to explain how genders typically pattern in languages like Arabic and Russian for instance. Therefore, the question that remains unanswered by the reviewed studies and that the current thesis attempts to answer is when do language learners become sensitive to such principle if the system of the language works according to it. I will go back to this question when discussing the results of the three experiments in chapter 7.

Related to rule hierarchy or rule ordering in gender assignment is the notion of default. The default gender is the least specified gender and the gender supplied by the elsewhere principle unless other more specified genders are required. I will discuss the different notions of default and the relation of default to gender acquisition in the next section.

2.4. The notion of default

Rice’s and Steinmetz’s proposals discussed in the previous section rely on markedness hierarchies as one of their cornerstones. Though it was not made clear in their proposals on what basis the markedness hierarchies of different languages are constructed, it seems that the least marked gender in their view is the gender with more members. That is, the markedness hierarchy depends on type frequency. For example, in German, masculine gender has more members than feminine gender and feminine gender has more members than neuter gender. Therefore,
masculine is considered the least marked gender while neuter is considered the most marked gender. Markedness hierarchies that depend on frequency counts would not fare well in languages where the least marked gender does not correspond to the gender with more members as I will show in the case of Dutch below. The term markedness can mean many things (Haspelmath, 2006 distinguished twelve senses in which markedness has been used in 20th-century linguistics). In the following, I will introduce the notion of default which is very related to the term unmarked. The two terms, default and unmarked, are sometimes used interchangeably by some linguists.

The notion of default is very central in morphological theory and description (Gisborne & Hipsley, 2017). Default has been used in the linguistic literature to refer to different meanings. The first is default as synonymous to the ‘regular’ form, i.e. the form supplied by a rule unless other ‘irregular’ or exceptional forms are available. This is very common in the literature on English past tense, as in Embick and Marantz (2005, p. 244): “In the case of Regular Verbs, the [past] feature is supplied, or spelled-out, with the phonological signal -ed, the default [emphasis added] instantiation of [past]”. It is also found in the debate about German plural as in Dąbrowska (2004, p. 125): “…German-speaking adults and children learning German know that -s is the default [emphasis added] ending and therefore apply it whenever a plural form cannot be retrieved from memory”. Marcus et al. (1995) listed a number of circumstances where the default, i.e. the regular inflection, is applied. This notion of default as correspondence to the regular form is sometimes
used interchangeably with *default* as the most frequent form. This introduces the second sense for the notion of *default* in the linguistic literature, namely default as the most frequent form. Though, for instance, the English default plural morpheme 
*-s* happens to be the most frequent one, frequency is not necessarily a defining feature of default as is the case in German plural. Default as the most frequent form is synonymous to Rice’s and Steinmetz’s ‘unmarked’. A third and a fourth meaning for the notion of *default* are clear in McFadden’s (2007, p. 231) discussion about the default case: “**Default** [emphasis added] case is not the case assigned when other cases fail, but the actual lack of case. Now, this is not meant in the morphological sense that **default** [emphasis added] case is the lack of an overt case-marker.” McFadden mentioned two different ways in which the notion of *default* is used. One of them is default as the ‘underspecified’ form or the form that lacks any feature specifications. The other is default as the ‘basic’ form or the form with no morphological marking. In some situations, McFadden’s two meanings correspond to each other as I will show for Arabic gender. A fifth sense of *default* is default as the form used in cases of ‘resolution’, which is seen in the agreement with coordinated nouns that have a mismatch in feature specification. For example, when a predicative adjective has to agree in gender with a coordination of a masculine and a feminine noun. Still another sense of *default* corresponds to the form used with loan words when they share no semantic or formal properties with other words in the host language as in the case of assigning masculine gender to new words coming into French (Ayres-Bennett & Carruthers, 2001; Violin-Wigent, 2006).
One of the uses of default in morphology stems from Kiparsky’s (1973) *Elsewhere Principle* in phonology which, simply put, states that a general rule (feature/form in morphology) applies unless a more specified rule (feature/form) is available. As I mentioned in the preceding chapter, Neeleman and Weerman (1999) argued that the default form, which is the feature-underspecified form, is acquired earlier than the non-default counterparts. This happens before children learn or discover the feature specifications of the other more specified forms. There is converging evidence for example that children overgeneralize nominative case before they learn the feature specifications required for producing other cases such as accusative and dative. This however could be due to the fact that nominative, in some languages, is the base (morphophonologically unmarked) form and therefore acquired before inflected/derived forms.

As I stressed in the preceding chapter, *default gender is nothing but default agreement*. Especially when studying the acquisition of gender, it is difficult, and may be impossible, to tease apart gender assignment from gender agreement. Because of the non-canonicity of gender as a morphosyntactic feature (Corbett, 2012), i.e. that every noun has a single inherent gender feature that has to be learned as part of the noun, the only thing that tells us what gender a child assigns to a noun are the agreeing elements that they use with it. Ergo, if the default forms in a morphological paradigm are acquired earlier than more specified forms (Neeleman and Weerman, 1999), it is predicted that before children acquire the feature specifications required for producing specified gender agreeing elements, they will
tend to produce the agreement elements that lack gender specifications. This prediction was confirmed by findings from different languages.

In his study of the acquisition of Arabic gender by English, French, and Japanese L2 learners of Arabic, Alhawary (2009) found that all the participants did better with the masculine gender than with the feminine gender. He looked at longitudinal as well as cross-sectional data. He also studied demonstrative-noun agreement, verbal agreement, and noun-adjective agreement. The same pattern was found across the different agreement targets. That is L2 learners of Arabic were more successful in producing masculine agreement than in producing feminine agreement. This could be due to the masculine being the default gender according to Alhawary. The same patterns were documented for other L2 learners of Arabic and using different tasks (See Alamry, 2014; Aljadani, 2019, inter alia).

In Karmiloff-Smith’s (1979) study on French, children older than 9 years were found to attribute masculine gender to all novel words. Pérez-Pereira (1991) also found that Spanish children tended to attribute masculine gender more often than feminine gender because according to him, masculine gender is the unmarked gender in Spanish (Harris (1991) also argued that masculine is the morphologically unmarked gender). Boloh and Ibernon (2010, 2013) and Boloh et al. (2012) argued that the role of morphophonological ending in acquiring French gender has been overstated by Karmiloff-Smith. They claimed that what French speakers seem to do is acquire masculine gender as a default and gradually learn the association between feminine gender and feminine endings. They argued that children do not
need to bother learning dozens of endings and their association with gender because
nouns almost always appear in the context marked for gender by means of
determiners. “Masculine” nouns are nothing but nouns that must co-occur with a
given set of mutually predictive elements, namely the “masculine” determiners.”
(p.3). According to them, sensitivity to word endings might appear as a by-product
of gender learning. One of the reasons they gave for the argument that masculine
gender and masculine suffixes might be learned faster and eventually contribute to
having masculine as default in French is that “masculine nouns might benefit from
a clearer within-similarity space than feminine nouns. The number of vocalic
ending types is far below the number of consonantal ending types and vocalic
endings showing a gender bias of 0.90 at least in the P&S input\(^{20}\) are always biased
towards the masculine. By contrast, consonant endings with 0.90 predictive values
are sometimes biased towards the feminine gender and sometimes towards the
masculine gender. Thus, similarities in the vocalic/masculine case do not even seem
to require that differences between vocalic endings types be necessarily processed,
which could not be dispensed with in the consonant/feminine case.” (Boloh &

Children (4–10 years) and adults participated in Boloh and Ibernon’s (2010)
study. Novel nouns with suffixes that had a high predictive value of gender were

\(^{20}\) P&S corresponds to Philippe’s lexical input and Stephane’s lexical input. It is a database created by the authors by combining the verbal production of two French children, Philippe and Stephane. The verbal production was recorded from age 1;1 to 3;3 for Philippe and from 2;3 to 4;9 for Stephane. Both corpora are available through the Child Language Data Exchange System (CHILDES, MacWhinney, 2000).
presented as 3D artificial objects. The novel nouns were presented with no accompanying articles and the gender attributed to each noun was determined by the form of the article produced by the participants. They found that in the three experiments, gender attribution to masculine nouns was at ceiling level in all the age groups with no difference between the different suffixes. However, gender attribution to feminine suffixes was at or below chance level with significant differences between the different suffixes.

In Boloh et al. (2012), the novel nouns (referring to inanimates) were presented with an indefinite article the gender of which contradicts the gender suggested by the ending. They found that children between 4 and 10 years predominantly attributed gender according to the determiner not according to the endings. In Boloh and Ibernon (2013), children (4-12), adolescents (14-15), and adults (17 years old) were presented with novel nouns with contradicting information, i.e. the suffix of the noun suggested one gender whereas the natural gender indications of the puppets suggested the opposite gender. They found that gender-suffix congruent attribution by the children was at ceiling level when the suffixes were masculine but at chance level when the suffixes were feminine. Adolescents and adults attributed gender according to the natural gender of the puppets.

The tasks used by Boloh and Ibernon (2010, 2013) and Boloh et al. (2012) were very complicated compared to the tasks used by Karmiloff-Smith. In Karmiloff-Smith’s five tasks, the participants were simply presented with two
identical drawings which differ in their colours. One of the drawings was then hidden in order to elicit some gender agreeing forms that describe that hidden drawing. Boloh and Ibernon and Boloh et al. however, provided the participants with sets of three different coloured objects and three bags with different sizes for which the participants had to describe the colour of the object in addition to the size of the bag in a storytelling setting, though what was counted in the results was the form of the determiner produced by the children. All that could have added to the processing load (as also noted in Kerkhoff, 2013). It is also not clear how and why children would start by setting masculine as a default gender while learning feminine nouns item by item. If, as they mentioned, French children do not need to bother learning dozens of suffixes and their association with gender because the nouns in most cases are used in combination with a gender marked determiner which itself indicates the gender of the noun, feminine and masculine nouns are expected to proceed in similar fashion until a particular stage where children might decide on using masculine as a default depending on other cues like the form of the adjectives or coordinated nouns.

To sum up, though looking at the role of default in gender acquisition sheds light on the order of acquisition of the gender classes, there is no clear definition of what constitutes a default gender in the aforementioned studies. If as argued by Neeleman and Weerman (1999), the forms underspecified for features are acquired earlier by children, why in some languages is this not the case? For instance, the neuter *het* is the linguistic default in Dutch, i.e. the form that is unspecified for
gender features. Therefore, *het* is the determiner used with nominalized categories and also as expletive subject pronoun (Roodenburg & Hulk, 2008, 2010). Yet, as argued by Tsimpli and Hulk (2013), *de* seems to be the learner’s default in Dutch up until six years of age because *de* is more frequent in the input than *het*. The frequency of a particular gender in the language the child is exposed to might also be affected by the biological gender of the child themselves. A boy might be exposed to more masculine agreeing forms in the language addressed to him while a girl might be exposed to more feminine agreeing forms in the language addressed to her. Nonetheless, *default gender* cannot simply mean the form that is acquired earlier by the learner. Rather *default gender* is nothing but *default agreement*, i.e. gender agreeing forms that lack feature specifications. In other words, default gender is the gender that requires the default agreeing forms, i.e. the agreeing forms underspecified for gender features. Whether the *default gender* is acquired earlier or not can be language-specific. Other properties of the language and its gender paradigm may play a role in the order of acquisition. The frequency of the default gender and its morphological marking/lack of marking are among those properties. A perfect language for testing the prediction that the default gender is acquired earlier than the non-default gender(s) would be a language where the default gender is not the most frequent gender and where the default gender is morphologically marked. Such a language would allow us to control the potential confound of frequency and morphological marking which are plausibly related to the order of acquisition of a morphological paradigm.
2.5. Chapter Summary

This review has shown that there is a general consensus on two issues regarding the type of cues children are sensitive to in the early years of language development. First, children appear to be biased to rely on morphophonological cues when attributing gender to nouns especially in the case of conflicting cues. Second, children will be more sensitive to semantic cues from an early age if this type of cue is consistent in the input.

In this chapter, I summarised the studies that looked at the acquisition of gender by L1 learners and occasionally referred to adult L2 learners. Those studies show that all the types of cues to grammatical gender are consulted by language learners when assigning gender to nouns though in different degrees. Those studies also show converging evidence that children unlike adults are biased to assign gender according to the form of the noun rather than according to its meaning, particularly in the case of contradicting cues.

As can be seen from this review, many of the studies, with very few exceptions, investigated Romance and Germanic languages with obligatory determiners. The obligatoriness of the determiners in languages like French and Spanish makes learning determiner + noun combinations as chunks highly possible, especially in the first stage of language acquisition which is characterized by rote learning. Ergo, testing the acquisition of gender through determiners does not seem to be the best method especially when familiar nouns are used. Unless novel nouns
are used, evidence of the early acquisition of gender that is based only on the use of determiners seems to be far from conclusive. Needless to say, that the way determiner + noun combinations are initially learned is the topic of an ongoing debate between nativists and constructivists (see Pine et al., 2013 and Valian et al., 2009 inter alia).

The different accounts proposed for the ordering of the assignment principles make different predictions about the end state of the gender system as manifested in the adults’ system. Those predictions depend on whether the account assumes a type-wise ordering of the assignment principles or not.

Finally, the last section in this chapter presented the notion of default in morphology. It is clear from the discussion in this section that the use of default in morphology is a thorny issue. One needs a clear definition of what default gender is in order to understand the order in which the gender paradigm of a particular language is acquired.
Chapter 3: The Arabic gender system

3.1. Introduction

In this chapter, I will present the background information necessary for understanding the Arabic gender system. I will start by defining what I mean by Arabic in this thesis (section 3.2). Section 3.3 will present the gender system of Arabic, highlighting the discrepancies between Hasawi Arabic (the parents’ input language that children are exposed to at home) and Standard Arabic (the formal system children are introduced to in school). The principles of assigning gender to singular nouns will be presented in section 3.3.1 and section 3.3.2 will demonstrate how gender manifests itself in gender agreement between the noun and other elements syntactically related to it. Section 3.4 will present the distribution of the gender classes in Arabic as well as the distribution of the noun-internal gender correlating information that could be used by the learner to facilitate their mission in acquiring the gender system (Braine, 1987; Braine et al., 1990; Brooks et al., 1993) as revealed from the analysis of four different corpora. Section 3.5 will present four pieces of evidence in support of the status of masculine agreement as the default agreement in (Hasawi) Arabic. Finally, section 3.6 will conclude this chapter.

3.2. Definition

Before presenting the gender system of Arabic, I need first to state clearly what I mean by Arabic in this study. A distinction needs to be made between two poles of
Arabic: Classical/Standard Arabic21 ‘fuṣḥā’ in one end and spoken or vernacular Arabic ‘dārija’ in the other end. Standard Arabic is used in this thesis to refer to the system of grammatical rules established by the Arab grammarians that should be followed by anyone who writes or delivers speech in formal contexts (Gothenburg, 2011). Spoken Hasawi Arabic (HA) on the other hand refers to the spoken variety of Arabic in the Alhasa22 province in the eastern part of Saudi Arabia23 where all the participants in this study reside and were borne. Therefore, when I use Arabic with no modification (standard or spoken), I imply that spoken Arabic (at least the variety under investigation in this project) does not differ substantially from SA with regard to the issue being discussed.

It is important to keep in mind this distinction between SA and spoken Arabic especially when looking at the acquisition of Arabic because as stated by Gothenburg (2011, p. 783), “an important characteristic of the Arabiyya/Classical Arabic/Modern Standard Arabic is that it is not the mother tongue of anyone. Furthermore, its structure differs considerably from all spoken varieties in the

21 The distinction between classical Arabic (CA) and Modern Standard Arabic (MSA) as in Badawi (1985) or what Haeri (2003) calls “contemporary Classical Arabic” and “old Classical Arabic” is irrelevant for the discussion here. The differences between the two are only stylistic and lexical. The rules of morphology and syntax are very similar in both CA and MSA. They are both referred to as fuṣḥā in Arabic (Ryding, 2011). Therefore, I will use Standard Arabic (SA) ‘fuṣḥā’ to refer to both throughout this thesis in order to distinguish the ‘formal’ language from the spoken vernacular (or dialect) under investigation.

22 Notice that the name of the province is interchangeably used in the literature as Al-Hasa or Al-Ahşah. Throughout this thesis, I will use Alhasa version for consistency’s sake.

23 For a background in the sociohistorical context of AlHasa, see Smeaton (1973). Also see Abdelatti et al. (2017) For the geographical set up of AlHasa.
modern Arab countries”. Thus, SA is not the input language for any child acquiring Arabic as a mother tongue. It is not the language spoken by Arab persons in their everyday life and in their everyday communication. What children receive as the input language from their parents, caregivers, and the people around them is the spoken vernacular. This spoken vernacular varies not only from one country to another as we find between Saudi Arabic, Egyptian Arabic, and Moroccan Arabic for example, but also from one region to another within the same country as between Hasawi Arabic and Hijazi Arabic, both of which are dialects spoken in Saudi Arabia.

Speaking of Hasawi children (and the Saudi children in general), their real exposure to SA begins when they start primary school around age six. The main focus in the first three years (six to eight years) of primary schools in Saudi Arabia is developing the literacy skills of the children. Children are also introduced to some grammatical patterns in those years that they have to imitate and copy (Siddig, 2022). The explicit teaching of the grammatical rules of SA starts in primary 4, i.e. around the age of nine.

In the following sections I will present the gender system of Arabic highlighting the discrepancies between Standard Arabic and Hasawi Arabic.
3.3. The Arabic gender system

3.3.1. Gender assignment in Arabic

Arabic, like many Romance languages such as French, Spanish, Portuguese, and Italian, has a two-gender system; with every noun being either masculine or feminine. Animate as well as inanimate nouns are assigned gender in Arabic. Masculine is the morphophonologically unmarked gender whereas feminine is the morphophonologically marked gender as the contrasts between masculine and feminine nouns in (17) illustrate. Feminine nouns usually end in one of the three feminine morphophonological endings -a(t)\(^{24}\) as in the feminine examples in (17), -ā’ as in saḥr-ā’\(^ {25}\) (desert), and -ā as in ḥalwā (sweet). Those endings are replaced by a single feminine morpheme, i.e. /-a/ in most Arabic dialects. In addition, -a(t) is more common than the other two endings and it is the ending found in the group of feminine nouns derived from their masculine counterparts by means of suffixation (Hachimi, 2007).

---

\(^{24}\) \(t\) is placed between parentheses here because it is not realized in the pausal forms neither in Hasawi Arabic nor in Standard Arabic. The only context in which \(t\) is realized is the Construct State. To explain, when the noun is the first part of a Construct State as in the example below, \(t\) is realized.

- **kātib-at ‘al-riwāya**  
  writer-F the-novel(F)  
  ‘The writer of the novel/ the novel writer.’

\(^{25}\) In this chapter and throughout this thesis, I will transcribe the Arabic words in their pausal forms; i.e. without the short vowels that indicate the case marking. Including case adds nothing for the purpose of this study. I will follow, to the best possible extent, the scientific transcription system of the Encyclopedia of Arabic Language and Linguistics (EALL) when transcribing SA examples. As “there is no formal ‘standard’ for dialectological transcription of Arabic” (Reichmuth, 2011), I will try to transcribe HA examples as accurately as possible following the same system of EALL where applicable.
When it comes to gender assignment to singular nouns in Arabic, Standard Arabic (see Alhawary, 2009; Cowan, 1958; Ryding, 2005, *inter alia*) and all spoken varieties of Arabic, including Hasawi Arabic26 (see Cowell, 1964 for Syrian Arabic; Hachimi, 2001 for Moroccan Arabic; Ingham, 1994 for Najdi Arabic; Omar, 1973 for Egyptian Arabic, *inter alia*) agree on the following two points:

- all nouns denoting single human males are grammatically masculine, and
- all nouns denoting single human females are grammatically feminine.

Those two generalizations work perfectly in Arabic with no exceptions no matter what morphophonological ending the noun has. To illustrate, in (18)27 below, the nouns on the left are all masculine though some of them end in -

---

26 As there is no scholarly reference grammar for Hasawi Arabic, I, as a native speaker of this variety who was born of two Hasawi parents and lived in Alhasa for my whole life, will rely on my own judgement when comparing the Hasawi dialect to SA.

27 The nouns in (18) are given in the HA version only for ease of presentation. Notice however that the grammatical gender as well as the endings of those nouns are not different from the SA versions.
While the nouns on the right are all feminine though some of them lack the aforementioned feminine markers.

18) **HA**

<table>
<thead>
<tr>
<th>Arabic</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>walad</em></td>
<td>(a boy)</td>
</tr>
<tr>
<td><em>rajjāl</em></td>
<td>(a man)</td>
</tr>
<tr>
<td><em>xalīfā</em></td>
<td>(a caliph)</td>
</tr>
<tr>
<td><em>ʾallāma</em></td>
<td>(a savant)</td>
</tr>
<tr>
<td><em>ʿābu</em></td>
<td>(a father)</td>
</tr>
<tr>
<td><em>bint</em></td>
<td>(a girl)</td>
</tr>
<tr>
<td><em>marah</em></td>
<td>(a woman)</td>
</tr>
<tr>
<td><em>ʿarūs</em></td>
<td>(a bride)</td>
</tr>
<tr>
<td><em>ʾixt</em></td>
<td>(a sister)</td>
</tr>
<tr>
<td><em>ʿumm</em></td>
<td>(a mother)</td>
</tr>
</tbody>
</table>

One comment is due here, namely that a great majority of the female-denoting nouns in Arabic are morphophonologically marked with the -a ending like the feminine nouns in (17) above. So, a case in which the semantic properties of the noun align with its morphophonological properties is not uncommon in Arabic.

In addition to the two generalizations that account for the assignment of gender to singular human-denoting nouns, the two generalizations below account for the vast majority of non-human-denoting nouns in Arabic:

---

28 Henceforth, I will use -a, which is the pausal form of all the feminine morphophonological markers, when discussing the endings of the feminine nouns in Arabic.

29 Instead of stating, in the main text, the variety of Arabic (SA vs HA) for each example and instead of giving examples from both varieties for each point (unless an area of discrepancy exists); henceforth, I will mention the variety next to the number of each example.
• almost all singular nonhumans that are morphophonologically marked are grammatically feminine, and

• the vast majority of the singular morphologically unmarked nonhuman nouns are grammatically masculine.

In order to illustrate these two generalizations, I will focus on SA and HA, which both adhere to them. That is, when a non-human noun ends in one of the feminine morphophonological markers, it is assigned feminine gender in both varieties, like all the nouns in (19). Similarly, when a noun lacks a feminine morphophonological marker, like the ones in (20), it is assigned masculine gender in the vast majority of the non-human nouns in both varieties.

19) SA

ḥadīqa (a garden)
ṭāwila (a table)
tuffāḥa (an apple)
madrasa (a school)
kura (a ball)
20) **SA**

- *kitāb* (a book)
- *kursi* (a chair)
- *qalam* (a pen)
- *bāb* (a door)
- *masjid* (a mosque)

A small number of the non-human denoting nouns that lack the feminine morphophonological markers are always feminine in Arabic, such as the ones in (21).

21) **HA**

- *šams* (a sun)
- *ʿein* (an eye)
- *yad* (a hand)
- *ʿarḍ* (an earth, a ground, a land)
- *nafs* (a self, a soul)

Another small subset of the nouns that lack the feminine morphophonological markers, such as the ones in (22), can be used with both genders in SA. They are *dual gender nouns* (Ryding, 2005). Yet, this subset of *dual gender nouns* has shifted gender to masculine in many Arabic
In HA for example, the nouns in (22) are heard with masculine agreement almost always. This reflects a tendency toward uniformity among nonhuman nouns; the nouns that have the -a marker are feminine whereas the nouns that lack such marker are masculine in HA. Notice, however, that this tendency does not have the same degree of effect on all the unmarked non-human nouns. That is, the gender of nouns like the ones in (21) is more fixed than the ones in (22).

22) **SA**

\[
\begin{aligned}
\text{sūq} & \quad \text{(a market)} \\
\text{tarīq} & \quad \text{(a road, a path)} \\
\text{ḥāl} & \quad \text{(a condition)} \\
\text{lisān} & \quad \text{(a condition)}
\end{aligned}
\]

To summarise, semantic and morphophonological properties of nouns in Arabic interact with each other in determining the different gender classes nouns belong to. One exceptionless generalization in Arabic is that nouns that refer to male humans are grammatically masculine and nouns that refer to female humans are grammatically feminine. Epicene nouns, which are masculine in Arabic, such as the ones given in (23) are not against this perfect correlation between natural

---

30 See Procházka (2004) for an investigation of the unmarked feminine nouns in Arabic dialects which revealed that the gender of some of those nouns is fixed among the Arabic dialects while the gender of others is very diverse and dialect-specific. I will not go into details of this issue and I will focus on how those nouns are used in HA compared to SA.
gender and grammatical gender in Arabic because whenever those nouns are used, they have generic meaning where the sex of the referent is not important as the sentence in (24) illustrates. Among the non-human nouns in Arabic, there is a robust correlation between noun ending and grammatical gender. That is, when a nonhuman noun ends in –a, it is almost always feminine whereas when a nonhuman noun lacks the –a ending, it is masculine most of the time with the exception of a small number of nouns, such as the examples in (21).

23) SA/HA

a. šahš (a person)
b. fard (an individual)
c. 'insaan (a human-being)

24) HA

\[
\begin{array}{ccc}
\text{ya-jgib-ni} & 'il-šahš & 'il-mibtisim \\
3\text{MS-attract-me} & \text{the-person(M)} & \text{the-smiling(M)}
\end{array}
\]

‘The smiley person attracts me.’

Nonetheless, the way in which semantics and morphophonology interacts is hierarchical: semantics always overrides morphophonology in Arabic. The grammatical gender of nouns referring to humans is determined by the biological sex of the referent regardless of the morphophonological shape of the noun. For non-human nouns, an -a ending implies feminine, although the reverse is not
necessarily true (i.e., not all feminine nouns end in -a). To put it differently, though 
–a is a morphophonological marker of feminine gender in Arabic (especially for 
non-humans), the absence of -a is not an indicator that the noun is grammatically 
masculine.

Though the discussion above shows the regularity of the Arabic gender 
system as the gender of the majority of the nouns can be determined either by their 
meanings or their forms with few exceptions, there is an issue that might have an 
effect on the way this system is acquired. In Arabic, unlike in Spanish for example, 
feminine gender is overtly marked by the -a ending\textsuperscript{31} whereas masculine gender 
lacks any gender marking. In contrast, both genders are overtly marked in languages 
like Spanish, Italian, and Portuguese, with masculine nouns ending in –o and 
feminine nouns ending in -a. Even in languages where the gender system is less 
transparent than in the aforementioned languages, such as in French (see Boloh & 
Ibernion, 2010) or German (see Szagun et al., 2007), both masculine and feminine 
tend to correlate with overt endings. In Arabic however, there is a kind of imbalance 
in the type of information the two genders tend to correlate with. Feminine gender 
correlates with an overt ending whereas masculine gender correlates with the lack 
of a specific ending. The effect of this imbalance will be discussed in chapter 8.

\textsuperscript{31} Bearing in mind the very few exceptions such as the nouns mentioned in (21), a situation that is 
very common in languages. That is, the correlation between morphophonological endings or noun 
markers does not have to be perfect to be used by children as a cue to noun gender.
The following section will present how the two gender classes are manifested in the agreement between the nouns and different targets.

3.3.2. Gender agreement

Agreement is the means via which gender manifests itself in a language and it is through agreement that one can establish how many genders a language has (Corbett, 1991). Before presenting gender agreement patterns in Arabic, a few words about Arabic morphology are in order here. Arabic, like other Semitic languages, has a non-concatenative morphology. Unlike in English for example, where a word like *truthfulness* has three continuous morphemes (truth-ful-ness), content words (nouns, adjectives, and verbs) in Arabic are composed of a set of discontinuous morphemes: a root and a pattern (see among others Alhawary, 2009; Ryding, 2005; Watson, 2007). The root carries the core semantic meaning while the pattern usually consists of the vocalic template. These two morphemes must combine together to form a word in Arabic; both are bound morphemes that cannot occur separately. Table 3.1 illustrates the Arabic root-and-pattern morphology using the widely cited example *k-t-b*. As this table illustrates, a new word is derived every time the consonantal root *k-t-b* combines with a different vocalic template. Arabic also has some derivational and inflectional affixes. The feminine suffix -*a* for instance, can be added to the word *kātib* ‘writer’ to generate the feminine *kātib-a* ‘female writer’.
Table 3.1. Example of root and pattern morphology in Standard Arabic.

<table>
<thead>
<tr>
<th>Root</th>
<th>Vocalic template</th>
<th>Word and lexical category</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-a-C-a-C</td>
<td><em>katab</em> (verb)</td>
<td>wrote</td>
<td></td>
</tr>
<tr>
<td>C-ā-C-i-C</td>
<td><em>kātib</em> (noun)</td>
<td>writer</td>
<td></td>
</tr>
<tr>
<td>C-u-CC-ā-C</td>
<td><em>kuttāb</em> (noun)</td>
<td>writers</td>
<td></td>
</tr>
<tr>
<td>C-i-C-ā-C</td>
<td><em>kitab</em> (noun)</td>
<td>book</td>
<td></td>
</tr>
<tr>
<td>C-u-C-u-C</td>
<td><em>kutub</em> (noun)</td>
<td>books</td>
<td></td>
</tr>
</tbody>
</table>

With this in mind, I will now discuss gender agreement in Arabic. Unlike Spanish and French for example, Arabic does not have gender marked articles that obligatorily accompany the noun. The definite article *ʿal*32 in Arabic is invariable for gender as the following two sentences show. Though *kursi* (street) in (25a) is masculine whereas *ḡurfa* (room) in (25b) is feminine, the definite article does not take different forms according to the gender of the noun.

---

32 See footnote 4.
25) SA

a. ʾal-kursi  
   ʾnaḏīf
   the-chair(M)  clean(M)
   ‘The chair is clean.’

b. ʾal-ġurfa  
   ʾnaḏīf-a
   the-room(F)  clean-F
   ‘The room is clean.’

In Arabic, adjectives as in (26) agree with the noun in gender. Because ʾil-kirsi (the chair) in (26a) and ʾabīb (doctor) in (27a) are masculine, the adjectives lack agreement markers whereas in (26b) and (27b) the adjectives are feminine because ʾil-kanaba (the sofa) and ʾabīb-a are feminine. Noun-adjective agreement in Arabic exhibits, to a great extent, *alliterative concord* (Corbett, 1991). That is, “the noun itself includes a form which is identical to the gender agreement marker and which clearly indicates the gender of the noun” (Corbett, 1991, p. 117). As the examples in (26b) and (27b) illustrate, both the nouns and the adjectives end in –a.
26) HA

a. ʾil-kirsi  

murīḥ

the-chair(M)  comfortable(M)

‘The chair is comfortable.’

b. ʾil-kanaba

murīḥ-a

the-sofa(F)  comfortable-F

‘The sofa is comfortable.’

27) SA

a. ṭabīb  

muxlis

docor(M)  dedicated(M)

‘A dedicated doctor’

b. ṭabāb-a  

muxlis-a

docor-F  dedicated-F

‘A dedicated doctor.’

It is worth mentioning that the formation of some feminine colour adjectives in Arabic involves in addition to suffixation, changes in the vocalic pattern. As the example in (28) illustrate, the masculine adjective has the template [ʾaC-C-a-C] whereas the feminine adjective has the template [C-a-C-C-ā'] (SA)/ [C-a-C-C-a] (HA). Colour adjectives that do not have the pattern [ʾaC-C-a-C] in their masculine forms do not involve templatic changes like the colour pink in (29). This change in
the vocalic template does not contradict the fact that noun-adjective agreement exhibits *alliterative concord*, at least in HA, even in the case of colour adjectives.

28) ʾahmar → ḥamr-ā (SA)/ ḥamr-a (HA)
   red(M) → red-F
   ‘red.’

29) wardi → wardiyy-a
   pink(M) → pink-F
   ‘pink.’

In addition to adjectives, demonstratives as in (30) take different forms according to the gender of the noun. In (30a) the demonstrative is in its masculine form because the noun *galam* (pen) is masculine whereas in (30b), the demonstrative is feminine because the noun *fūṭa* (towel) is feminine.

30) HA
   
   a. hāḍa ʾil-galam nāšif
   This(M) the-pen(M) dry(M)
   ‘This pen is dry.’

33 The insertion of yy is due to phonological and syllabification constraints and has nothing to do with the vocalic template.
b. ħāḍī 'il-fūṭa nāšf-a

This(F) the-towel(F) dry-F

‘This towel is dry.’

Imperatives in Arabic also agree with the implied subject in gender as the examples in (31) illustrate. The imperative in (31a) lacks a gender marker because the addressee is masculine whereas the imperative in (31b) shows feminine agreement because the addressee is feminine.

31) SA

a. ʾuktub 'al-dars

(you(M)) write(M) the-lesson

‘Write the lesson.’

b. ʾuktub-i 'al-dars

(You(F)) write-F the-lesson

‘Write the lesson.’

Also, verbs in Arabic agree in gender with their subjects. The examples in (32) show subject-verb agreement in the perfective aspect and the examples in (33) show subject-verb agreement in the imperfective aspect. In (32a) and (33a), the

---

34 Though subject-verb agreement in Arabic depends on word order (VS / SV), it is only number agreement that is affected by word order but gender agreement is always realized no matter what word order the subject and verb have.
verbs have no gender marking because the subjects 'il-gemar (the moon) and 'al-walad (the boy) are masculine, whereas in (32b) and (33b), the verbs show feminine agreement because the subjects 'il-šams (the sun) and 'al-'umm (the mother) are feminine.35

32) HA
a. 'il-gemar ṭala'
   the-moon(M) appeared
   ‘The moon appeared.’

b. 'il-šams ṣlaʿa-t
   the-sun(F) appeared-F
   ‘The sun appeared.’

33) SA
a. 'al-walad ya-lʿab
   The-boy(M) 3MS-play
   ‘The boy is playing.’

35 Notice that in the perfective verbs, gender agreement is realized by means of suffixation while in the imperfective verbs gender agreement is realized by means of prefixation.
b. ‘al-ʾumm  ta-qra’  ‘al-qiṣṣ-a

The-mother(F)  3FS-read  the-story-F

‘The mother is reading the story.’

Relative pronouns also agree with nouns in gender but only in Standard Arabic. Different relative pronouns are used according to the gender of the antecedent as the SA examples in (34a) and (35a) illustrate. In (34a) the relative pronoun ʾallati is feminine because ʾal-tuffāḥa (the apple) is feminine whereas in (35a) the relative pronoun ʾallaḏi is masculine because ʾal-qalam (the pen) is masculine. In most Arabic dialects however, a single relative pronoun that is invariable for number and gender is used with all nouns (Ángeles, 2011). Ferguson (1959) listed this as one of the fifteen features present in all modern Arabic dialects. The most common relative pronoun among the Arabic dialects which is also the relative pronoun used in HA is ʾilli. We can think of this relative pronoun as resembling that in English. As (34b) and (35b) illustrate, the same relative pronoun ʾilli is used though ʾil-tuffāḥa (the apple) is feminine while ʾil-qalam36 (the pen) is masculine.

---

36 Some Hasawi speakers might pronounce this word as ʾil-galam.
34)  

a. man 'allaḍi 'akala 'lt-tuffāḥa 'allati fī 'al-ṭallāja

who that(MS) ate the-apple(F) that(FS) in the-fridge(F)?

b. mīn kal 'il-tuffāḥa 'illi fī 'il-ṭallāja?

Who ate the-apple(F) that in the-fridge(F)?

‘Who ate the apple that is in the fridge?’

35)  

a. 'al-qalam 'allaḍi 'ala 'al-maktab lī

The-pen(M) that(MS) on the-desk for-me

‘The pen that is on the desk is mine.’

b. 'il-qalam 'illi 'ala 'il-maktab ḥaggī.

The-pen(M) that on the-desk mine

‘The pen that is on the desk is mine.’

As I mentioned in the first chapter, this project will look into singular nouns only because only singular agreement provides reliable evidence to the gender of the nouns in Arabic. The discussion above shows how gender agreement with singular nouns is ubiquitous in Arabic and that the only discrepancy found between SA and HA is in the relative pronouns. When it comes to gender agreement with
plural nouns, gender interacts with number in a special way in Arabic and this interaction is conditioned by animacy, or more precisely, humanness. In (36a), a masculine plural adjective is used when the noun is a human-referring masculine plural. Similarly, a feminine plural adjective is used in (36b) because the noun is a human-referring feminine plural.

36) SA

a. ‘al-ṭullāb mujtahid-īn

the-students(MPl) hard-working-MPL

‘The (male) students are hard-working.’

b. ‘al-ṭālib-āt mujtahid-āt

The-students-FPl hard-working-FPl

‘The (female) students are hard-working.’

Nonetheless, when plural nouns refer to non-humans in Arabic, gender and number agreement are neutralized. All non-human plural nouns require singular feminine agreeing forms as shown in examples (37) and (38). Though kitāb (book) is masculine while nāfiḍa (window) is feminine in Arabic as indicated by the

---

37 See Benduhaish (2018) for the effect of word order (VS/ SV) on this interaction in Standard Arabic.

38 This ‘full’ agreement is not always the case in HA. Some speakers do use masculine plural adjectives in both (36a) and (36b). See Albirini et al. (2013) for the same patterns in Palestinian and Egyptian Arabic.
demonstratives and adjectives in (37a) and (38a) respectively, when these two nouns are in their plural forms, both require feminine singular forms of the agreeing demonstratives and adjectives as in (37b) and (38b).

37) SA

a.  

<table>
<thead>
<tr>
<th>Arabic</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>hāḏa</td>
<td>this(M)</td>
</tr>
<tr>
<td>'al-kitāb</td>
<td>the-book(M)</td>
</tr>
<tr>
<td>mufīd</td>
<td>useful(M)</td>
</tr>
</tbody>
</table>

(This book is useful.)

b.  

<table>
<thead>
<tr>
<th>Arabic</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>hāḏihi</td>
<td>this(F)</td>
</tr>
<tr>
<td>'al-kutub</td>
<td>the-books</td>
</tr>
<tr>
<td>mufīd-a</td>
<td>useful-F</td>
</tr>
</tbody>
</table>

‘These books are useful.’

38) SA

  c.  

<table>
<thead>
<tr>
<th>Arabic</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>hāḏihi</td>
<td>this(F)</td>
</tr>
<tr>
<td>'al-nāfiḍa</td>
<td>the-window(F)</td>
</tr>
<tr>
<td>naḍīf-a</td>
<td>clean-F</td>
</tr>
</tbody>
</table>

‘This window is clean.’

  d.  

<table>
<thead>
<tr>
<th>Arabic</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>hāḏihi</td>
<td>this(F)</td>
</tr>
<tr>
<td>'al-nawāfiḍ</td>
<td>the-windows</td>
</tr>
<tr>
<td>naḍīf-a</td>
<td>clean-F</td>
</tr>
</tbody>
</table>

‘These windows are clean.’

One comment is in order here: though this neutralized feminine singular agreement with non-human nouns is strict in Standard Arabic, there is a noticeable degree of
variation in this pattern of agreement among Hasawi speakers. All the sentences in (39) are common in HA. To explain, some Hasawi speakers use feminine singular agreeing forms with non-human plurals similar to SA while others use plural agreement with such nouns. The plural adjectives in (39b) and (39c) are not feminine though the singular noun nāḏa (window) is feminine.\(^{39}\)

\(39)\) HA

a. 'il-newāfid ṉēḏīf-a

the-windows clean-F

b. 'il-newāfid ṉḏāf

the-windows clean(Pl)

c. 'il-newāfid ṉēḏīf-in

the-windows clean-MPl

‘The windows are clean.’

This neutralized agreement does not apply to collective plurals, which are more similar to mass nouns in English, such as ward (flowers), samak (fish).

\(^{39}\) In order to make sure that these are not personal speculations, I asked a group of Hasawi adults, who were born in Alhasa of two Hasawi parents and who lived their whole lives there, how often they use singular/plural adjectives with non-human plurals. I provided them with thirteen non-human plural nouns, with each noun used in a pair of sentences. One of the sentences had a singular feminine adjective and the other had a plural adjective. I asked them to rate each sentence in every pair according to how often they use it in their everyday life and communication by giving a percentage (0-100). There was a high degree of variations in the ratings the Hasawi adults gave to each sentence in every pair with all the sentences ranging from 0% by some speakers to 100% by other speakers. There was a high degree of variation even between speakers from the same family.
Collective plurals trigger masculine singular agreement. As illustrated in (40), the adjective *jamīl* (beautiful) is masculine singular.

40) HA

<table>
<thead>
<tr>
<th>'il-ward</th>
<th>'illi</th>
<th>jib-t-i</th>
<th>l-ī</th>
<th>jamīl</th>
<th>jiddan</th>
</tr>
</thead>
</table>

The flowers that brought-2S-F for-me beautiful very

‘The flowers you brought me are very beautiful.’

In sum, this discussion of the agreement patterns in Arabic demonstrates very clearly how gender and number interacts in determining the form of the agreeing targets and how this interaction is conditioned by other factors like humanness and word order. It is only the singular form of the noun that provides a reliable cue to its gender both in SA and HA. Therefore, this thesis will investigate the assignment of gender to singular nouns.

3.4. Distribution of the gender classes and noun-internal gender-correlating information

I consulted four different corpora in order to have an estimate of the distribution of the two gender classes in Arabic. In addition, the distribution of the gender-correlating noun-internal information was estimated.
• The first corpus was extracted from the Arabic Communicative Development Inventory (CDI) word list\textsuperscript{40} (Abdelwahab et al., 2021). Arabic CDI word list is a tool used to assess language development in Arabic-learning children between 8 and 30 months. It contains 100 words and it is available in 17 different Arabic dialects\textsuperscript{41}. This list can be freely accessed at \url{http://www.psy.plymouth.ac.uk/OpenArabicCDI/}.

• The second and third corpora were extracted from Salama’s Egyptian Arabic corpus (Salama & Alansary, 2016). This corpus has data from 10 Egyptian children (1;7-3;8 years) and Egyptian mothers or an instructor interacting with the children. The children in this corpus produced 2,383 wordform types (10,650 tokens). The mothers and the instructor produced 955 wordform types (4,951 tokens). In this thesis, I refer to the corpora extracted from Salama & Alansary (2016) as Salama child corpus and Salama mother corpus. This corpus is publicly available through the CHILDES project (MacWhinney, 2000). The Salama corpus in combination with the Arabic CDI word list constitutes a good estimate of the words produced by Arabic-speaking children in their early years of development.

\textsuperscript{40} For the words that appear in the plural form in the Arabic CDI word list, a singular form is included in the analysis; for example, the word ʿayn (eye) is listed as ʿuyūn (eyes) but I included the singular form, rather than the plural one, in the analysis because as I mentioned earlier, my project is an attempt to investigate the acquisition of gender in itself, not the interaction between gender and other features like number. Also, collective nouns like fulūs (money) are not included if there is no corresponding unit noun in Arabic. However, if the collective plural like burtuqāl (orange) has a unit noun burtuqāla (an orange), the latter is included in the analysis. The same method was followed when extracting the nouns from the Salama corpus.

\textsuperscript{41} Notice that only the Saudi word list was consulted.
The fourth and last corpus was extracted from Buckwalter and Parkinson’s (2011) Frequency Dictionary\textsuperscript{42} of Arabic. It is a list of the 5000 most frequently used words in Modern Standard Arabic. This dictionary was compiled from a corpus of 30 million words. Only 10\% of the corpus was based on spontaneous speech data. The remaining 90\% of the corpus was from written sources of five genre types: “(i) daily newswire; (ii) newspaper editorials, opinion essays, regular columns; (iii) learned prose, consisting mostly of articles in academic and scientific journals, including “Islamic guidance” essays, and popular but formal magazines and publications; (iv) posting on internet discussion forums; and (v) literature and fiction, made up of short stories, novels, and plays.” (Buckwalter & Parkinson, 2011, p.3). This dictionary provides an estimation of the distribution of the gender classes and the types of noun-internal gender-correlating information available in the most frequent nouns in Arabic.

\textsuperscript{42}A number of things were taken into account when extracting nouns from this corpus;

a) When a noun is used with both genders, it is included in both feminine and masculine lists unless its occurrence in one gender is very rare. In the latter case the noun is included only under the gender it is more associated with. For example, the noun \textit{sīq} (market) is feminine and rarely used as masculine; therefore, it is included under the feminine nouns list only. However, the noun \textit{ṭūrūq} (road/way) is used with both genders and therefore it is included in both lists,

b) Some feminine nouns are derived from the masculine nouns simply by suffixing the (-a) ending to the masculine form; such as \textit{ṣādiq/ṣādiq-a} (friend(M)/friend-F). This class of words is most of the time listed in the dictionary in the masculine form. In my analysis however, I also included the feminine forms because they are among the most common nouns in Arabic,

c) for collective nouns like \textit{samak} (fish.Pl), \textit{bayd} (eggs), only the unit form is included, i.e. \textit{samak-a}, \textit{bayd-a}, and

d) Some noun entries in the dictionary are not included in the analysis because (for me) they are not qualified to be categorized as nouns in Arabic. To illustrate, some noun entries are prepositions and adverbs, e.g. \textit{fawrān} (at once), \textit{ṭawāl} (throughout).
I extracted only the singular nouns from those corpora because as I mentioned earlier, the focus of this thesis is the acquisition of gender, not the way it interacts with number. The numbers given here represent the type counts of the nouns in those corpora.

3.4.1. The distribution of the gender classes in Arabic

As Fig. 3.1 illustrates, the gender classes are almost evenly distributed in Arabic with rather more masculine (a mean of 57.4%) than feminine nouns (a mean of 42.6%). Below, I will give details of the numbers of nouns in each of the corpora.

![Bar chart showing the distribution of gender classes in Arabic corpora](image)

**Fig. 3.1.** The distribution of the gender classes (masculine & feminine) in Arabic.
➢ *The Saudi Arabic CDI Word list*

There were 54 noun types; 32 masculine (59%) and 22 feminine (41%) in this corpus.

➢ *Salama (2016) Egyptian Arabic corpus*

The children in this corpus produced 297 singular noun types. There were 157 (52.9%) masculine nouns compared to 140 (47.1%) feminine nouns. The mothers/instructor in this corpus produced 159 singular noun types; 84 nouns (53%) were masculine and 75 nouns (47%) were feminine.

➢ *Frequency dictionary of Arabic (Buckwalter and Parkinson, 2011)*

There are 2178 singular noun types in this dictionary (1406 masculine and 772 feminine). In other words, 64.5% of the most frequent Arabic nouns according to this dictionary are masculine whereas 35.5% are feminine.

3.4.2. *The distribution of the noun internal gender correlating information*

As I discussed in section 3.2., gender in Arabic tends to correlate with the internal properties of the nouns. First, semantic properties manifested in the perfect correlation between grammatical gender and natural gender in Arabic. Nouns denoting male and female humans represent 10% of the Arabic lexicon according to the four corpora consulted in this thesis. Second, morphophonological properties, manifested in the correlation between feminine gender and the –a ending and the correlation between masculine gender and the lack of the –a ending.
Fig. 3.2 shows that when the nouns end in –a in Arabic, they are feminine about 99% of the time. On the other hand, when the nouns lack the –a ending in Arabic, they are masculine about 92.8% of the time. This shows that the correlation between -a and feminine gender and between the lack of -a and masculine gender in Arabic is very strong regardless of the register. It is a pattern that is pervasively available in the speech addressed to children.

Fig. 3.2. The distribution of the correlation between morphophonological ending and grammatical gender.

As I mentioned in section 3.3.1, it is not uncommon in Arabic for the morphophonological properties and the semantic properties of the noun to align together suggesting the assignment of the same gender class. That is, it is very common for male-human-referring nouns to lack the feminine morphophonological
marker -a. Table 3.2 shows that in the four corpora, it is almost always the case that male-human referring nouns lack the feminine morphophonological marker. Considering the female-human-referring nouns however, there is a high degree of variation between the four corpora consulted. The generalization that I can draw on the basis of the three spoken corpora (CDI word list, Salama child, and Salama mother) is that female-human-referring nouns do not end in -a in almost half of the cases.\textsuperscript{43}

Table 3.2. The correlation between the natural gender information and the morphophonological information in the four corpora.

<table>
<thead>
<tr>
<th>Corpus</th>
<th>Male humans</th>
<th>Female humans</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No -a ending</td>
<td>-a ending</td>
<td>No -a ending</td>
<td>-a ending</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDI</td>
<td>100.0%</td>
<td>0.0%</td>
<td>100.0%</td>
<td>0.0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salama child</td>
<td>95.0%</td>
<td>5.0%</td>
<td>44.0%</td>
<td>56.0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salama mother</td>
<td>100.0%</td>
<td>0.0%</td>
<td>43.0%</td>
<td>57.0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequency</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dictionary</td>
<td>98.4%</td>
<td>1.6%</td>
<td>8.0%</td>
<td>92.0%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{43} As I mentioned in footnote 42, most of the -a ending female-human referring nouns were not originally listed in the frequency dictionary consulted but were added by the researcher. This may have an effect on the high percentage of -a ending female-human referring nouns in this corpus.
3.4.3. Types of endings other than -a

In addition to estimating the distribution of –a vs. no –a endings in the corpora and their correlation with the two gender classes, I also calculated the percentages of the different endings found in the no –a words. These estimates will be used in this thesis to evaluate two strategies that Arabic-learning children might follow when learning the gender system of their language. The first strategy suggests learning the association between -a ending and feminine gender and assigning masculine gender by default. The other strategy suggests learning all the different endings found in the no -a ending nouns as cues for masculine gender in addition to learning the association between -a and feminine gender. As shown in table 3.1, the majority (34.0%) of the no –a nouns end in stops (b, t, d, k, q, ’, ŋ, q). There is also a fair proportion of nouns (20.0%) ending in fricatives (f, t̪, ṭ, s, z, x, ð, h, ‘, h, ḍ, š). 16.0% of the no –a nouns end in nasals and 15.0% end in the rhotic (r). The liquid (l) is found at the end of 8.0% of the no –a nouns. The long vowels (ā & ū) are found at the end of 5.0% of the no –a nouns. 2.0% of the no –a nouns end in the affricate (j). Finally, the approximants (w & y) are found at the end of less than 0.1% of those nouns. So, basically, 95.0% of the masculine nouns end in consonants in contrast to the feminine nouns which end in vowel(s) more than 99.0% of the time. Whether Arabic-acquiring children learn the separate consonants as cues to masculine gender and -a as a cue to feminine gender or whether they learn the association between -a and feminine gender and assign masculine by default will be explored.
in the next chapters. In the next section, I will provide a range of evidence for the status of masculine as the default gender in Arabic.

Table 3.3. The distribution of the endings found in the no -a words.

<table>
<thead>
<tr>
<th></th>
<th>CDI</th>
<th>CDS mother</th>
<th>CDS child</th>
<th>Frequency dictionary</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stops</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b, t, d, k, q, ʾ, ṭ, ḏ)</td>
<td>45.5%</td>
<td>28.6%</td>
<td>28.0%</td>
<td>31.8%</td>
<td>34.0%</td>
</tr>
<tr>
<td><strong>Affricates</strong></td>
<td>3.0%</td>
<td>2.4%</td>
<td>1.9%</td>
<td>1.9%</td>
<td>2.0%</td>
</tr>
<tr>
<td>(j)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Fricatives</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(f, t, d, s, z, x, ǧ, h, ʾ, h, ḏ, ẓ)</td>
<td>15.2%</td>
<td>19.0%</td>
<td>23.0%</td>
<td>21.1%</td>
<td>20.0%</td>
</tr>
<tr>
<td><strong>Nasals</strong></td>
<td>12.0%</td>
<td>20.2%</td>
<td>18.0%</td>
<td>14.2%</td>
<td>16.0%</td>
</tr>
<tr>
<td>(m, n)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Liquids</strong></td>
<td>3.0%</td>
<td>7.0%</td>
<td>9.3%</td>
<td>11.4%</td>
<td>8.0%</td>
</tr>
<tr>
<td>(l)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Rhotics</strong></td>
<td>18.2%</td>
<td>14.3%</td>
<td>13.0%</td>
<td>14.6%</td>
<td>15.0%</td>
</tr>
<tr>
<td>(r)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Approximants</strong></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.3%</td>
<td>&lt; 0.1%</td>
</tr>
<tr>
<td>(w, j)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(w, ǧ)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Long vowels</strong></td>
<td>3.0%</td>
<td>8.3%</td>
<td>6.8%</td>
<td>2.4%</td>
<td>5.0%</td>
</tr>
<tr>
<td>(ū/ī)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

44 Notice that the Arabic consonantal inventory is dialect dependent. For example, many Arabic dialects neutralize the difference between ḏ and ẓ. The Egyptian dialect has ʾ for q and g for j. For ease of presentation, the Standard Arabic equivalents of the different consonants are included in this table.
3.5. Evidence for masculine agreement as the default agreement

As I discussed in chapter 1, gender as a morphosyntactic feature is non-canonical because every noun can only, in most cases, have a single inherent gender feature (masculine or feminine in Arabic for example). Gender is unlike other morphosyntactic features like number where a noun has available to it all values of this feature: singular, plural, dual and so on depending on the number system of the language. The language learner either learns the gender of a particular noun or not. Thus, when talking about default gender we are in fact talking about default agreement. That is, the pattern of agreement that is underspecified for gender features and which is expected to be learned earlier (Neeleman and Weerman, 1999) and used by learners when the gender of a particular noun has not yet been learned or in cases where there is no gender controller. In this section, I will provide a range of evidence to support the status of masculine agreement as the default agreement pattern in Arabic.

1. Morphophonological marking

As discussed in section 3.3.1, it is almost always the case that grammatically masculine nouns lack any gender markers whereas feminine nouns end in the feminine marker -a. Furthermore, a fair proportion of feminine nouns are derived from their masculine counterparts by the suffixation of -a as the examples in (17) above illustrate. There is no single case in Arabic in which masculine nouns are
derived from their feminine counterparts by means of suffixation. This demonstrates clearly that masculine gender is the morphophonologically unmarked (base) gender whereas feminine gender is the morphophonologically marked gender. Similar to the nouns themselves, masculine agreeing adjectives, imperatives and verbs lack morphophonological marking as the examples in (25, 26, 31, 32, and 33) show.

2. Coordination of different-gender nouns

Masculine agreement is used in the case of coordination between nouns of different genders as in (41) below. ‘ābā’ (fathers) is masculine and ‘ummah-āt (mothers) is feminine but the coordination of them requires a masculine agreeing adjective.

41) SA

‘al-‘ābā’ wa ‘al-‘ummah-āt mad’uww-ūn li ḥafl

The-father(M.Pl) and the-mother-F.Pl invited-M.Pl to ceremony

‘at-taxarruj

the-graduation

‘The fathers and the mothers are invited to the graduation ceremony.’

3. Agreement when there is no gender-carrying controller

Masculine agreement is also the default agreement when there is no gender-carrying controller as in (42). The adjective muhimm (important) in (42) is
in its default masculine form because the that-phrase has no gender feature specifications.

42) HA

\[
\text{muhimm} \quad \text{jiddan} \quad \text{l-ī} \quad \text{ʾinn-ak} \quad \text{ta-ḥḍir} \quad \text{ʾil-ḥafil}
\]

important very for-1S that-2M 2-attend the-ceremony

‘It is very important for me that you attend the ceremony.’

4. Agreement neutralization

There is a tendency in many Arabic dialects\(^{45}\) including Hasawi Arabic to use masculine plural verbs and adjectives whether the controlling noun refers to a group of female humans, male humans or non-humans as the examples in (43) and (44) illustrate. As I mentioned in section 3.3.2, agreement with non-human plurals in dialectal Arabic does not have to be feminine singular as in Standard Arabic. Plural masculine agreement is often used with non-human plurals regardless of the gender of the singular noun. The neutralization direction is almost always towards the masculine but not the feminine plural.

\(^{45}\) See Albirini et al., 2013 for the patterns of agreement in Palestinian and Egyptian Arabic among others.
43) HA

a. 'il-ban-āt  ʧāh-aw
   The-girl-FP  fell(down)-MP
   (The girls fell down.)

b. 'il-ʾawlād  ʧāh-aw
   The-boys  fell(down)-MP
   (The boys fell down.)

c. 'il-kutub  ʧāh-aw
   The-books  fell(down)-MP
   (The books fell down.)

44) HA

a. 'il-ban-āt  ḥilw-īn
   The-girl-FP  sweet-MP
   (The girls are beautiful.)

b. 'il-ʾawlād  ḥilw-īn
   The-boys  sweet-MP
   (The boys are beautiful.)

c. 'il-kutub  ḥilw-īn
   The-books  sweet-MP
   (The books are beautiful.)
3.6. Conclusion

From this presentation of the Arabic gender system, it is clear that Arabic has a very predictive system that can be accounted for by the interaction of two assignment principles. Those principles accurately predict the gender of more than 90% of the Arabic nouns. I also showed how gender manifests itself through the agreement between nouns and different elements. Gender agreement is ubiquitous in Arabic, which contributes to the discovery of the gender system and ultimately learning it. Nonetheless, an area where the Arabic gender system exhibits a kind of imbalance is that feminine gender correlates with overt uniform morphophonological marking whereas masculine gender correlates with the lack of a uniform morphophonological marking. The lack of a morphophonological marker could in itself be used as a cue to masculine gender by children learning Arabic. As revealed from the analysis of the types of endings other than -a, the majority of the masculine nouns (95%) end in consonants. Though the consonantal space exhibits a lower degree of homogeneity compared to the vowel space, it is possible that children could use this information as cue to gender class membership. All these possibilities in addition to the effect of the seaming imbalance between feminine and masculine gender in Arabic will be considered in the next chapters the focus of which will be describing the developmental trajectory of the Arabic gender system and the interaction between the noun internal and the noun-external information over the course of development.
Chapter 4: Screening criteria: MGIT test and Experiment I

4.1. Introduction

In order to establish the ground for my argument, I used two criteria to filter the young participants: the Michigan Gender Identity Test (MGIT) (Dull et al., 1975) and a semi-spontaneous production task (Experiment I). In the existing literature, it is not clear whether the youngest children’s failure to use the semantic properties and particularly natural gender (male vs. female) to classify nouns into different gender classes is due to the fact that they have not yet acquired the cognitive concepts associated with those properties or that they have already acquired such concepts but have not yet learned that those concepts correlate with grammatical gender in their language and therefore they do not use them as cues in gender attribution. Therefore, an important step towards clarifying this point was establishing that the children who will take part in this study had learned the cognitive concepts that correlate with the grammatical gender of the language being studied as this presupposes using these concepts in classification. Applying MGIT prior to the experimental tasks was a good step to find out whether children have the cognitive concept necessary to use natural gender as a cue in gender assignment to the new nouns or not. It also allowed me to find out WHEN L1 learners become able to use semantic concepts, once they have been acquired, in gender assignment. In the previous studies that report a bias in young children to prioritise morphophonological cues over semantic cues in grammatical gender assignment, children were not screened for their understanding of natural gender. This leaves
the possibility that the reported bias was due to lack of cognitive distinction between males and females (see Schütze, 2009) rather than a failure in weighting the cues in an adult-like fashion.

Similarly, Experiment I was a very important and informing task because it allowed me to find out whether or not the participant was able to switch between masculine and feminine genders when talking about the familiar nouns. It was essential to check that because in both Experiment II (Chapter 5) and Experiment III (chapter 6), I used agreement as the test for gender assignment. The gender the participants assigned to the nouns in those experiments was determined by the agreeing forms they produced. It was therefore necessary to ensure that the participants were able to use correct gender agreement for familiar words. To put it differently, participants who overgeneralize one gender by producing only one form of demonstratives, adjectives, verbs, or any other agreeing forms when talking about familiar nouns are not expected to perform in a different way when it comes to the novel nouns. They are also likely to show clear distinctions in their gender, but because of a problem in agreement. Experiment I was conducted to rule out such a possibility. This measure has not always been taken in previous studies, making it difficult to interpret the reported results. In Karmiloff-Smith (1979), Pérez-Pereira (1991), Boloh & Ibernon (2010, 2013), and Boloh et al.’s (2012) studies for example, it was found that children’s responses were better with masculine gender than with feminine gender. This could be due to the insufficient
knowledge of gender agreement which was used to test gender assignment in those studies similar to the current study.

In this chapter, both screening criteria will be presented and discussed. Section 4.1 presents the MGIT and section 4.2 presents Experiment I.

4.2. Screening criterion I: MGIT

4.2.1. Method

4.2.1.1. Participants

The participants were 126 typically-developing Hasawi Arabic-speaking children of two age groups. The first group consisted of 50 toddlers (23 females) aged 2;06 - 3;11 (Mean age: 3;0). The second age group consisted of 76 pre-schoolers (45 females) aged 4;0 - 5;11 (Mean age: 4;9). All the children were local residents in Ahsa (a city in the eastern province in Saudi Arabia). Both parents of all the children were speakers of HA.

4.2.1.2. Test materials

The test materials consisted of 22 (3.5 by 4.5 inch) pictures as follows;

Picture sort I: dogs and balls

a. line drawings: a dog (D) and a ball (B).

b. Trial 1: 2 dogs and 2 balls (B D B D).

c. Trial 2: 2 dogs and 3 balls (B D B D B).
Picture sort II: boys and girls

   a. line drawings: a boy (B) and a girl (G) (B G).
   b. Trial 1: 2 boys and 2 girls (B G B G)
   c. Trial 2: 2 boys, 2 girls, and child own photograph\textsuperscript{46} (B G B G C).

4.2.1.3. Procedure\textsuperscript{47}

The test was carried out in a quiet room in the experimenter’s house or in the participant’s house whichever was more convenient for the participant. Each child was interviewed individually except for 12 children who were too shy to be left with the experimenter alone and therefore they were accompanied by their mothers.

First, the child was presented with the line drawings of the ball and the dog and was asked “do you know what these are?” If the child did not respond, they were told “this is a dog and this is a ball.” Then, the child was told “now I will show you more pictures and I want you to tell me which one is a dog and which one is a ball.” The pictures were placed in front of the child below the line drawings. As indicated in Dull et al. (1975), Trial 2 of the picture sorting I is optional according to the authors and it is only used if the experimenter has questions about the child’s performance\textsuperscript{48}. Then, the child was presented with the line drawings of the boy and

\textsuperscript{46} Each child’s mother was asked to bring a recent photograph of the child. I asked the child ‘who is this’ when I placed their photograph under the line drawings in this trial in order to make sure that the child is aware that it is a picture of themselves.

\textsuperscript{47} The same procedure described in Dull et al. (1975) was followed which was recommended by one of the authors who was contacted to obtain a more recent version of the test or a complete set of it.

\textsuperscript{48} When applying the test, trial 2 was never needed.
the girl and was told “here I have a boy and a girl.” After that, the pictures of the two boys and two girls were placed below the line drawings and the child was asked “now tell me which one of these is a boy and which one is a girl.” Finally, trial 2 was done following the same procedure by placing the pictures of the two boys and the two girls in addition to the child’s own picture below the line drawings. The experimenter took notes of each child’s performance after the test.

4.2.2. Results

The criterion for passing MGIT was to pass all the trials successfully. So, only those who did so in addition to passing the screening criterion II (Experiment I) were qualified to take part in the two subsequent experiments.

Only 27 out of the 50 younger children passed the test successfully. Those children ranged in age between 2;6 and 3;11 (mean = 3;3, sd = 0.47) Though the minimal age of the children who passed the MGIT was 2;6, there were children as old as 3;6 who failed the test. The children who failed the test ranged in age between 2;6 and 3;6 (mean = 2;8, sd = 0.48). On the other hand, 75 out of the 76 participants in the older age group passed the MGIT test successfully. The children who passed the test ranged in age between 4;0 and 5;11 (mean = 4;9, sd = 0.57). The only child who failed the test was 4;4.

4.3. Screening criterion II: Experiment I

The aim of this experiment is two-fold. First, it provides a picture of how the gender system looks in the early stages of language development, what sort of errors are
produced by children during those stages, and whether one gender is overgeneralized or used by learners as a default gender. Second, it allows the researcher to find out whether or not the participant is able to switch between masculine and feminine agreeing forms when talking about familiar items. As mentioned in the introduction, participants who overgeneralize one gender over the other when talking about familiar nouns are not expected to behave differently when describing novel nouns. Checking that the participants have sufficient knowledge of gender agreement is essential to eliminate the possibility that their responses in experiments II and III are affected by the lack of this knowledge.

4.3.1. Method

4.3.1.1. Participants
The participants were the same participants who took part in the MGIT (section 4.1).

4.3.1.2. Materials
A toy box was used to elicit sentences with gender agreeing elements. The toys included objects, animals, and human-like characters mostly known by young Arabic-speaking children (Fig. 4.1 shows some of the toys used). Toys (and 3D characters in experiments II and III), not pictures, were used in this project to ascertain that the participants describe the character itself not the picture\(^{49}\). In

\(^{49}\)The word ṣūra 'picture' is feminine in Arabic.
addition to the familiarity criterion, the items were selected to represent all the possibilities that Arabic-speaking children encounter in the input language. To put it differently, the items represented -a ending and non -a ending nouns, they also represented humans, animals, and objects. Furthermore, the toy box contained pairs of toys that represent different characteristics, e.g. different colours or different sizes (see Appendix E for pictures of the full set of the toys used in this experiment). Half of the items were grammatically masculine and half of them were grammatically feminine. The 18 items are given in Table 4.1.

Fig. 4.1. Examples of the toys used in experiment I.
Table 4.1. List of the familiar items used in experiment I.

<table>
<thead>
<tr>
<th>Humans</th>
<th>Animals</th>
<th>Objects</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Masculine</strong></td>
<td><strong>Feminine</strong></td>
<td><strong>Masculine</strong></td>
</tr>
<tr>
<td><em>walad</em></td>
<td><em>bint</em></td>
<td><em>fil</em></td>
</tr>
<tr>
<td>‘boy’</td>
<td>‘girl’</td>
<td>‘elephant’</td>
</tr>
<tr>
<td><em>bāba</em></td>
<td><em>māma</em></td>
<td><em>’asad/ nimir</em></td>
</tr>
<tr>
<td>‘father’</td>
<td>‘mother’</td>
<td>‘lion/tiger’</td>
</tr>
<tr>
<td><em>diktūr/ṭabīb</em></td>
<td><em>’arūs</em></td>
<td><em>ḥiṣān</em></td>
</tr>
<tr>
<td>‘(male medical)’</td>
<td>‘bride’</td>
<td>‘horse’</td>
</tr>
</tbody>
</table>

4.3.1.3. **Procedure**

The experiment was carried out in a quiet room in the experimenter’s house or in the participant’s house whichever was more convenient for the participant. Each

---

Notes:

50 It was the child who first named the object (in response to the question: “What do I have?”), and that form was used for the subsequent task. In case the child did not know or did not name the object, the experimenter provided the noun that is shown in this table. These forms are HA, not SA, varieties.

51 For both *baba* and *mama*, the child was first shown a picture with a mother and a daughter/ a father and a son and was asked ‘can you tell me what is happening here?’ Those two pictures (which are included in Appendix D) were used to provide the context for presenting the pairs of the 3D characters as fathers and mothers.

52 The toy box contained tiger toys but some of the children were calling them tigers and others were calling them lions. The gender attributed to both animals is included in the results because both animals have masculine grammatical gender in Arabic and both of them are non -a ending nouns.

53 One girl pronounced this as ‘bgara’ which is a common pronunciation in HA. However, the variation in the way this word is pronounced by different speakers has no effect in the argument made here as the main concern is the -a ending which is present in all the variations.
participant was interviewed individually by the experimenter. However, 12 children from the young age group were accompanied by their mothers because they refused to be left alone with the experimenter. The mothers were instructed not to speak to their children during the experiment and to keep quiet. The answers were recorded using a Zoom recorder.

The task was carried out in the following scenario:

The experimenter presented the child with a toy and asked them “what do I have?” and then the experimenter presented a similar item and said “look! I have an X more.” Then, the experimenter asked the child “which X do you want?” or “which X do you like?” If the participant simply pointed to the item or just used a demonstrative to point to it, the experimenter asked “why?” to give an opportunity for producing more agreeing elements. Such questions prompted the production of utterances with gender agreeing elements instead of just naming the items especially in the case of the youngest subjects. The 18 pairs were presented randomly.

4.3.2. Results

The dependent variable was the grammatical gender (masculine/feminine) attributed to each item which was indicated by the agreeing element used by the

---

54 These questions do not contain any gender marked elements which might prompt the way the subjects could respond.

55 See Appendix A for the Arabic script of the prompt questions with some example answers.
participant to describe the item in question. Table 4.2. shows the distribution of the different types of agreeing elements produced by each age group.

Table 4.2. The distribution of the agreeing elements by age group.

<table>
<thead>
<tr>
<th>Type of agreeing element</th>
<th>Age group</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2-3 years</td>
<td>4-5 years</td>
<td></td>
</tr>
<tr>
<td>Demonstratives</td>
<td>76.6%</td>
<td>56.3%</td>
<td></td>
</tr>
<tr>
<td>Adjectives</td>
<td>19.0%</td>
<td>38.1%</td>
<td></td>
</tr>
<tr>
<td>Verbs</td>
<td>3.7%</td>
<td>3.8%</td>
<td></td>
</tr>
<tr>
<td>Pronouns</td>
<td>0.7%</td>
<td>1.8%</td>
<td></td>
</tr>
</tbody>
</table>

To give some examples\footnote{Henceforth, when presenting example answers from the children, the prompt questions will be given only in English with the noun being tested underlined and its gender in parentheses. Full transcription, annotation and English translation of the child’s answer will be presented. The agreeing forms will also be underlined in the English translation of the answers. As stated in the previous footnote, Appendix A presents the Arabic script of the prompt questions used in this experiment with example answers from one of the children.}: when the children were asked which item of the pairs they liked, the most common answers among the younger age group (76.6%) were the answers in (45) and (46) below, which contained a demonstrative. Yet, some children from the same age group gave more elaborate answers that included adjectives, verbs, and personal pronouns as illustrated in examples (47) and (48) which were given by a three-year-and-four-month boy.
45) Which elephant(F) do you like?

\[ \text{hāḍa/ \ ḏa} \]
this(M)/ this(M)

‘This one.’

46) Which fish(F) do you like?

\[ \text{hāḍi / ḏi} \]
this(F)/ this(F)

‘This one.’

47) Which elephant(M) do you like?

\[ 'a-ḥib \ hāḍa \ 'ašān \ kibīr \ yi-ṭali' \ ṣawt \]
1S-love this(M) because big 3MS-produce sound

‘I love this elephant because it is big (and) makes sound.’

48) Which fish(F) do you like?

\[ 'a-ḥib \ 'il-samaka \ hāḍi \ 'ašān \ 'ihya \ kibīr-a \ tā-kil \]
1S-love the-fish(F) this(F) because she big-F 3FS-eat

\[ 'il-nās \]
the-people

‘I love this fish because it is big (and) eats people.’

Though demonstratives also represent more than half of the responses given by the older age group, more adjectives were produced by this age group than by the younger age group (38% vs. 19%). The sentences in (49) and (50) give some
example answers from the older age group with the age of the participant given in parentheses.

49) Which sun(F) do you like?

\[
dī \quad li’an-h-a \quad farhān-a \quad w-dī \quad za’lān-a \quad (4;3 \text{ girl})
\]

this(F) because-it-F happy-F and-this(F) sad-F

‘This (sun) because it is happy and this one is sad.’

50) Which doctor(M) do you like?

\[
ga \quad li’an \quad ‘hwa \quad ‘inda-h \quad ‘adaw-āt \quad (5;8 \text{ girl})
\]

this(M) because he has-3MS tool-FPI

(This doctor because he has tools.)

As one of the aims of this experiment, in addition to screening the participants towards the main experiment, is to provide a picture of the gender system of Arabic-speaking children in its early stages, the results will be presented first by correct gender attributed by each age group (Fig. 4.2). Then the results will be presented by the two commonalities found between grammatical gender and nouns in Arabic, i.e. word ending (Fig. 4.3) and natural gender (Fig. 4.4). For each participant, the mean of feminine responses attributed to grammatically feminine nouns and grammatically masculine nouns was calculated. Then, the mean of the feminine responses attributed by each age group was calculated.
Table 4.3 shows the mean correct grammatical gender attributed to each of the eighteen nouns by the two age groups. On average, the younger age group attributed more correct responses to grammatically masculine nouns than to grammatically feminine nouns. The older age group were better at assigning the correct gender to the different nouns especially in the case of grammatically feminine nouns.
Table 4.3. Mean correct grammatical gender attributed to the different items by each age group.

<table>
<thead>
<tr>
<th>Noun</th>
<th>Grammatical gender</th>
<th>Mean of correct grammatical gender assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>2 – 3 years</td>
</tr>
<tr>
<td>walad</td>
<td>Masculine</td>
<td>0.84</td>
</tr>
<tr>
<td>bāba</td>
<td>Masculine</td>
<td>0.87</td>
</tr>
<tr>
<td>diktūr</td>
<td>Masculine</td>
<td>0.87</td>
</tr>
<tr>
<td>bint</td>
<td>Feminine</td>
<td>0.65</td>
</tr>
<tr>
<td>mama</td>
<td>Feminine</td>
<td>0.65</td>
</tr>
<tr>
<td>ʿarūs</td>
<td>Feminine</td>
<td>0.59</td>
</tr>
<tr>
<td>fil</td>
<td>Masculine</td>
<td>0.87</td>
</tr>
<tr>
<td>nimir</td>
<td>Masculine</td>
<td>0.81</td>
</tr>
<tr>
<td>ḥiṣān</td>
<td>Masculine</td>
<td>0.91</td>
</tr>
<tr>
<td>baṭṭa</td>
<td>Feminine</td>
<td>0.58</td>
</tr>
<tr>
<td>samaka</td>
<td>Feminine</td>
<td>0.59</td>
</tr>
<tr>
<td>baqara</td>
<td>Feminine</td>
<td>0.53</td>
</tr>
<tr>
<td>kirsi</td>
<td>Masculine</td>
<td>0.81</td>
</tr>
<tr>
<td>bāṣ</td>
<td>Masculine</td>
<td>0.88</td>
</tr>
<tr>
<td>kās</td>
<td>Masculine</td>
<td>0.84</td>
</tr>
<tr>
<td>milʿaga</td>
<td>Feminine</td>
<td>0.59</td>
</tr>
<tr>
<td>siyyāra</td>
<td>Feminine</td>
<td>0.68</td>
</tr>
<tr>
<td>šams</td>
<td>Feminine</td>
<td>0.44</td>
</tr>
</tbody>
</table>
The behaviour of the two age groups is demonstrated in Fig. 4.2 More feminine responses were given to grammatically feminine nouns by the older age group (mean = 0.83, sd = 0.23) than by the younger age group (mean = 0.59, sd = 0.41). Similarly, fewer feminine responses were given to grammatically masculine nouns by the older age group (mean = 0.07, sd = 0.18) than by the younger age group (mean = 0.16, sd = 0.27).

As illustrated in Fig. 4.3, when the nouns ended in -a, more ending consistent responses were given by the older age group (mean = 0.74, sd = 0.21) than by the younger age group (mean =0.54, sd = 0.40). However, when the nouns did not end in -a, the same rate of feminine responses was given by both age groups;
(mean = 0.27, sd = 0.27) and (mean = 0.27, sd = 0.16) by the younger and the older group respectively.

Fig. 4.4 shows that both age groups gave more feminine responses to nouns denoting female persons and more masculine responses to nouns denoting male persons. Yet, more natural-gender-consistent responses were given by the older age group (mean = 0.83, sd = 0.30) in the case of female-person-referring nouns and (mean = 0.08, sd = 0.22) in the case of male-person-referring nouns compared to (mean = 0.61, sd = 0.45) in the case of female-person-referring nouns and (mean = 0.14, sd = 0.29) in the case of male-person-referring nouns by the younger age
group. The rate of feminine responses given to non-persons is not of interest because there is no semantic correlation between non-persons and grammatical gender in Arabic as discussed in chapter 3.

Two by-participant mixed analyses of variance (ANOVA) were carried out to examine the interactions between the different variables. The first analysis was run with age (2-3 vs. 4-5) as a between-subject variable and grammatical gender (masculine vs. feminine) and morphophonological ending (no -a vs. -a) as within-subject variables. No significant three-way interaction was found between age, 

---

57 Separate analyses were carried out because grammatical gender and natural gender were not orthogonal in the design.
ending, and grammatical gender $F(1, 101) = 0.20, p = 0.65$. There were two-way significant interactions between morphophonological ending and grammatical gender $F(1, 101) = 6.04, p = 0.02$ and between age and grammatical gender $F(1, 101) = 22.22, p < 0.0001$. There was also a significant main effect of grammatical gender $F(1, 101) = 285.52, p < 0.0001$. The interaction between age and morphophonological ending, the main effect age, and the main effect of morphophonological ending were not significant.

The second analysis was run with age (2-3 vs. 4-5) as a between-subject variable and natural gender (male vs. female)$^{58}$ and morphophonological ending (no -a vs. -a) as within-subject variables. Similar to the first analysis, there was no significant three-way interaction between age, natural gender, and morphophonological ending $F(1, 100) = 0.00, p = 0.97$. The interaction between age and natural gender was significant $F(1, 100) = 8.16, p = 0.005$. There was also a main effect of natural gender $F(1, 100) = 211.95, p < 0.0001$. The two-way interactions between age and morphophonological ending and between age and ending as well as the main effects of age and morphophonological ending were not significant.

Since the second and more important aim of this experiment is to filter the participants towards the main experiment, the results of those participants who

---

$^{58}$ Non-human nouns were not included in the analysis because of the lack of a semantic cue to grammatical gender in this group of nouns.
passed the experiment above chance level\textsuperscript{59} will now be presented by the two commonalities (noun ending and natural gender). Only 12 out of the 50 (2;6 - 3;11) children who participated in this experiment passed the task above chance level. The children who passed the task above chance level ranged in age between 2;11 and 3;11 (mean = 3;3, sd = 0.43). The children who failed the task ranged in age between 2;6 and 3;9 (mean = 2;10, sd = 0.55). Only 10 out of the 12 participants who passed this task above chance level also passed the MGIT test meaning that only those 10 participants passed the screening criteria towards Experiments II and III. Considering the older age group (4;0-5;11), 61 out of the 76 participants in this group passed the Screening experiment above chance level all of whom also passed the MGIT successfully. Those participants ranged in age between 4;0 and 5;11 (mean = 4;10, sd = 0.58). The 15 children who failed the task ranged in age between 4;2 and 5;8 (mean = 4;8, sd = 0.61).

Fig. 4.5 shows that both age groups gave a high rate of feminine responses when the nouns ended in \textit{-a}; (mean = 0.80, sd = 0.10) by the older age group and (mean = 0.83, sd = 0.10) by the younger age group. They also gave similar rates of feminine responses to the nouns that did not end in \textit{-a}; (mean = 0.25, sd = 0.07) and (mean = 0.27, sd = 0.12) by the older and the younger age groups respectively.

\textsuperscript{59} 14 out of the 18 items is the minimal number of items each participant needs to assign the correct grammatical gender to in order to pass this experiment. This was determined on the basis of a chi-square test for an 18-item binary response task.
As illustrated in Fig. 4.6., both age groups gave a very high rate of feminine responses to the nouns referring to female persons and a very low rate of feminine responses to the nouns referring to male persons. To explain, the younger age group gave (mean = 0.90, sd = 0.16) feminine responses to female-person-referring nouns and (mean = 0.10, sd = 0.22) feminine responses to male-person-referring nouns. Similarly, the older age group gave (mean =0.92, sd = 0.20) feminine responses to female-person-referring nouns and (mean = 0.03, sd = 0.10) feminine responses to male-person-referring nouns.
Two mixed analyses of variance (ANOVA) were also carried out to examine the interactions between the different variables in the behaviour of the participants who passed the task above chance level. The first analysis was run with age (2-3 vs. 4-5) as a between-subject variable and grammatical gender (masculine vs. feminine) and morphophonological ending (no -a vs. -a) as within-subject variables. No significant three-way interaction was found between age, ending, and grammatical gender $F(1, 68) = 0.02, p = 0.89$. There was a two-way significant interaction between morphophonological ending and grammatical gender $F(1, 68) = 4.96, p = 0.02$. There were also a significant main effect of grammatical gender $F(1, 68) = 952.82, p < 0.0001$ and a significant main effect of morphophonological
ending $F(1, 68) = 9.46, p = 0.003$. The interaction between age and morphophonological ending, the interaction between age and grammatical gender, and the main effect of age were not significant.

The second analysis was run with age (2-3 vs. 4-5) as a between-subject variable and natural gender (male vs. female) and morphophonological ending (no -$a$ vs. -$a$) as within-subject variables. There was no significant three-way interaction between age, natural gender, and morphophonological ending $F(1, 68) = 1.50, p = 0.22$. The only two-way significant interaction was between age and ending $F(1, 68) = 5.43, p = 0.02$. There was also a significant main effect of natural gender $F(1, 68) = 487.19, p < 0.0001$. The two-way interactions between age and natural gender and between natural gender and morphophonological ending as well as the main effects of age and morphophonological ending were not significant.

Before moving to the discussion, I will present the types of errors produced by the participants of the two age groups in this experiment. As Table 4.4 demonstrates, both age groups attributed more masculine responses to grammatically feminine nouns than feminine responses to grammatically masculine nouns. The majority of the errors (76.9%) produced by the younger age group were masculine demonstratives with grammatically feminine nouns as the examples in (51) and (52) illustrate. In very few cases, they produced other types of agreeing elements like the one in (53) where a masculine adjective was used to describe a grammatically feminine noun. Only 11.3% of the errors produced by the younger
participants were feminine agreeing forms with grammatically masculine nouns as in example (54).

51) Which sun(F) do you like?

ʾaḥib šams ḡa
1S-love sun(F) this(M)
‘I love this sun.’

52) Which duck(F) do you like? (2;9 boy)

ʾāna ʾaḥib ḡa
I 1S-love this(M)
‘I love this one.’

53) Which spoon(F) do you like?

ʾil-kibīr (3;8 boy)
The-big
‘The big one.’

54) Which bus(M) do you like?

ʾil-bāṣ ʾil-kibīr-a ʾašān ta-mši (2;10 girl)
The-bus the-big-F because 3FS-walk
‘The big bus because it moves.’
The older age group produced comparable percentages of masculine demonstratives (49.4%) and masculine adjectives (37.2%) with grammatically feminine nouns. Examples (55) and (56) show those types of errors. Only in 10.6% of the errors produced by this age group, feminine forms were used with grammatically masculine nouns as the example in (57) shows.

55) Which mother(F) do you like?

\[
\begin{align*}
\text{mama} & \quad \text{il-‘ahmar} \\
\text{mother} & \quad \text{the-red}
\end{align*}
\]

‘The red mother.’

56) Which fish(F) do you like?

\[
\begin{align*}
\text{hāda} & \quad \text{samaka} \\
\text{this(M)} & \quad \text{fish(F)}
\end{align*}
\]

‘This fish.’
57) Which tiger(M) do you like?

ʾil-kibīr-a (5:8 girl)

The-big-F

(The big one.)

Table 4.4. The types of errors produced by the children in experiment I.

<table>
<thead>
<tr>
<th>Grammatical gender assigned</th>
<th>Percentage of each assigned gender by age</th>
<th>Type of errors produced</th>
<th>Percentage of each type of error by age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Masculine gender to feminine nouns</td>
<td>2 - 3</td>
<td>4 - 5</td>
<td>Masculine demonstrative with -a ending female human</td>
</tr>
<tr>
<td></td>
<td>88.7%</td>
<td>89.4%</td>
<td>Masculine demonstrative with no -a ending female human</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Masculine demonstrative with no -a ending non-human</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Masculine demonstrative with -a ending non-human</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Masculine adjective with -a ending non-human</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Masculine adjective with -a ending female human</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Masculine adjective with no -a female human</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Masculine adjective with no -a ending non-human</td>
</tr>
<tr>
<td></td>
<td>Masculine verb with -a ending non-human</td>
<td>1.4%</td>
<td>/</td>
</tr>
<tr>
<td>------------------------</td>
<td>----------------------------------------</td>
<td>------</td>
<td>----</td>
</tr>
<tr>
<td></td>
<td>Masculine verb with no -a ending non-human</td>
<td>/</td>
<td>2.7%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Feminine gender to masculine nouns</th>
<th>Feminine demonstrative with -a ending male human</th>
<th>1.4%</th>
<th>/</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Feminine demonstrative with no -a ending male human</td>
<td>1.4%</td>
<td>2.7%</td>
</tr>
<tr>
<td></td>
<td>Feminine demonstrative with non-human masculine</td>
<td>4.2%</td>
<td>5.3%</td>
</tr>
<tr>
<td></td>
<td>Feminine adjective with non-human masculine</td>
<td>3.5%</td>
<td>2.7%</td>
</tr>
<tr>
<td></td>
<td>Feminine verb with non-human masculine</td>
<td>0.7%</td>
<td>/</td>
</tr>
</tbody>
</table>

4.4. Discussion

Two screening criteria were used to filter the participants towards the main experiment. The first criterion was the MGIT to ascertain that all the participants who will take part in experiments II and III have already acquired ‘gender identity’ and ‘gender constancy’ as cognitive concepts. That was a crucial step in order to pave the ground for the argument. To put it differently, cognitively distinguishing between males and females is a prerequisite for using such a distinction in noun classification. Hence, to argue that young children are biased to use morphophonological information and that they are not sensitive to natural gender as a cue in noun gender classification (Culbertson et al., 2019; Gagliardi & Lidz, 2014; Karmiloff-Smith, 1979; Levy, 1983; Pérez-Pereira, 1991, for example), it is
important to find out first whether those children have the cognitive concept required for that or not. As outlined by Schütze (2009), one of the potential sources for agreement errors in child language is the lack of the conceptual or perceptual knowledge required for producing this type of agreement. A child for example might fail to produce number agreement because they do not yet make the conceptual distinction between one and more than one. The same applies for gender agreement. The child may fail to produce the correct gender agreement because they do not yet make the distinction between males and females. As per the results of MGIT presented in section (4.2.2), we can conclude that ‘gender identity’ and ‘gender constancy’ are acquired by children around the age of three\textsuperscript{60}.

The second criterion was the semi-spontaneous production experiment with familiar nouns (Experiment I). One of the aims of this experiment was to find out how grammatical gender looks starting from the two-word utterance stage up to pre-school age. Therefore, the responses of all the participants in addition to the participants who passed the experiment above chance level were presented and analysed. The results based on all participants are valuable in and of themselves because they offer insights into whether children treated the two gender classes similarly or they did better with one gender class compared to the other. The results revealed a significant effect of gender class as the children were attributing more

\textsuperscript{60} Notice that this conclusion should not be taken for granted; there were children as old as 3;6 and 4;5 who failed MGIT. Therefore, finding out that the same individuals who will take part in the next experiments have also passed MGIT successfully is highly recommended due to the individual differences between children.
masculine responses to grammatically masculine nouns than feminine responses to grammatically feminine nouns (see Table 4.3). The same behaviour is also reported for L2 learners of Arabic (see Alhawary, 2009; Alamry, 2014) and for heritage Arabic speakers (Albirini, 2013). They do better with masculine nouns than with feminine nouns. The behaviour of both L1 Arabic-learning children and L2 Arabic-learning adults gives support to learning masculine gender as a default before learning the feature specifications of the feminine gender (Neeleman and Weerman, 1999). There is a caveat to this interpretation, however, since the feature underspecification of the masculine gender in Arabic is confounded by the fact that this gender is morphophonologically unmarked. It could be that masculine gender agreement is acquired before feminine gender agreement because masculine gender agreement is morphophonologically unmarked in Arabic.

The responses of all the participants also provide information about the types of errors produced by the children in their early language development stages. Two types of errors were attested in the children’s responses in this experiment:

a) masculine gender attributed to grammatically feminine nouns; and

b) feminine gender attributed to grammatically masculine nouns.

The first type of errors was more common than the second type (see Table. 4.4). Several possible reasons could be behind the patterns of the responses and the errors attested in this experiment. One possible reason for that is the way care-givers talk to the young children. In (Alqattan, 2015) corpora of Kuwaiti Arabic for example, some care-givers used masculine demonstratives most of the time when asking their
children about things, persons, or animals whether they were asking about grammatically masculine or feminine items\textsuperscript{61}. Haggan (2002) reported that Kuwaiti adults modify and simplify their language when talking to young children. Ferguson (1956) and Omar (1973) also found out that (Syrian in the former and Egyptian in the latter) Arabic baby talk has no plural or gender agreement markers. Whether the tendency to use the masculine demonstrative by Arabic adults is due to the fact that the words for thing (šay), animal (ḥaywān), person (šaxṣ) are all masculine in Arabic or whether this has to do with the modifications that caretakers do when they try to communicate with children (Haggan, 2002), is something that needs further investigation. Another potential reason is that masculine gender is the default underspecified gender that could be acquired earlier by Arabic L1 learners and used before acquiring the feature specifications of the feminine gender or in cases where the gender of a particular noun is not yet known as I just mentioned above. Children’s behaviour in this experiment cannot be due to the lack of sufficient input necessary for learning the grammatical gender of some of the items used in the experiment because if that was the case, we would not expect a bias in the answer pattern attested in the results. In other words, if the pattern attested in the results is simply due to the patterns children are exposed to in the input language, this would imply that children have sufficient input for learning the gender of the masculine items but not the feminine items. This cannot be the case.

\textsuperscript{61} This should not be taken as something special about Kuwaiti Arabic. From my personal experience, I can say that some of the members in Alahsa community follow the same pattern when talking to young children.
in HA (the variation of Arabic under investigation in this project) because feminine agreeing forms should be as ubiquitous as masculine agreeing forms as feminine forms are used by some speakers with inanimate plurals\textsuperscript{62}.

Of additional interest is grammatical gender correlates. As discussed in chapter 3, grammatical gender in Arabic, as in many other languages, correlates with word ending and natural gender. Though the effect of these correlations is clear in the responses of all the participants, it is manifested more clearly in the performance of the children who passed the experiment above chance level\textsuperscript{63}. The use of the nouns internal properties (semantic and morphophonological) in assigning gender to them suggests that the association between those inherent properties and grammatical gender are internalized by the children as part of the gender system. The low rate of feminine responses given to the word šams (sun) by both age groups (see Table 4.2.) demonstrates very clearly the association that children make between word ending and grammatical gender in Arabic. To explain, šams is a grammatically feminine noun that lacks an -a ending, which is not the case for most of the non-human feminine nouns in Arabic. Though ʿarūs (bride) and bint (girl) also lack the feminine ending, they were given very high rate of feminine responses especially by the older age group. This shows that both the

\textsuperscript{62} Note that there is a high degree of variation among HA speakers on the forms of agreement used with inanimate nouns as discussed in Chapter 3.

\textsuperscript{63} Note that when the mean responses of all the participants and the mean responses of the participants who passed the experiment above chance level were compared, the differences were found significant in all the cases except in the correlation between grammatical gender and natural gender in the younger age group (\(P = 0.05\)).
semantic and the morphophonological properties of the nouns are entertained by the children, albeit more clearly by the older children, when attributing gender to nouns.

This stage does not necessarily mark the discovery of the gender system, i.e. that Arabic has a gender system in the first place, which could have been discovered earlier, but it marks the system in use as it is manifested in the production of the very young children. The correlation between grammatical gender and word ending as well as between grammatical gender and natural gender seems to be discovered earlier than 2;6-3;0 years. Yet, those correlations require sometime to be integrated as a system keeping in mind the individual differences between the children. That is, some children might discover and build the system earlier than others. This is confirmed by the fact that there were children from both age groups who could not pass Experiment I above chance level, though that was more common among the younger participants than among the older participants.

One comment is in order here: children’s performance in Experiment I does not necessarily reflect the underlying abstract representation of their gender system. Their performance could be due to their rote learning (Pizzuto & Caselli, 1994; Tomasello, 2003). It is fairly likely that a child produced, for instance, feminine adjectives/demonstratives/verbs when talking about the noun ‘ʾbint’ (girl) because they learned by rote that ‘bint’ is used with these forms, because that is how the child encountered this noun in the speech they are hearing not because the child has discovered that there is a correlation between feminine natural gender and feminine
grammatical gender. Similarly, the child probably described the word ‘ḥiṣa:n’ (horse) with a masculine agreeing form because that is how they learned this word not because they have discovered that the absence of –a or the presence of endings other than -a is a cue to masculine gender in Arabic. Thus, the children’s behaviour in this experiment reflects their production patterns but not necessarily their sensitivity to the different cues for gender. In order to gain better insight into the abstract representations of gender in its early stages, nonce nouns will be used in Experiment II (chapter 5) and Experiment III (chapter 6). Nonetheless, the aim of Experiment I was not to find out what characterizes the abstract representations of the gender system in the early stages. Rather, this experiment was used to test children’s knowledge of gender agreement and to ascertain that they can alternate between masculine and feminine agreeing forms when talking about familiar items. Basically, this experiment was used to minimize, to the best possible extent, the possibility that children’s agreement patterns in the subsequent experiments could be due to the lack of the required syntactic knowledge.

Overall, the different performance of the two age groups shows the developmental trajectory that learning grammatical gender follows as children grow from the two-word utterance stage to pre-school age. The pre-schoolers produced more correct gender-agreeing forms than their younger pairs.
Chapter 5: Experiment II: Cue interaction

5.1. Introduction

This chapter aims at drawing a comprehensive picture of the developmental trajectory of grammatical gender in Arabic starting from the two-word utterance to adulthood. It presents an experiment that investigates the interaction between the noun-internal (morphophonological and semantic) information and the noun-external gender-defining information, i.e. agreement. This experiment tries to measure the weight given by the learners to the different types of information at different developmental stages. Many previous studies examined how children understand grammatical gender and how they use various cues to determine gender at a particular age (Brooks et al., 1993; Culbertson et al., 2019; Gagliardi & Lidz, 2014; Karmiloff-Smith, 1979; Levy, 1983; Pérez-Pereira, 1991; Szagun et al., 2007, inter alia). However, there is no clear picture of how cue preference develops from morphophonology to semantics and how the cues interact over the course of development. By using the same experimental task, I will first describe the representation of gender in its end state as reflected in the behaviour of the adult participants. Once this is done, I will be able to trace the developmental trajectory that this system undergoes until it reaches its end state.

Though many studies have addressed this issue, it is hard to find a developmental study that considered all the types of cues or that showed a complete

---

A condensed version of this chapter was published in the proceeding of the 46th annual Boston University Conference on Language Development BUCLD as Alali (2022).
To mention some examples, Pérez-Pereira (1991) has a study design quite similar to the one used here. However, the participants in his experiment were between 4 and 11 years only. Also, in Pérez-Pereira’s study, every novel noun was used with a fixed combination of cues as if the grammatical gender of the noun was presupposed in the experiment. That presupposition was clear in the general result table (Table 5, p. 581) as there was an expected gender for each novel item according to which the results were explained. For example, the item ‘lodena’ was always presented with a feminine determiner and a character that has no natural gender indication. It is also not clear on what basis the expected gender was decided. With some items, it is the gender defined by the article while with other items, it is the gender suggested by the noun ending. Furthermore, the method used for analysing the results of the experiment was a simple comparison of the observed gender assignment distribution to the expected distribution. This approach does not reveal the factors contributing to the developmental changes. By contrast, my analysis employed a multi-factor approach that identifies changes in the cues across ages.

Gagliardi and Lidz’s (2014) experiment had participants of many age groups (4 - 7), (8 - 12) and adults; however syntactic agreement was not a variable in their experiment. Though syntactic agreement has a status different from the semantic and morphophonological properties of the noun as it provides a noun-external cue to gender, including agreement is important when studying the acquisition of gender. It is through agreement that grammatical gender manifests
itself. Therefore, syntactic agreement was included as a variable in this experiment to find out how this defining element of grammatical gender interacts with the weight that L1 learners/speakers ascribe to morphophonological and semantic cues for gender. If, as argued by Monaghan et al. (2005), distributional cues are used by the learners to categorize high-frequency nouns whereas morphophonological cues are used to categorize low-frequency nouns, we would predict that the bias to use morphophonological cues in noun class learning will not be overcome by the syntactic cues in this experiment. All the nouns here are novel nouns that have never been heard by the participants, i.e. they are low-frequency nouns.

Furthermore, in most of the studies where syntactic agreement was a variable, the syntactic cue was provided by an article (Karmiloff-Smith, 1979; Pérez-Pereira, 1991). Arabic does not have a closed class of indefinite articles that obligatorily co-occur with the nouns and though Arabic has a closed class of demonstratives that are used with the nouns, nouns are mostly used with no accompanying elements which makes the possibility of learning demonstrative-noun combinations as chunks unlikely in Arabic. Also, the most frequently co-occurring gender agreeing forms, namely adjectives, do not constitute a closed class of items. Rather, every noun can be described by many adjectives. This

---

65 Szagun et al. (2007) is more problematic because children’s assignment of grammatical gender was determined by the correct use of the article in the naturalistic speech corpora. Given the obligatoriness of the articles in languages like German, French, and Spanish and the controversy around learning article + noun combinations (Pine et al., 2013 and Valian et al., 2009 inter alia), interpreting the patterns of gender assignment found among children in natural speech is problematic. It is worth mentioning that Karmiloff-Smith (1979) and Pérez-Pereira (1991) did not have this issue because novel nouns were used in their experiments.
characteristic of Arabic minimizes the confounding effect that determiner + noun combinations might be learned as whole chunks and any link might be established between the article and the noun ending. In my design, the novel nouns were presented in a sentence that contained a demonstrative and an adjective. Thus, this experiment was designed to provide the pieces missing from some studies and to overcome the drawbacks found in others.

5.2. Method

The ultimate goal of this study is to draw a comprehensive picture of the abstract representation of the grammatical gender system over the course of development, making inferences about what learners/speakers know on the basis of what they produce (Lust, 2006). Therefore, the experimental design used novel nouns. Novel nouns provide accurate insights into language learners/speakers’ ability to productively generalize their existing knowledge to newly encountered items. Using novel nouns also eliminates the possibility that some items have been encountered or learned by some participants but not others and makes the exposure period equivalent among all the participants. Furthermore, following this experimental method gives the opportunity for the results to be compared to results obtained from other languages.
5.2.1. Participants

The participants were 202 typically-developing Hasawi Arabic speakers of five age groups. The youngest age group consisted of seven toddlers (5 females) aged 2;6-3;11 (Mean age: 3;3) who passed the two Screening criteria (Experiment I and MGIT) presented in chapter 4. Though it was found that children are able to track gender regularities in their languages even before they start producing a minimum length of two-word utterances (Cyr & Shi, 2013; Mariscal, 2009; Van Heugten & Christophe, 2015), the method I used in my investigation required a minimum length of two-word utterance. For that reason, two-and-a-half-year-old children were the youngest participants to take part in the study. The second group consisted of forty-seven (29 females) pre-schoolers aged 4;0-5;11 (Mean age: 4;9) who also passed the screening criteria. The third group consisted of fifty-two (35 females) school children aged 6;0-10;11 (Mean age: 8;2). The fourth group consisted of forty-five (22 females) adolescents aged 11;0-15;11 (Mean age: 12;9). The last group consisted of fifty-one (26 females) adults aged 16 years and over (Mean age: 25;3). All the participants were local residents in Ahsa (a city in the eastern province in Saudi Arabia).

Twenty-six additional children (three 2-3, fourteen 4-5, nine 6-10), five adolescents, and four adults were tested but excluded from the final analyses due to failure to understand the task. To explain, some of those participants were giving masculine forms of the colours of the characters across the board because they were
referring to the word ‘colour’ which is masculine in Arabic.\textsuperscript{66} This was explicitly reported by the excluded adults and adolescents when they were asked about their responses after the experiment. Other children overgeneralized the feminine demonstrative ‘(hā)ḏi’ to all the items in the experiment because they were referring to the hand with which the experimenter was holding the item, not to the item itself. The word ‘hand’ is feminine in (Hasawi) Arabic and requires a feminine demonstrative.\textsuperscript{67} Some children were simply pointing to the hidden item and no verbal responses could be elicited from them. Finally, one 4-year-old girl was describing all the experimental items using the feminine form of the word (sweet) ‘ḥilw-a’ in Arabic.

5.2.2. Materials
The verbal stimuli consisted of 24 novel nouns (see Table 5.1.), which were created in accordance with Hasawi Arabic phonotactics.\textsuperscript{68} The same nouns were used with -a ending and without -a ending to form two patterns; half of the participants had nouns (1-12) with -a and nouns (13-24) without -a, the other half of the participants

\textsuperscript{66} Mariscal (2009) also found that his four children between 1;10 and 2;0 tended to produce masculine adjectives to describe feminine nouns when they were asked ‘What colour is + NP?’ because the word ‘colour’ is masculine in Spanish. In my experiments, the excluded participants were giving masculine responses across the board even though I did not use the word ‘colour’ in my questions.

\textsuperscript{67} This is familiar to the children because one of the games that parents usually play with their children is holding something in one hand and putting both hands behind the parent’s back and asking the child in which hand is the hidden thing. It was clear from the excluded children’s responses that they were thinking that I was playing the same game with them.

\textsuperscript{68} Unlike in standard Arabic where consonant clusters are inhibited in onset positions (Abu-Salim, 1982), HA does allow consonant clusters as argued by Altaisan (2022).
had nouns (1-12) without -a and nouns (13-24) with -a. Moreover, the nouns were assigned randomly to the different cue combinations (see the twelve conditions mentioned in the design below) to control for item effect. In addition to those nouns, the following two sentences were used to provide the subjects with the linguistic context that serves as the defining element of the gender of the nouns. The demonstratives and adjectives are real Arabic words.

- *hāḏi* X *wa* hāḏi X *ṭāny-*a\(^69\)

  this(F) X and this(F) X second-F

  ‘This is X and this is a second X.’

- *hāḏā* X *wa* hāḏā X *ṭānī*

  this(M) X and this(M) X second(M)

  ‘This is X and this is a second X.’

\(^69\) *ṭānī/ṭāny-*a are adjectives in Arabic.
Table 5.1. List of novel nouns used in Experiment II.

<table>
<thead>
<tr>
<th></th>
<th>-a ending novel nouns</th>
<th>Non -a ending novel nouns</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>fādir-a</td>
<td>fādir</td>
</tr>
<tr>
<td>2</td>
<td>nūšiy-a</td>
<td>nūšī</td>
</tr>
<tr>
<td>3</td>
<td>šandal-a</td>
<td>šandal</td>
</tr>
<tr>
<td>4</td>
<td>ġādiy-a</td>
<td>ġādī</td>
</tr>
<tr>
<td>5</td>
<td>da ’būs-a</td>
<td>da ’būs</td>
</tr>
<tr>
<td>6</td>
<td>šnīf-a</td>
<td>šnīf</td>
</tr>
<tr>
<td>7</td>
<td>lqeim-a</td>
<td>lqeim</td>
</tr>
<tr>
<td>8</td>
<td>šmeil-a</td>
<td>šmeil</td>
</tr>
<tr>
<td>9</td>
<td>zammāl-a</td>
<td>zammāl</td>
</tr>
<tr>
<td>10</td>
<td>šahrūr-a</td>
<td>šahrūr</td>
</tr>
<tr>
<td>11</td>
<td>zāš-a</td>
<td>zāš</td>
</tr>
<tr>
<td>12</td>
<td>fawān-a</td>
<td>fawān</td>
</tr>
<tr>
<td>13</td>
<td>la ’būs-a</td>
<td>la ’būs</td>
</tr>
<tr>
<td>14</td>
<td>ḥatr-a</td>
<td>ḥatar</td>
</tr>
<tr>
<td>15</td>
<td>nāš-a</td>
<td>nāš</td>
</tr>
<tr>
<td>16</td>
<td>z’ēif-a</td>
<td>z’ēif</td>
</tr>
<tr>
<td>17</td>
<td>fleil-a</td>
<td>fleil</td>
</tr>
<tr>
<td>18</td>
<td>mandar-a</td>
<td>mandar</td>
</tr>
<tr>
<td>19</td>
<td>šbeil-a</td>
<td>šbeil</td>
</tr>
<tr>
<td>20</td>
<td>šbeib-a</td>
<td>šbeib</td>
</tr>
<tr>
<td>21</td>
<td>kammūš-a</td>
<td>kammūš</td>
</tr>
<tr>
<td>22</td>
<td>sbeim-a</td>
<td>sbeim</td>
</tr>
<tr>
<td>23</td>
<td>ġādm-a</td>
<td>ġādīm</td>
</tr>
<tr>
<td>24</td>
<td>šleiw-a</td>
<td>šleiw</td>
</tr>
</tbody>
</table>
In addition to the verbal stimuli, 24 pairs of 3D characters of two different colours\textsuperscript{70} or sizes, which either have indications as being male/female humans or nonhumans were created. Eight characters represented male-like humans (having moustaches and/or short hair), eight represented female-like humans (wearing dresses, skirts, having long hair or bows), and eight represented non-humans (Martian like characters that do not have arms or legs, have three or more eyes, have a hollow body). Fig. 5.1 shows examples of the 3D characters used\textsuperscript{71} (see Appendix F for pictures of the full set of the 3D characters used in this experiment. 3D characters rather than pictures were used to ascertain that the participants were describing the characters themselves and not the pictures, which might affect their answers.

\textsuperscript{70} This is inspired by Karmiloff-Smith (1979).

\textsuperscript{71} The 3D characters were further evaluated by native Arabic speakers to make sure that they are classified as they were intended to be in the experiment (see the results section below).
In addition to the test items, four 3D lifelike animal figures; two feminine and two masculine (a cow, a duck, a horse, and an elephant) were used as fillers.

5.2.3. Design
A factorial design \((2 \times 2 \times 3)\), noun ending \((-a\) versus no \(-a\)) \(\times\) agreement (masculine versus feminine) \(\times\) natural gender (male, female, or non-human), was used to get all the possible combinations that an Arabic-speaking learner might encounter in the input.\(^{72}\) This results in the following 12 conditions:

---

\(^{72}\) This design is similar to the design of Pérez-Pereira’s (1991).
1. masculine agreement + -a ending + male human
2. *masculine agreement + -a ending + female human
3. masculine agreement + -a ending + non-human
4. masculine agreement + non -a ending + male human
5. *masculine agreement + non -a ending + female human
6. masculine agreement + non -a ending + non-human
7. *feminine agreement + -a ending + male human
8. feminine agreement + -a ending + female human
9. feminine agreement + -a ending + non-human
10. *feminine agreement + non -a ending + male human
11. feminine agreement + non -a ending + female human
12. feminine agreement + non -a ending + non-human

As indicated by the asterisk (*), some combinations are impossible in Arabic. Nevertheless, those combinations were included to see how participants behave with regard to such conditions.

5.2.4. Procedure
The experiment was carried out in a quiet room in the experimenter’s house or in the participant’s house whichever was more convenient for the participant. Each participant was interviewed individually by the experimenter except for two (2 - 3) children and three (4 - 5) children who were shy to be left with the experimenter alone and therefore they were accompanied by their mothers. The mothers were
instructed to keep silent throughout the experiment. The answers were recorded using a Zoom recorder.

Each of the 12 conditions was tested twice in the following scenario:

The experimenter explained the task to the participants by saying “we are going to play a game in which you will have to describe some characters.” Then, the experimenter presented the participant with an item and said “This is X” and then the participant was asked “who is this?” to make sure that the participant got the novel noun. And then the experimenter said “and this is a second X” holding the pair of 3D characters. Then the experimenter repeated “This is X and this is a second X” while holding the pair. Then one of the items was hidden and the participant was asked “which X did I hide?” The gender of the demonstratives and the adjectives were manipulated depending on the condition being tested.

5.3. Results

The dependent variable was grammatical gender (masculine/feminine) attributed to each item. The grammatical gender was determined on the basis of agreement, i.e. the agreeing forms produced by the participant.

For each participant, the mean rate of feminine responses attributed in each condition was calculated. Then, the mean rate of feminine responses attributed by each age group in each condition was calculated (Table 5.2) and (Fig. 5.2).

73 The Arabic script for this experiment with example answers are given in Appendix B.
Table 5.2. The mean rate of the feminine responses attributed by the five age groups in each condition.

<table>
<thead>
<tr>
<th>Condition</th>
<th>2;6 - 3</th>
<th>4 - 5</th>
<th>6 - 10</th>
<th>11 – 15</th>
<th>Adults</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 M agreement + -a + male human</td>
<td>0.64</td>
<td>0.31</td>
<td>0.14</td>
<td>0.09</td>
<td>0.04</td>
</tr>
<tr>
<td>2 M agreement + -a + female human</td>
<td>0.86</td>
<td>0.64</td>
<td>0.69</td>
<td>0.69</td>
<td>0.69</td>
</tr>
<tr>
<td>3 M agreement + -a + non-human</td>
<td>0.86</td>
<td>0.41</td>
<td>0.26</td>
<td>0.20</td>
<td>0.21</td>
</tr>
<tr>
<td>4 M agreement + no -a + male human</td>
<td>0.07</td>
<td>0.00</td>
<td>0.08</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>5 M agreement + no -a + female human</td>
<td>0.21</td>
<td>0.22</td>
<td>0.35</td>
<td>0.44</td>
<td>0.37</td>
</tr>
<tr>
<td>6 M agreement + no -a + non-human</td>
<td>0.07</td>
<td>0.04</td>
<td>0.07</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>7 F agreement + -a + male human</td>
<td>0.79</td>
<td>0.70</td>
<td>0.67</td>
<td>0.59</td>
<td>0.68</td>
</tr>
<tr>
<td>8 F agreement + -a + female human</td>
<td>0.93</td>
<td>0.81</td>
<td>0.89</td>
<td>0.94</td>
<td>0.97</td>
</tr>
<tr>
<td>9 F agreement + -a + non-human</td>
<td>1.00</td>
<td>0.84</td>
<td>0.86</td>
<td>0.81</td>
<td>0.89</td>
</tr>
<tr>
<td>10 F agreement + no -a + male human</td>
<td>0.43</td>
<td>0.45</td>
<td>0.43</td>
<td>0.41</td>
<td>0.39</td>
</tr>
<tr>
<td>11 F agreement + no -a + female human</td>
<td>0.86</td>
<td>0.73</td>
<td>0.86</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td>12 F agreement + no -a + non-human</td>
<td>0.50</td>
<td>0.56</td>
<td>0.75</td>
<td>0.70</td>
<td>0.72</td>
</tr>
</tbody>
</table>
Fig. 5.2. Mean rate of feminine responses by age × word ending × agreement × natural gender.
The way the cues (semantic, morphophonological, and syntactic) provided in each condition interact varies among the different age groups. A number of interesting patterns emerge from this analysis:

- The youngest participants assigned feminine gender disproportionately when the items ended in the feminine marker -a.
- Though the absence of -al the presence of the masculine endings resulted in assigning masculine gender by the youngest participants, this was not as predominant as the assignment of feminine gender when the -a ending was present.
- The other four groups were not biased by the morphophonological endings. The pattern of their responses varied depending on the other cues (semantic/syntactic).
- More natural gender-consistent responses are assigned as the participants progress in age.
- Syntactic agreement seems to play an important role in the gender attributed to the novel items by all the age groups. However, the effect of syntactic agreement seems to be conditioned by the morphophonological ending in the youngest age group.

To tease apart the developmental effect of each type of cue, I summarised the data by age and syntactic agreement, by age and morphophonological ending,
and by age and natural gender. As can be seen in Fig. 5.3, the mean rate of feminine responses attributed to the novel nouns when the agreement in prompt was feminine is nearly the same across all the age groups ranging between (0.68) by the pre-schoolers and (0.77) by the adults. On the other hand, the rate of feminine responses when the agreement in prompt was masculine decreased (i.e. the rate of masculine responses increased) as participants grew older. In the latter case, the mean rate of feminine responses decreased from (0.45) by the youngest age group to (0.22) by the adults.

![Fig. 5.3](image)

Fig. 5.3. Mean rate of feminine responses by age and syntactic agreement in prompt. Error bars represent standard errors of the mean (SEM).

Fig. 5.4 shows that the highest rate of feminine responses attributed to -a-ending nouns was by the youngest age group. The association between -a and
feminine gender decreased gradually as participants grew older. It decreased from (0.85) by the youngest age group to (0.55) by the adolescents. On the other hand, the rate of feminine responses attributed to novel nouns when they did not end in -a was very similar among all the age groups ranging between (0.34) by the preschoolers and (0.42) by the school children and the adolescents.

Fig. 5.4. Mean rate of feminine responses by age and noun ending. Error bars represent standard errors of the mean (SEM).

Fig. 5.5 shows the interaction between age and natural gender information when the novel nouns were represented as male vs. female humans. The mean rate of feminine responses attributed to female-human-like characters was very high among all the age groups ranging between (0.76) buy the adolescents and (0.60) by
the pre-schoolers. Similarly, the mean rate of the feminine responses attributed to male-human-like characters was relatively low by all the age groups ranging between (0.48) by the youngest age group and (0.27) by the adolescents.

Fig. 5.5. Mean rate of feminine responses by age and natural gender. Error bars represent standard errors of the mean (SEM).

A by-participant mixed analysis of variance (ANOVA)\(^74\) was run with age (2-3, 4-5, 6-10, 11-15, and adults) as a between-subject variable and

---

\(^74\) I tried to run a mixed-effect model using the `glmer` function in R but the model failed to converge. It seems that my experimental design is too complex (4-way interaction) for the amount of data collected. I tried to simplify the model so it could converge. The successful model did not have a four-way interaction between age, agreement, morphophonological ending, and natural gender and it did not have item as a random effect. This model, the results of which essentially do not differ from the mixed-ANOVA reported here, and its results are included in appendix B.
morphophonological ending (-a vs. no -a), syntactic agreement in the prompt (masculine vs. feminine), and natural gender (male human vs. female human vs. non-human) as within-subject variables. Statistical analyses were performed by participant. A three-way significant interaction was found between word ending, natural gender and agreement $F(2, 386) = 17.46, p < 0.0001$. There were also two-way significant interactions between age and agreement $F(4,193) = 3.43, p = 0.010$, between age and natural gender $F(7, 337.86) = 5.72, p < 0.0001$, between age and word ending $F(4, 193) = 9.05, p < 0.0001$, between natural gender and agreement $F(1.91, 367.95) = 11.63, p < 0.0001$, and between word ending and agreement $F(1, 193) = 17.05, p < 0.0001$. There were also significant main effects of agreement $F(1, 193) = 454.69, p < 0.0001$, natural gender $F(1.75, 337.86) = 153.65, p < 0.0001$, and word ending $F(1, 193) = 202.38, p < 0.0001$.

When the two-way interaction between age and agreement was further investigated, a significant effect of age was found when the agreement in prompt was masculine $F(4, 197) = 3.32, p = 0.011$, but not when it was feminine $F(4, 197) = 0.96, p = 0.432$. Simple pairwise comparisons were run between the different age groups for feminine responses in the case of masculine agreement. A Bonferroni adjustment was applied. The mean of feminine responses was significantly different between the youngest age group and school-age children ($p = 0.042$), between the youngest age group and adolescents ($p = 0.014$), and between the youngest age group and adults ($p = 0.006$).
Considering age and noun ending two-way interaction, an effect of age was found when the nouns ended in -a, $F(4, 197) = 3.59, p = 0.007$. No effect of age was found when the novel nouns did not end in -a, $F(4, 197) = 1.80, p = 0.130$. Pairwise comparisons between the age groups revealed significant differences between the youngest age group and pre-schoolers ($p = 0.034$) between the youngest age group and school-age children ($p = 0.010$), between the youngest age group and adolescents ($p = 0.002$), and between the youngest age group and adults ($p = 0.008$). Further post hoc tests were run to investigate the effect of ending within each age group. One-way within subject ANOVA revealed that the rate of feminine responses attributed when the nouns ended in –a was significantly different from the rate of feminine responses attributed when the nouns did not end –a in all the age groups; $p = 0.002$ in the youngest age group and $p < 0.0001$ in all the other age groups.

Moving to the two-way interaction between age and natural gender, one-way between subjects ANOVA showed a significant main effect of age $F(4, 197) = 3.89, p = 0.004$ in the case of feminine responses attributed to female persons. Pairwise comparisons revealed significant differences in the rate of feminine responses attributed to female persons between pre-schoolers and adolescents ($p = 0.004$) and between pre-schoolers and adults ($p = 0.009$). As illustrated in fig. 5.5., the rate of feminine/masculine responses attributed to novel nouns in the case of male persons is also different between the different age groups $F(4, 197) = 2.98, p = 0.020$. However, pairwise comparisons between the age groups in the case of male
persons revealed no significant differences between the different age groups. No effect of age was found when the referent of novel nouns was non-human.

Furthermore, the effect size of the different cues in each age group was calculated using partial eta-squared test ($\eta_p^2$). The results are given in table 5.3. Partial eta-squared, which is recommended for mixed designs (Richardson, 2011), was chosen because I am interested in finding out the amount of variation in the dependent variable accounted for by each type of information, morphophonological vs. syntactic vs. semantic in each age group when non-error sources of variation are partialed out. The results of the $\eta_p^2$ that was run for the youngest age group shows that morphophonological information accounts for the majority of the variation in the responses of this age group (45%). Considering the pre-schoolers group, morphophonological information accounts for 13%, syntactic information accounts for 26%, and natural gender information accounts for only 7% of the variations in the responses in this age group ($\sum \eta_p^2 = 46$). For the school children, adolescents, and adults, syntactic agreement as well as natural gender information account for great proportions of the variations in the responses while morphophonological information accounts for only a small proportion of the variations in the responses of these age groups.

---

75 Eta squared ($\eta^2$) is also reported for the sake of comparison as a recommended practice in scientific research (Cohen, 1973; Levine & Hullett, 2002).
Table 5.3. The effect size of each type of information in each age group ($\eta^2_p$ and $\eta^2$).

<table>
<thead>
<tr>
<th>Age group</th>
<th>Type of information</th>
<th>$\eta^2_p$</th>
<th>$\eta^2$</th>
<th>$\eta^2_p$</th>
<th>$\eta^2$</th>
<th>$\eta^2_p$</th>
<th>$\eta^2$</th>
<th>$\sum \eta^2_p$ computed</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-3</td>
<td>Morphophonological Ending</td>
<td>0.45</td>
<td>0.34</td>
<td>0.23</td>
<td>0.13</td>
<td>0.11</td>
<td>0.05</td>
<td>79%</td>
</tr>
<tr>
<td>4-5</td>
<td>Syntactic Agreement</td>
<td>0.13</td>
<td>0.09</td>
<td>0.26</td>
<td>0.21</td>
<td>0.07</td>
<td>0.05</td>
<td>46%</td>
</tr>
<tr>
<td>6-10</td>
<td>Natural gender</td>
<td>0.07</td>
<td>0.04</td>
<td>0.37</td>
<td>0.30</td>
<td>0.19</td>
<td>0.12</td>
<td>63%</td>
</tr>
<tr>
<td>11-15</td>
<td></td>
<td>0.05</td>
<td>0.02</td>
<td>0.41</td>
<td>0.31</td>
<td>0.32</td>
<td>0.20</td>
<td>78%</td>
</tr>
<tr>
<td>Adults</td>
<td></td>
<td>0.08</td>
<td>0.04</td>
<td>0.47</td>
<td>0.35</td>
<td>0.30</td>
<td>0.18</td>
<td>85%</td>
</tr>
</tbody>
</table>

Before moving to the discussion section, I need to consider two potential artefacts that could have caused the pattern of the observed results; the 3D prompt and the endings of the no -a words.

➢ The 3D prompts

It is possible that the observed effect of natural gender was there because the 3D prompts were not classified by the participants during the experiment as they were intended to be. To explain, it could be that some participants attributed feminine/masculine gender to some characters because they were referred to with a feminine/masculine demonstrative and/or adjective during the experiment or
because they were represented with a non–āl/ -ā ending nouns not because they represented male/female humans. Therefore, I had all the 16 human-like prompts evaluated by twenty 3-year-old children, twenty (4-5) children and twenty adults who were native speakers of Hasawi Arabic. They were simply asked to classify the characters into two groups: boys/males vs. girls/females. They were also told to leave out any character that was hard to classify.

The results of this evaluation showed that 14 out of the 16 human-like prompts were classified correctly (as they were intended to be) by all the 60 participants. 2 characters however were either put in the wrong group or were left out.

Accordingly, all the test items that were represented using those two prompts were identified. All the responses to those items were marked as unavailable responses (NAs). The data was then reanalysed following the same steps of analysing the data before deleting those responses (see above) to detect any discrepancies that could have been caused by those prompts.

Fig. 5.6 shows the resulting pattern after deleting the “unreliable” responses. This pattern is very similar to the pattern found before deleting those responses and this is confirmed by the results of the mixed analysis of variance (ANOVA) which revealed the same three-way and two-way interactions which were found before deleting those responses.

---

76 I did not ask school-children and adolescents to evaluate the prompts because no significant differences were found between them and between the adults in the interaction between age and natural gender.
When the two-way interaction between age and agreement was reconsidered (Fig. 5.7), the only discrepancy found was that the difference between the youngest age group and school age children in the rate of feminine responses attributed when the agreement in prompt was masculine did not reach statistical significance but approached it very closely ($p = 0.053$), (cf. $p = 0.042$ before deleting the ‘unreliable’ responses).

Fig. 5.6. Mean rate of feminine responses by age × word ending × agreement × natural gender (re-analysed).
Moving to the two-way interaction between age and word ending (Fig. 5.8), post-hoc tests revealed no discrepancies between the results before and after deleting the ‘unreliable’ responses.
Finally, Fig. 5.9 shows the two-way interaction between age and natural gender after deleting the ‘unreliable’ responses. Reanalysing the data revealed statistically significant differences between pre-schoolers and adolescents and between pre-schoolers and adults in the rate of feminine/masculine responses attributed to male-like humans; \((p = 0.034)\) in both cases.
Selecting the no -a words endings

After running the experiment, I found out that the endings of the no –a words were not representative of how those endings are distributed in Arabic according to the four corpora consulted (See Chapter 3)\(^\text{77}\) as shown in Table 5.4.

---

\(^{77}\) This issue was resolved when creating the novel words for experiment III (next chapter).
Table 5.4. The distribution of the no -a endings in the four corpora vs. in the experimental items.

<table>
<thead>
<tr>
<th>Ending</th>
<th>Distribution in the corpora</th>
<th>Distribution in the experimental items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approximants</td>
<td>0.1%</td>
<td>4.2% (N = 1)</td>
</tr>
<tr>
<td>Fricatives</td>
<td>20.0%</td>
<td>29.2% (N = 7)</td>
</tr>
<tr>
<td>Liquids</td>
<td>8.0%</td>
<td>20.8% (N = 5)</td>
</tr>
<tr>
<td>Long vowels</td>
<td>5.0%</td>
<td>8.2% (N = 2)</td>
</tr>
<tr>
<td>Nasals</td>
<td>16.0%</td>
<td>16.7% (N = 4)</td>
</tr>
<tr>
<td>Rhotics</td>
<td>15.0%</td>
<td>16.7% (N = 4)</td>
</tr>
<tr>
<td>Stops</td>
<td>34.0%</td>
<td>4.2% (N = 1)</td>
</tr>
</tbody>
</table>

To explain, some of the endings found in masculine words in Arabic were under-represented in the experimental items, namely stops. As I discussed in chapter 3, stops represent 34.0% of the masculine words I found in the four corpora. Yet, they constituted only 4.2% of the no –a experimental items. Similarly, approximants which I found in only about 0.1% of the Arabic nouns are over-represented in the experimental items, constituting 4.2% of the no –a ending items. Therefore, the types of endings were further investigated to see what potential effect they might have on the overall pattern of the results. As shown in Fig. 5.9, it seems that the endings were treated differently by the participants with approximants assigned the lowest rate of feminine responses (i.e., the highest rate of masculine responses) (mean = 0.18, sd = 0.14) and long vowels (mainly _trap) assigned the highest rate of feminine responses (mean = 0.59, sd = 0.14).
In order to make Fig. 5.10 more comprehensible, I tried to find out how these types of endings interact with age. Fig. 5.11 shows that some but not all endings were treated differently by the different age groups. One way between subjects ANOVA was run to investigate the effect of age on the rate of feminine responses attributed to each type of ending. No significant effect of age was found on the rate of feminine responses attributed to nasal, liquid, rhotic, fricative, approximant, and stop endings. However, a significant effect of age was found on the rate of feminine responses attributed to long vowels $F(4, 98) = 1.60, p = 0.013$. Post hoc comparisons revealed that the difference between the age groups was only significant between school-age children and adults ($p = 0.027$).
In order to check the effect of the different endings in each age group, I ran one-way within-subjects ANOVA tests, with endings as the independent variable for each age group. The prediction is that if the participants are sensitive to the association between the masculine endings and masculine gender in Arabic, they will attribute significantly lower rate of feminine responses to the nouns that have those endings than to the nouns that ended in -a. The results of those tests are summarized in Table 5.5 and explained below. However, a few preliminary caveats must be made before presenting the results. First, any results from the youngest age group must be interpreted with caution. It is very likely that the lack

---

78 Notice that in this table, I am only presenting the results of the comparisons between -a and the other endings which is my interest here but not between the masculine endings themselves. Statistical significance is represented by an asterisk (*).
of statistical power due to the small number of the participants (N = 7) in this age group makes it hard to detect any significant results. Second, the underrepresentation of stops in the experimental items (N = 1) limits the scope of the results coming from the one item that ends in a stop. However, almost all the other endings are very well represented in the experimental items (see table 5.4). Therefore, drawing some conclusion about how L1 learners/users of (Hasawi) Arabic treat those endings should not be an issue here. Third, these tests show the effect of noun endings away from the effect of syntactic agreement and natural gender cues. All the experimental items were presented to the participants with agreeing demonstratives and adjectives in addition to a natural gender cue in some cases. Therefore, determining the subtle effect of those endings in the attested patterns in this experiment does not seem to be a straightforward task.

Table 5.5. Summary of the one-way within-subjects ANOVA for the effect of noun ending in each age group.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>2.6 – 3 years</th>
<th>4 - 5 years</th>
<th>6 - 10 years</th>
<th>11 - 15 years</th>
<th>Adults</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approximants vs. -a</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Fricatives vs. -a</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>Liquids vs. -a</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long vowels vs. -a</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasals vs. -a</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>Rhotics vs. -a</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stops vs. -a</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>
The results of the one-way within-subjects ANOVA revealed no significant effect of noun endings in the youngest age group. As I mentioned above, this could well be due to the lack of statistical power. A significant effect of noun endings was found in the pre-schooler group $F(7, 292) = 9.48, p < 0.0001$. Post hoc comparisons revealed that the difference between the rate of feminine responses attributed to the items with the masculine endings and the rate of the feminine responses attributed to the items that ended in -a was significant except in the case of the long vowel $\acute{i}$ and the rhotic $r$.

Moving to the school age children, the effect of noun ending was also significant in this age group $F(7, 322) = 3.54, p = 0.001$. Post hoc comparisons revealed that the difference between the rate of feminine responses attributed to the items with the masculine endings and the rate of the feminine responses attributed to the items that ended in -a was significant except in the case of the long vowel $\acute{i}$ and the rhotic $r$, and liquid $l$.

For the adolescents, the effect of noun ending was also significant $F(7, 272) = 7.09, p < 0.0001$. Post hoc comparisons revealed that the difference between the rate of feminine responses attributed to the items with the masculine endings and the rate of the feminine responses attributed to the items that ended in -a was only significant in the cases of the approximant $w$ and the stop $b$.

Finally, the effect of noun ending in the adult group was significant as well $F(7, 314) = 10.35, p < 0.0001$. Post hoc comparisons revealed that the difference
between the rate of feminine responses attributed to the items with the *masculine* endings and the rate of the feminine responses attributed to the items that ended in *-a* was significant only in the cases of the approximant *w*, fricatives, and nasals.

To summarise, the four older age groups only agree on treating the approximant *w* significantly different from the feminine ending *-a*. Approximants are not very common endings in Arabic as revealed from the corpora analysis. They are only found in (0.1%) of the nouns in the four corpora. In addition, only one experimental item ended in an approximant. Fricatives and nasals which were very well represented in the experimental items were treated similarly by all the age groups except for the adolescents. Stops, the under-represented endings in the experimental items, were also treated similarly by all the age groups except for the adults. Only pre-schoolers assigned significantly lower feminine responses to the items that ended in the liquid *l* than to the items that ended in *-a*. The rhotic *r* and the long vowel *ī* did not differ significantly from the *-a* ending in all the age groups.

Overall, the lack of a uniform treatment of the *masculine* endings is a reflection of the lack of a uniformity among those endings. It is highly probable that this diversity and lack of uniformity can lead to the masculine endings being a less reliable cue for masculine gender than the uniform feminine ending *-a* for feminine gender.
5.4. Discussion

This experiment was set up to investigate the interaction between gender-correlating information and gender attribution. All the available information for the learner, i.e. noun-internal information (natural gender and morphophonological ending) and noun-external information (syntactic agreement), were considered. The ultimate aim was to find out how these sorts of information interact and how they are weighed by L1 speakers over the course of development, beginning from two-word-utterance stage to adulthood.

The picture that emerged from this investigation was shown in Fig. 5.2 (repeated here as Fig. 5.12 for ease of reference).

Fig. 5.12. Mean rate of feminine responses by age × word ending × agreement × natural gender.
Several patterns emerged from this investigation. Depending on the degree of the reliability of the cues presented in each tested condition and the detectability of those cues, the attested patterns varied among the five age groups. Considering the morphophonological cues, the feminine ending -a is a very reliable and deterministic cue to the noun gender in (Hasawi) Arabic; it is never misleading or ambiguous (MacWhinney et al., 1984). Contrary to this, and despite the high degree of correlation between the absence of -a (or the presence of other endings as will be discussed below) and masculine gender in Arabic, this cue is not as reliable as the feminine gender cue. In some cases, the masculine gender cue/lack of cue is not a predictor of the noun gender. Furthermore, the uniform marking of the feminine gender not only on the nouns themselves, but also on the agreeing adjectives increases the noticeability and the detectability of it as a cue to feminine gender (Frigo and McDonald, 1998). All these characteristics of the feminine gender marker compared to the masculine gender markers/lack of marker appear to contribute to the bias attested among the youngest participants in this experiment to attribute gender in accordance with this marker whenever it is present regardless of the other types of cues (semantic and syntactic). In most cases, they even ignored the deterministic cue provided by syntactic agreement. The presence of -a was weighted heavily by the youngest participants in deciding what gender they assign to the experimental items. The absence of -a however, was not as heavily weighted as its presence in determining gender, i.e. the absence of -a does not override
feminine semantic and syntactic cues as easily as presence of –a overrides masculine semantic and syntactic cues in the youngest group. This pattern is demonstrated in the youngest participants’ responses to conditions 11 and 12. In both of those conditions, the experimental items lacked the -a ending and were presented with feminine agreeing elements. In condition 11, the novel items were also presented as female humans which seems to offer a preliminary explanation for the high rate of the feminine responses attributed by the youngest participants, like all the other participants, in this condition. In condition 12, the novel items were presented as non-humans. The youngest participants were not sure what to do in this condition. They attributed feminine gender half of the times and masculine gender half of the times. The behaviour of the youngest participants in those two conditions show that the lack of -a is not as strong and reliable as its presence. Thus, though young children are biased by morphophonological cues to noun gender class, the overt coding of the feminine gender and the uniformity of the -a ending in the feminine nouns in Arabic makes -a a stronger cue than the heterogeneous endings of the masculine nouns. The overreliance on the morphophonological cues was not attested in the other age groups. Overcoming the morphophonological bias followed a developmental path. It was more obvious in the behaviour of the three older groups than in the pre-schoolers.

79 More about the role of the overt coding and cue uniformity and also about the implications this might have for the representation of gender features in general will be discussed in chapter 7.
Moving to the semantic cues provided by natural gender, this cue is not only reliable but deterministic and never misleading in Arabic. Whenever this cue is present, the grammatical gender of the noun comes for free. In addition, all the younger participants (2-3 and pre-schoolers) do have the conceptual meaning of gender identity and gender constancy which was one of the filtering criteria to take part in this experiment (see chapter 4). This cue, however, is only present in a small subset of nouns in Arabic (10% of the Arabic nouns according to the four corpora consulted). Additionally, in almost half of the human-referring nouns, the female-human-referring nouns are morphophonologically marked by the -a marker in contrast to their male-human-referring counterparts which lacked this marker. This low frequency of human-referring nouns and the correlation between the morphophonological and the semantic cues to the noun gender may affect the detectability of the semantic cue by the young children. The design of this experiment makes it hard to tease the effect of natural gender out because of the cue interaction. It is clear that the youngest participants weigh the morphophonological cue heavier than the other cues because whenever -a was present, they assigned feminine gender disproportionately. The only condition that shows an emerging sensitivity to natural gender by the youngest participants is condition 11 [F agreement, no -a, female human]. Yet, in this condition where the youngest children were not biased by the morphophonological ending, the novel items were also

---

Note that there was a high degree of variation between the four corpora regarding the morphophonological marking of the human-referring nouns.
presented with feminine agreeing elements. Whether the youngest children were sensitive to the natural gender cue in this condition or whether they were assigning the gender suggested by the syntactic agreement is an issue that will be investigated further in the next chapter. For all the other age groups, they assigned the gender suggested by the natural gender cue except in two conditions (5 and 7). In these two conditions all the age groups ignored the natural gender cue because the natural gender cue was competing against the syntactic and the morphophonological cues both of which suggesting the same gender. All the participants were more likely to attribute the gender suggested by the two cues in combination not by the natural gender cue alone. In general, natural gender cue seems to be a very salient cue for Arabic speakers. As I mentioned above, the strength of this cue will be further determined in next chapter.

Regarding the syntactic cue, syntactic agreement is, by definition, the defining element of the gender of a noun. This means that syntactic agreement is a perfectly reliable cue when present. Yet, as I showed in chapter 3, masculine agreement is the default agreement in Arabic. In addition to that, it is almost always the case that masculine agreeing elements are not morphophonologically marked whereas feminine agreeing elements are marked morphophonologically (see chapter 3 for some examples). The effect of the markedness of feminine agreement was clear in the behaviour of all the age groups. When the novel items were presented with feminine agreeing elements all the participants were very likely to assign feminine gender. Notice however that the effect of feminine agreement on
the behaviour of the youngest participants was more noticeable when the novel items ended in -a. On the other hand, when the novel items were presented with masculine agreeing elements, the behaviour of the youngest age group was different from the other age groups. This behaviour of the youngest participants is explicable when knowing that in most cases the competition was between masculine agreement and -a ending. On the other hand, whether feminine or masculine, syntactic agreement was weighed heavier than the morphophonological cues by the other four age groups. This is demonstrated very clearly in their responses to non-human nouns. They assigned the non-human nouns the gender suggested by the syntactic cue not by the morphophonological cue.

To summarize the pattern observed in each age group:

- 2-to-3-year-olds showed a bias for morphophonological cues when the nouns ended in -a. Yet, they were sensitive to natural gender cues and syntactic cues when nouns did not end in -a.

- 4-to-5-year-olds were no longer weighting morphophonological cues heavier than syntactic cues. Syntactic agreement was used as the defining element of the noun gender by participants in this age group. Yet, this age group is more influenced by the interaction of the cues rather than by each type of cue in isolation.

- 6-to-10-year-olds showed to a great extent an adult-like pattern. They were not biased to assign gender in accordance with the morphophonological ending of
the nouns. Rather, natural gender and syntactic agreement were given greater weight than morphophonological ending.

- 11-to-15-year-olds were not different from the previous group. No difference was ever attested between them and the school-age children in all the results.
- Adults’ behaviour was almost the same as the behaviour of 6-to-10-year-olds with syntactic agreement being the most deterministic cue to the gender of the nouns and an interaction between it and natural gender was attested.

Another issue that arises from the current investigation is regarding how L1 Arabic learners approach learning the association between morphophonological endings and grammatical gender. Two strategies can be followed. I will present and discuss each strategy in the following paragraphs.

The first strategy I am suggesting here is that L1 Arabic learners need to learn the association between feminine gender and the -a ending. Masculine gender on the other hand can be assigned to everything else, i.e. everything that lacks the -a ending. Though the greatest majority of the Arabic masculine nouns end in consonants as shown in the corpora analysis in chapter 3, none of those consonants has a suffixal status similar to the status of the feminine marker -a. There is no single case in Arabic where a masculine noun can be derived from its feminine counterpart by the addition of one of those consonants. In contrast, a fair proportion of feminine nouns are derived from their masculine counterparts by the addition of -a (see chapter 3). Thus, there is an asymmetry in the way masculine nouns and feminine nouns are morphophonologically
marked in Arabic. Feminine nouns are marked by a uniform marker which is highly reliable whenever it is present while masculine nouns end in a heterogenous collection of markers which differ in their degree of reliability. Additionally, the feminine marker has an affixal status whereas none of the masculine endings has such a status. It is not unlikely therefore for the L1 Arabic learners/speakers to be sensitive to such an asymmetry. In fact, learning the association between the feminine marker -\textit{a} and feminine gender and the association between the lack of this marker and masculine gender is a very economical strategy. L1 Arabic learners do not need to bother learning the association between masculine gender and a number of heterogenous endings since this can come as a by-product of learning the feminine marker alone. This strategy is very compatible with the results of this experiment. All the age groups attributed significantly more feminine responses to the items that ended in -\textit{a} than to the nouns that lacked this ending. As I discussed in chapter 2, Boloh and Ibernon (2010, 2013) and Boloh et al. (2012) argued that L1 French learners can learn the gender system by learning masculine as default and learning the association between feminine gender and some endings\textsuperscript{81}. Such a strategy would be more feasible in Arabic than in French\textsuperscript{82} where feminine gender is associated with a single, uniform, and perfectly reliable ending whereas masculine gender is associated with a number of heterogenous endings.

\textsuperscript{81} Notice that in French, vocalic endings are associated with masculine gender whereas consonantal endings are associated with feminine gender.

\textsuperscript{82} See chapter 2, section 2.4 for a discussion of boloh and Ibernon and Boloh et al. proposals.
The second strategy that L1 Arabic speakers may follow in learning the morphophonological association with grammatical gender is by learning that -a correlates with feminine gender whereas stops, fricatives, liquids, rhotics, nasals, approximants, long vowels, and affricates correlate with masculine gender. However, not only will the learning load be increased by following this strategy, but also the learning task will be affected by the different frequencies and reliabilities of those endings. As revealed from the corpora analysis in chapter 3, some of the masculine endings are more frequent in Arabic than others. To explain, the greatest majority of the Arabic masculine nouns end in stops. There are also fair proportions of the Arabic masculine nouns ending in fricatives, nasals, and the rhotic r. So, the prediction is that the association between masculine gender and the high frequent endings would be stronger than the association between masculine gender and low frequent endings, namely approximants, long vowels, and the liquid l. In languages like German (Walter et al., 2021) and French (Boloh and Ibernon, 2010), it was found that the more frequent the noun ending is, the stronger the association between this ending and grammatical gender is. The results of this experiment do not provide conclusive evidence for this strategy from the youngest age group. The results suggest that the youngest participants were not following this second strategy. But due to the small number of the participants (N =7) in this age group, this result should not be overstated. Considering the behaviour of the other four age groups, mixed results emerged from this experiment regarding this second
strategy. Those groups, except adults, treated the stop b as a strong cue to masculine gender as they all attributed significantly lower feminine responses to the item that ended in a stop than to the items that ended in -a. Remember however, that only one item in this experiment ends in a stop compared to 12 items that end in -a. This under-representation of stops in the experimental design makes it hard to draw any conclusions. The same applies for the approximant w; only one item has this ending. Yet, the prediction is that L1 Arabic learners/users will not weigh this ending as a strong cue to masculine gender due to its low frequency. The results however, show that all the four age groups do attribute more masculine responses to the items that ended in the approximant w. The nasals and the fricatives which are very common masculine endings in Arabic and which are very well represented in the experimental items, (N = 4) and (N = 7) respectively, show the same pattern in the results. All the four age groups, except adolescents, weigh those endings as strong cues for masculine gender. Surprisingly, the rhotic r, which has the same frequency as the nasals both in the corpora and the experimental items, was not weighed similarly. All the age groups, including the youngest participants, did not treat the rhotic r as a cue to masculine gender. The rate of feminine responses attributed to the items that ended in r did not differ significantly from the rate of feminine responses attributed to the items that ended in -a. The prediction that the long vowels will not be treated as strong cues for masculine
Thus, the results of this experiment strongly suggest that the L1 Arabic learners may follow the first strategy in learning the morphophonological association between grammatical gender and nouns’ endings. Nonetheless, there is no sufficient evidence coming from the results to rule out the possibility that they may follow the second strategy. Deciding on which strategy the Arabic learners follow needs further investigation through an experimental design that considers the distribution and the frequency of the different masculine endings which indeed was not one of the aims of this experiment.

Overall, this data from Arabic demonstrates that, despite early bias for morphophonological cues, sensitivity to semantic and syntactic cues exists as early as 2 to 3 years. The development of grammatical gender from that stage therefore involves adjustments in cue weighting rather than discovery of the cues. In the following chapter, the role of natural gender will be further investigated to have a clearer picture of how it is used by participants of different ages as a cue to the gender of the noun.
Chapter 6: Experiment III: the effect of natural gender

6.1. Introduction

The results of experiment II presented in the previous chapter revealed that the youngest children weigh the feminine marker -a very heavily. That is, whenever the novel items had this marker, the youngest participants attributed feminine gender out of proportion, mostly ignoring all the other cues (syntactic and semantic). This behaviour of the youngest participants confirms the reported bias among young children to rely on the morphophonological cues in learning noun gender classes in many languages (Culbertson et al., 2019; Gagliardi & Lidz, 2014; Karmiloff-Smith, 1979; Levy, 1983; Pérez-Pereira, 1991). However, when the novel items lacked the -a ending, the pattern attested in the behaviour of this group varied according to the other cues. In some conditions, namely when the agreement was masculine, they attributed masculine gender disproportionately suggesting that they may weigh the lack of -a/the presence of the masculine endings as a strong cue for masculine gender. Nonetheless, in other conditions, namely when the agreement was feminine, they showed sensitivity to the syntactic and/or semantic cues and attributed the gender suggested by those cues not the masculine gender suggested by the lack of -a/the presence of the masculine endings. Because of the high degree
of interaction in the previous experiment, it was not easy to decide whether those children were assigning feminine gender because the novel items were presented as female-human-like characters or because the syntactically agreeing forms were feminine.

Therefore, the goal of this experiment is three-fold. Firstly, it tries to further investigate the effect of the natural gender cue by presenting the novel items with no syntactic agreement. Secondly, it provides the opportunity to evaluate the strength of the argument that the lack of \textit{-a}/the presence of the masculine endings is not used by the L1 Arabic learners as a cue to masculine gender in the same way they use the feminine marker \textit{-a} as a cue to feminine gender. This second goal will be achieved by having 80\% of the experimental items with no \textit{-a} endings. If the above argument is in the right direction, the prediction is that the participants, at least the youngest participants, will not attribute masculine gender predominantly to all the items. I predict that their responses will depend to some extent on the natural gender cue. That is, they are predicted to attribute more masculine responses when the novel items are presented as male-human-like characters and more feminine responses when the novel items are presented as female-human-like characters. On the contrary, if the aforementioned argument is not strong enough, the prediction is that those participants will be biased to attribute masculine gender to all the experimental items regardless of the natural gender cue presented. Three items (20\% of the experimental items) that end in \textit{-a} will also be included in this experiment to further support the pattern attested in Experiment II. That is, the
youngest participants are predicted to attribute feminine gender to those items out of proportion, regardless of the natural cue presented, because of the heavy weight of the feminine marker -a. Thirdly, Experiment I (chapter 4) revealed that the two younger groups were sensitive to the semantic association between grammatical gender and natural gender. It was not entirely clear however whether the children’s responses in that experiment reflect an underlying representation of their gender systems or whether their responses were the result of rote learning. After all, all the nouns used in experiment I were real nouns known by the young children. Therefore, half of the no -a items in the experiment in this chapter will be familiar items and half of them will be novel items in order to gain a better understanding of the children’s behaviour in experiment I and the representation of the gender system in its early stages. If the children’s behaviour in Experiment I is a reflection of their underlying abstract representation of the gender system, no significant difference between familiar nouns and novel nouns is predicted. In contrast, if the children’s behaviour in experiment I was a result of rote and item-based learning, different patterns are predicted depending on the item familiarity. That is, an effect of natural gender is expected in the case of familiar nouns but not in the case of novel nouns.

All those predictions will be tested through the experiment presented in this chapter.
6.2. Method

6.2.1. Participants

The same 202 typically-developing Hasawi Arabic speakers who took part in experiment II completed this experiment.

6.2.2. Materials

The verbal stimuli consisted of 15 nouns (Table 6.1). Twelve nouns (6 familiar and 6 novel), do not end in -a while the other three nouns have the -a ending. The no-a novel nouns were created to reflect, to the greatest possible extent, the distribution of the endings found in Arabic masculine words\(^{83}\) to resolve the issue I had with the novel items I used in experiment II (chapter 5). Hence, 33.2% of the novel nouns ended in stops (labbāk, hamrūd), 16.7% ended in fricatives (laftīf), 16.7% ended in nasals (hašbān), 16.7% ended in liquids (daḡfāl), and 16.7% ended in rhotics (daʿbūr). Selecting the familiar items was based on their familiarity to young children and on meeting the ending criterion, i.e. no -a ending.

In addition to the verbal stimuli, 15 pairs of 3D characters of two different colours, which represent male/female humans were created (see Appendix G for

---

\(^{83}\) The distribution of the endings of masculine nouns in Arabic was determined by consulting four corpora as presented in chapter 3.
pictures of the full set of the 3d characters used in this experiment). Each novel noun was represented as a male character to half of the participants and as a female character to the other half of the participants to control for item effect. The 3D characters were assigned to the nonce words randomly taking into consideration the natural gender intended. Also, three 3D familiar animals (a fish, a horse, a duck) were used as practice items at the beginning of the experiment.

Table 6.1. List of the items used in Experiment III.

<table>
<thead>
<tr>
<th>No -a ending nouns</th>
<th>-a ending nouns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Familial nouns</td>
<td>Novel nouns</td>
</tr>
<tr>
<td>raǧǧāl (man)</td>
<td>laftī</td>
</tr>
<tr>
<td>duktūr/ṭabīb (male medical doctor)</td>
<td>daḡfāl</td>
</tr>
<tr>
<td>walad (boy)</td>
<td>labbāk</td>
</tr>
<tr>
<td>ʿarūs (bride)</td>
<td>ḥamrūd</td>
</tr>
<tr>
<td>bint (girl)</td>
<td>ʿabūr</td>
</tr>
<tr>
<td>ʿumm (mother)</td>
<td>hašbān</td>
</tr>
</tbody>
</table>

6.2.3. Procedure

The experiment was carried out in a quiet room in the experimenter’s house or in the participant’s house whichever was more convenient for the participant. Each participant was interviewed individually by the experimenter except for two
children from the youngest age group (2 - 3) and three (4 - 5) who were shy to be left with the experimenter alone and for that reason they were accompanied by their mothers. The mothers were instructed to keep silent throughout the experiment. The answers were recorded using a Zoom recorder.

Each item was presented to the participants in the following way\textsuperscript{84}:

“Here I have X” and then the participant was asked “who do I have?” to make sure that the participant learned the noun. And then the experimenter said “here I have X more” holding the pair of the 3D characters. Then the experimenter repeated while holding the pair “I have X and X.” Then one of the items was hidden and the participant was asked “which X did I hide?” All the questions used in presenting the items had no agreeing elements that could indicate the grammatical gender of the item.

6.3. Results

The dependent variable was grammatical gender (masculine/feminine) attributed to each item, which was determined by the agreeing elements produced by the participants. I will present the results of the no -\textit{a} items first and then the results of the three items that end in -\textit{a} will be presented separately.

\textsuperscript{84} The Arabic script for this experiment and example answers are given in Appendix C.
➢ No -a items

For each participant, the mean of feminine responses in each condition (male human vs. female human) was calculated. Then the mean of feminine responses attributed by each age group was calculated. A by-participant mixed analysis of variance (ANOVA) was run with age (2-3, 4-5, 6-10, 11-15, and adults) as a between-subject variable and natural gender (male human vs. female human) as a within-subject variable.

Fig. 6.1. summarises the resulting pattern. All the age groups attributed more feminine responses to female-like humans and more masculine responses to male-like humans. There was a statistically significant two-way interaction between age and natural gender $F(4, 200) = 9.96, p < 0.0001$. Post hoc tests revealed a significant effect of age when the natural gender was represented by a female-human-like character $F(4, 200) = 6.74, p < 0.0001$, and also when the natural gender was represented by a male-human-like character $F(4, 200) = 5.17, p < 0.001$. 
Pairwise comparisons were run between the age groups for the rate of feminine responses in the case of female referents. A Bonferroni adjustment was applied. The mean of feminine responses was found significantly different between the youngest age group and school-age children ($p = 0.002$), between the youngest age group and adolescents ($p < 0.001$), and between the youngest age group and adults ($p < 0.001$). It was also significant between pre-schoolers and adolescents ($p = 0.045$), and between pre-schoolers and adults ($p = 0.013$).

On the other hand, pairwise comparisons between the age groups for the rate of feminine responses in the case of male referents revealed significant differences between the youngest age group and school-age group ($p = 0.011$), between the
youngest age group and adolescents ($p = 0.001$) and between the youngest age group and adults ($p = 0.002$).

As one of the aims of this experiment is gaining a better insight into the underlying representation of the gender system in its early stages, especially regarding the sensitivity to the semantic gender cues attested in Experiment I (chapter 4), the data from this experiment was re-summarized by adding the familiarity of the items as a variable (Fig. 6.2) It is possible that the observed results in Fig. 6.1 have been caused by the familiarity of the stimuli in this experiment (50% familiar items vs. 50% novel items) especially in the younger age groups. My concern here is the effect of familiarity in using natural gender as a cue for grammatical gender. That is, if the children are sensitive to the correlation between grammatical gender and natural gender cues, their responses are expected to show similar patterns regardless of the familiarity of the items. However, if the young children are not sensitive to this semantic correlation with grammatical gender, their responses are expected to vary depending on the familiarity of the items.
As can be seen in Fig. 6.2, the rate of feminine responses attributed to familiar female humans was far higher than the rate of feminine responses attributed to novel female humans. A by-participant mixed analysis of variance (ANOVA) was run with age (2-3, 4-5, 6-10, 11-15, and adults) as a between-subject variable and natural gender (male human vs. female human) and familiarity (familiar nouns vs. novel nouns) as within-subject variables. The results revealed significant two-way interactions between natural gender and familiarity $F(1, 199) = 139.85, p < 0.0001$ and between age and natural gender $F(4,199) = 10.24, p <0.0001$ (cf. $F(4,$
200) = 9.96, p < 0.0001 above), but no significant interaction between age and familiarity.

When the interaction between natural gender and familiarity was investigated further, no significant effect of familiarity was found in the case of male persons. Yet, a significant effect of familiarity was found in the case of female persons $F(1, 407) = 191.5, p < 0.001$.

The two-way interaction between age and natural gender was further investigated to see if there is any difference between familiar and novel items. A mixed ANOVA was run for each type of words (familiar vs. novel). A significant interaction between age and natural gender was found in the case of familiar words $F(4, 200) = 12.54, p < 0.0001$ as well as in the case of novel words $F(4,199) = 5.23, p = 0.0005$.

Post hoc tests for familiar nouns revealed a significant effect of age in the case of male persons $F(4, 200) = 8.00, p < 0.0001$, with significant differences between the youngest age group and all the other groups ($p < 0.0001$ in all the cases). The effect of age was also significant in the case of female persons $F(4, 200) = 7.72, p < 0.0001$, with significant differences between the youngest age group and pre-schoolers ($p = 0.022$), between the youngest age group and school children ($p = 0.003$), between the youngest age group and adolescents ($p = 0.0001$), between the youngest age group and adults ($p < 0.0001$), between pre-schoolers and adolescents ($p = 0.040$), and between pre-schoolers and adults ($p = 0.011$).
Post hoc tests for novel nouns revealed a significant effect of age in the case of female persons $F(4, 199) = 3.84, p = 0.004$ but not in the case of male persons. Pairwise comparisons between the different groups in the case of female persons showed significant differences between the youngest age group and school children ($p = 0.046$), between the youngest age group and adolescents ($p = 0.023$), and between the youngest age group and adults ($p = 0.017$).

It is clear from this analysis that the overall pattern of responses observed in this experiment (Fig. 6.1), especially in the responses of the two young age groups, is affected by the familiar items. This artifact necessitates caution when interpreting the precise magnitude of the natural gender information in the different age groups.

Furthermore, the responses attributed to novel nouns were also analysed by the nouns’ endings. Fig. 6.3 shows the mean rate of feminine responses attributed by all the participants to each type of ending by the natural gender cue presented. It shows that the lowest rate of feminine responses (i.e. the highest rate of masculine responses) when the natural gender was represented by a female-human-like character was attributed to the noun that ended in the rhotic $r'da'bu'r'$ (Mean = 0.40, sd = 0.29). Surprisingly, $'da'bu'r'$ is also the noun that received the highest rate of feminine responses in the case of male humans (Mean = 0.027, sd = 0.29). A one-way repeated measure ANOVA was run with noun ending as a within-subject variable to find out whether noun ending has a statistically significant effect on the rate of the responses and the results revealed a marginal positive effect of
ending $F(1.9, 379.58) = 4.37, \ p < 0.05$. However, post-hoc analyses with a Bonferroni adjustment revealed that the pairwise differences were not statistically significant between any of the endings. Note however that this result suffers from a lack of statistical power because almost all the endings (except stops) were represented by one single item.

![Graph showing mean rate of feminine responses by natural gender and noun ending. Error bars represent standard errors of the mean (SEM).](image)

Fig. 6.3. Mean rate of feminine responses by natural gender and noun ending. Error bars represent standard errors of the mean (SEM).

The effect of ending was further summarised by age (Fig. 6.4). In order to see whether the endings were treated differently by the different age groups, one way between subjects ANOVA was run to see the effect of age on the rate of
feminine responses attributed to each type of ending. No significant effect of age was found on the rate of feminine responses attributed to any of the endings.

Fig. 6.4. Mean rate of feminine responses attributed to the different endings by the different age groups. Error bars represent standard errors of the mean (SEM).

➢ -a ending items

Now, I will present the results of the three -a ending items. For each participant, the mean of feminine responses in each condition (male human vs. female human) was calculated. Then the mean of feminine responses attributed by each age group was calculated. Fig. 6.5 shows that, as predicted, the presence of the -a ending increased the rate of feminine responses attributed to the novel items compared to the items that did not end in -a (cf. Fig. 6.1). The feminine responses reached a ceiling effect in the school-age children when the natural gender cue was
presented by a female-human-like character. Yet there is still a difference between the rates of feminine responses attributed to those items when they were presented as female vs. male characters by each age group. A by-participant mixed ANOVA revealed a significant interaction between age and natural gender $F(4,158) = 3.78$, $p = 0.006$. Post hoc tests were run within each age group to investigate the effect of natural gender. The differences in the rates of feminine responses attributed to male vs. female characters were found significant in all the age groups ($p < 0.0001$) except for the youngest age group.

Fig. 6.5. Mean rate of feminine responses attributed by the five age groups to the three nouns ending in -a. Error bars represent standard errors of the mean (SEM).
Before moving to the discussion section, I need to consider one potential artefact that could have caused the observed results in the no -a ending words, namely the 3D prompts.

➢ The 3D prompts

It is possible that the observed effect of natural gender is a by-product of the 3D characters used in the experiment. The grammatical gender attributed by each participant could reflect the natural gender they were attributing to each character, which could be different from the natural gender intended when designing and using the prompts. Therefore, I followed the same procedure I used in evaluating the prompts used in Experiment II (Chapter 5). To explain, all the 12 human-like prompts were evaluated by the same Arabic-speaking 20 3-year-old children, 20 (4-5) children and 20 adults. They were simply asked to classify the characters into two groups: boys/males vs. girls/females. They were also told to leave out any character that was hard to classify.

The results showed that 10 out of the 12 human-like prompts were classified correctly (as they were intended to be) by all the 60 participants. The other 2 characters (which were used to present the novel items) were either put in the wrong group or were left out.

As a result, all the test items that were represented using those two prompts were identified. All the responses to those items were marked as unavailable responses (NAs). The data was then reanalysed following the same steps followed
above to analyse the data before deleting those responses in order to detect any
discrepancies that could have been caused by those prompts. Since all the
‘unreliable’ responses were responses to the novel items, only the responses to those
items were reanalysed.

Fig. 6.6 shows that the resulting pattern after deleting the unreliable
responses does not differ from the pattern before deleting those responses (cf. Fig.
6.2). A mixed ANOVA run with the missing responses revealed the same
significant two-way interaction between age and natural gender $F(4, 199) = 5.46, p$
= 0.0003. Post hoc tests also revealed a significant effect of age in the case of female
person ($F(4, 199) = 3.87, p = 0.004$) but not in the case of male person. Pairwise
comparisons between the age groups in the case of female persons showed the same
significant differences between the youngest age group and adolescents ($p = 0.047$)
and between the youngest age group and adults ($p = 0.021$), but not between the
youngest age group and school children ($p = 0.086$). The difference between pre-
schoolers and adults was also found ‘marginally’ significant ($p = 0.049$).
A further post-hoc test was run to compare the rate of feminine responses attributed by the two younger age groups (2-3 & 4-5) to males vs. females. One-way within subjects ANOVA revealed that the effect of natural gender was significant in group 4-5 ($F(1, 49) = 50.27, p < 0.0001$) but not in group 2-3.

6.4. Discussion

This experiment was designed to further investigate the sensitivity of Arabic L1 learners/speakers to the semantic information provided by natural gender as a cue to grammatical gender. A range of interesting results emerged from this investigation. To start with, though the analysis of the no -a items showed that the participants in all age groups attributed more feminine responses to female-like
characters and more masculine responses to male-like characters, adding familiarity to the analysis revealed the reason behind this behaviour. The observed effect of natural gender in the youngest age group (2-3) was caused by the familiar items used as stimuli in this experiment. The effect of the familiar items can be seen in the different rates of feminine responses attributed to familiar female humans vs. novel female humans. It is very likely that such behaviour is due to rote learning. Though youngest participants showed an emerging tendency of attributing feminine gender to novel female humans, the rate of feminine responses attributed to female humans was not significantly different from the rate of feminine responses attributed to male humans.

Although pre-schoolers showed a behaviour similar to the youngest participants (i.e. behaving differently towards familiar vs. novel items), the feminine responses attributed to novel females vs. novel males were significantly different. This shows that it is at this early age that children start to internalise the correlation between natural gender and grammatical gender as part of their gender system.

The behaviour of the two younger age groups was confirmed by the data obtained by using the three –a ending words. To explain, though the presence of –a increased the rate of feminine responses attributed by the children in those two groups (and all the other groups), 2-3 children were found to disproportionately rely on the ending not on the natural gender when attributing gender to those items. Pre-schoolers on the other hand started to overcome the morphophonological bias and
attributed significantly fewer feminine responses to male persons compared to female persons.

Furthermore, the sensitivity to semantic cues followed a developmental path. The rate of feminine responses attributed to female-like humans, whether they were referred to with a novel item that ends in -a or one that lacks this ending, increased as participants grow older. Though adult participants attributed the highest rate of feminine responses to female-human-like characters, the differences between them and between adolescents and school children were not significant. We can conclude from this that it is at school-age that children approach adult-like sensitivity to the correlation between natural gender and grammatical gender. In other words, as early as 6-years of age children overcome the morphophonological bias and behave similar to adults in following the cue provided by natural gender.

With all these results in mind, two main conclusions can be drawn from this investigation. First, the results of this experiment in combination with the results from Experiment II rules out the possibility that the bias among young children to go with the morphophonological cues in deciding the grammatical gender of the noun could be due to the lack of the conceptual meaning necessary to use such a cue. To put it differently, all the youngest participants who took part in this investigation do have the concepts of gender identity and gender constancy as revealed from the results of the MGIT (chapter 4). Nevertheless, they are not sensitive to the correlation between this concept and grammatical gender in their language. It is not part of their abstract representation of grammatical gender system
and they are not able to generalize it to newly encountered nouns. Second, this investigation demonstrates that L1 Arabic learners are using the lack of -\textit{a}l/presence of the masculine endings as a cue to masculine gender. The prediction that the responses of the young participants will vary depending on the natural gender cue presented was not borne out by the results of this experiment. The rate of the masculine responses attributed to the no -\textit{a} ending nouns was very high among the two younger age groups. This suggests that they do use the lack of -\textit{a}/the presence of the masculine endings as a cue to masculine gender. Yet, the results of this experiment confirm the results of Experiment II regarding the asymmetry attested in treating the -\textit{a} versus the lack of -\textit{a}l the presence of the masculine endings. That is the rate of the masculine responses attributed to the novel nouns that lacked the -\textit{a} ending was never as high as the rate of the feminine responses attributed when the novel nouns ended in -\textit{a} (see Fig. 6.1 and Fig. 6.5). This asymmetry in treating the presence of -\textit{a} versus the lack of -\textit{a} also suggests that the learners may be attributing masculine gender by default to everything that lacks the -\textit{a} ending.

Furthermore, in my analysis I tried to investigate the existence of a potential tendency to associate masculine gender with the several phonological endings found in the masculine nouns in Arabic and not simply treating those endings as the lack of -\textit{a}\textsuperscript{85}, i.e. the second strategy suggested in chapter 5. Though the distribution of the endings of the no -\textit{a} novel nouns in this experiment mirrored to the greatest

\textsuperscript{85} Notice that deciding on which strategy L1 learners follow in learning the morphophonological association between nouns’ endings and grammatical gender was not one of the aims of the current study. It is for this reason that the experimental items were not designed in such a way that enables the evaluation of the two strategies.
possible extent the distribution of the endings found in masculine nouns in Arabic, the lack of statistical power is still an issue in this experiment, too. The results show that none of the endings was treated differently by the different age groups. Yet, all the endings, except the stops, are represented by a single item in this experiment. Thus, there is no sufficient evidence coming from this experiment to rule out the possibility that what L1 Arabic learners do is learning the masculine endings as separate phonological cues which vary in their degree of reliability. Still, the results provide additional support to the strength and reliability of the -a ending as a cue to feminine gender in Arabic.

To sum up, the results of this experiment confirm and give support to the results of Experiment II. Sensitivity to natural gender as a cue to grammatical gender seems to follow a developmental trajectory and it exists as early as 2-3 years. Yet, this sensitivity requires time to be internalised as a part of the gender system. As early as 4 years of age, children seem to start internalising the association between natural gender and grammatical gender as part of the gender system. Around 6 years, children approach adult-like weighting of morphophonological and semantic cues for grammatical gender and overcome the morphophonological bias for grammatical gender classes.
Chapter 7: General discussion

7.1. Introduction

The ultimate goal of this thesis was to draw a comprehensive developmental picture of how the grammatical gender system of an L1 speaker starts and the trajectory it follows until it reaches its adult-like state. In the attempt to provide this picture, I investigated the acquisition of the Arabic gender system. I tried to find out when and to explore how children overcome the well-documented morphophonological bias and approximate an adult-like weighting of the cues provided by noun-internal morphophonological and semantic properties and the noun-external syntactic agreement for grammatical gender classes.

Thus, I tried to generate developmental data from language learners/speakers as young as 2.5 years to adulthood using the same experimental tasks. I carried out two experiments using novel nouns in addition to a screening task (Experiment I) and MGIT (chapter 4). Experiment II (chapter 5) investigated the interaction between all the available types of information; morphophonological, semantic, and syntactic. Experiment III (Chapter 6) further investigated the effect of natural gender in gender assignment. Novel nouns were used in both experiments to provide accurate insights into language learners’/speakers’ ability to productively generalize their existing knowledge to newly encountered items.

The developmental trajectory of the Arabic gender system attested in the results will be described in section 7.2. Section 7.3 will address the asymmetry
observed in the weight ascribed to the presence of –a but not to its absence and to feminine agreement compared to masculine agreement.

7.2. The developmental trajectory of grammatical gender

As per the results of the three experiments, I can distinguish three patterns of gender assignment along the developmental trajectory of the gender system of Arabic speakers:

Pattern 1: morphophonological bias

This pattern of gender assignment was manifested in the behaviour of the youngest age group (2;06 – 3;11). The pattern those children followed in attributing gender to the test items confirms previously reported bias for morphophonological cues over semantic or natural gender cues in 2- to-3-year-old children (Karmiloff-Smith, 1979; Levy, 1983; Szagun et al., 2007; Walter et al., 2021). My youngest participants were found to disproportionately attribute gender to nouns according to their morphophonological endings. They even ignored the defining elements provided by the syntactically agreeing demonstratives and adjectives when the nouns ended in -a. A similar behaviour was reported by Walter et al. (2021) for German-learning three-year-old children. The children in their experiment attributed gender according to the two highly reliable phonological cues they provided. They relied on those cues in determining the gender of the novel nouns and they produced the definite articles according to them, not according to the gender-marked indefinite articles, adjectives, and demonstrative pronouns that were
used to introduce the items. This finding, i.e. that morphophonological cues were weighted even heavier than syntactic agreement is in contrast with what Correa and Name (2003) and Correa et al. (2011) found with Portuguese-learning children. They studied the behaviour of children younger than three (2-year olds) and older than three (4-year olds) making the young participants in their experiments comparable to the young participants in my experiments. They found that the young participants attributed gender to novel nouns (which was determined by the agreeing forms produced by the children in an elicited production task) according to the gender of the determiners not according to the nouns’ endings and that sensitivity to noun endings increased as children grew older. In my experiments, however, the opposite direction was observed. The bias for morphophonological endings decreased in effect and the effect of syntactic agreement increased as my participants grew older. A similar direction to mine was observed by Karmiloff-Smith (1979) and Pérez-Pereira (1991). The youngest participants in their experiments ignored the cues provided by the determiner or the sex indications and assigned gender according to the nouns’ endings. Sensitivity to the nouns’ endings decreased and following the gender suggested by the determiner increased as the participants grew older. Karmiloff-Smith, Pérez-Pereira, Correa and Name, and Correa et al. all studied Romance languages (French, Spanish, and Portuguese). The youngest participants in Karmiloff-Smith’s and Pérez-Pereira’s studies (3-year-old children and 4-year-old children respectively) were older than the youngest participants in Correa and Name’s and Correa et al.’s studies. Collectively, their
results may suggest that children learning Romance languages start learning gender through the closed class of determiners as suggested by Correa et al. According to them, children learn the gender of a new noun through the syntactic information provided by the accompanying determiner. As children grow older, they start to incorporate the correlation between gender and noun-internal information and give it prominence, as the pattern of responses of the older children in Correa and Name (2003) and Correa et al. (2011) and the pattern of responses of the young participants in Karmiloff-Smith (1979) and Pérez-Pereira (1991) suggest. After that, agreeing determiners gain importance over formal properties of the noun again as suggested by the behaviour of the older participants in Karmiloff-Smith’s and Pérez-Pereira’s studies.

Determiners in Romance languages have an obligatory status, and it is highly possible that learners of those languages acquire determiner-noun pairs as unanalysed chunks. There is a growing body of literature supporting the idea that children follow a start-big approach in language acquisition (Arnon & Ramscar, 2012; Peters, 1977; Skarabela et al., 2021; inter alia). Children acquiring languages with obligatory determiners may differ in how they approach the task of determiner-noun learning and how they learn determiner-noun gender agreement. They may start by learning determiner + noun combinations as unanalysed units. This may facilitate the discovery of the dependency between the determiners and the nouns. It may also suggest that learning those dependencies involves a high degree of memorization and rote learning. As their lexical knowledge expands, they may,
notice that the nouns occurring with different determiners have things in common; they end in particular sounds or they share particular meanings. Once the intrinsic similarities within each noun class have been discovered, some generalizations may start to be executed by the learners, especially in the case of new and low frequency items (Monaghan et al., 2005). Monaghan et al. (2005: p. 177) argued that distributional cues are used to categorize high-frequency words whereas phonological cues are useful in categorizing low-frequency words as “their use compensate for the absence of distributional information for words that occur rarely”. That is, the co-occurrence of the determiner + noun in the case of highly frequent nouns may get entrenched in the lexicon as a chunk while, in the case of low-frequency nouns, the association between the noun and the determiner is not expected to be entrenched in the lexicon as a chunk because there are not enough co-occurrences for the entrenchment to take place and in this case the role of the morphophonological cues comes into play. The same applies for new nouns that enter the language with no gender-defining determiners and for novel items in experimental situations. If determiner + noun combinations in these languages are indeed initially learned like whole chunks, using determiner-noun agreement as a test for the acquisition of grammatical gender may not be the best method in those languages, especially with highly frequent familiar nouns. Though, the aforementioned studies used novel nouns, it is possible that learners make a connection between definite and indefinite articles. Thus, introducing the noun with one determiner may prompt the other. Arabic on the other hand, does not have any
obligatory definite/indefinite determiners and nouns are mostly used with no accompanying determiners (see Chapter 3). Though Arabic has a closed class of gender-marked demonstratives, those demonstratives are not by any means obligatory. Also, there is a tendency among some Arabic speakers to overgeneralize the masculine demonstrative in child-directed speech [CDS] as in Alqattan’s (2015) Kuwaiti corpus, especially when asking “what is this?” because ‘thing’ in Arabic is masculine. Omar’s (2;06) mother for example was always using the masculine form of the demonstrative even when the noun she was referring to was feminine or ended in -a. Yet, in the same corpus, some parents use both masculine and feminine forms of the demonstratives in accordance with the gender of the thing they are asking about. Mariam’s (2;04) mother was using both forms of the demonstratives when talking to her daughter and the effect of this is clear in the different forms of demonstratives Mariam was producing.\textsuperscript{86} The most reliable cue within the noun phrase domain in Arabic comes from adjectives. As I mentioned in chapter 3, noun-adjective agreement in Arabic exhibits, to a great extent, alliterative concord. That is, “the noun itself includes a form which is identical to the gender agreement marker and which clearly indicates the gender of the noun” (Corbett, 1991, p. 117). This type of concord may gravitate children’s attention to noun endings as the most reliable cue for their gender class, which could give an explanation for the early morphophonological bias attested in the youngest age group. This alliterative

\textsuperscript{86} Though Mariam was using the English names of some animals and things like ‘fish’ and ‘car’, she was using the Arabic demonstratives in accordance to the gender of the Arabic translation equivalents when describing them.
concord is also attested in some Romance languages such as Spanish and Italian as the examples in (58) and (59) from Spanish and Italian respectively illustrate. Those two sentences show alliterative concord not only between the nouns and the adjectives but also between the nouns and the articles. It is very likely that this type of concord has an effect in gravitating the attention of the children learning those languages to word endings as a reliable cue for gender class membership. Notice however, that articles are more frequent than adjectives in those languages (Mariscal, 2009). Thus, though Arabic, Spanish, and Italian all exhibit alliterative concord between nouns and adjective, adjectives are expected to be a stronger cue for gender class in Arabic than in Spanish and Italian. In Spanish and Italian, nouns are almost always accompanied by gender defining determiners which might attract the learners’ attention away from adjectives. In addition, some highly frequent adjectives are gender-invariant in those languages such as grande (big).

58) Un-a muchaca alt-a

a-F Girl(F) tall

‘A tall girl’.

59) Un-a sedia rott-a

a-F chair(F) broken-F

‘A broken chair’.

Moreover, Experiment II revealed an asymmetry in the way young children treated the presence of the morphophonological ending –a and the way they treated
its absence. The -a ending was heavily weighed by young children as a cue to feminine gender whereas its absence was not similarly weighed as a cue to the opposite gender, i.e. masculine. When –a was present, they even ignored the gender suggested by the syntactic agreement as I mentioned above. This latter behaviour lends support to the accounts that argue for the importance of the shared similarities within each class in order for noun-class learning to take place successfully (Braine, 1987; Braine et al., 1990; Brooks et al., 1993; Brown et al., 2018; Gerken et al., 2005). They all found that learning arbitrary classes that have long-distance syntactic dependencies is very difficult, if not impossible, unless there is something in common between the members in each class. The commonality within noun class systems that children were repeatedly found biased to follow is morphophonological endings. When a reliable morphophonological ending (such as –a here) is present, young children seem to ignore all other sorts of information (semantic or distributional) no matter what statistical reliability those sorts of information have (Gagliardi & Lidz, 2014). However, when such morphophonological information is absent, young children seem to consult the other types of available information (semantic and distributional). My data showed that when the words did not end in –a in Experiment II (chapter 5), young children attributed more feminine responses to female persons than to male persons and also more feminine responses when the syntactic agreement was feminine than when it
was masculine. When the effect of natural gender was further investigated in Experiment III (chapter 6), children manifested an emerging sensitivity to natural gender. Yet, this sensitivity was, to a large extent, overridden by the morphophonological ending of the novel words.

It goes without saying that it is through co-occurrences that L1 learners discover the existence of noun classes such as grammatical gender classes in their language, because that is how gender manifests itself in gendered languages. Yet, the learners’ task is made achievable by the existence of shared similarities between nouns in each sub-class. The accounts that argue for the learnability of category and sub-category membership in artificial languages by relying solely on the distributional information such as Reeder et al. (2017) do not receive support from the acquisition literature. As also pointed out by Gagliardi and Lidz (2014), such accounts fail to capture the discrepancies between children and adults in their sensitivity to the different types of cues. That is to say, if noun classes are acquired by relying on the external distributional information provided by syntactic agreement and noun-internal (semantic and morphophonological) cues are integrated into the system later on after the lexicon has been built, children and adults are not expected to vary in their sensitivity to the different cues because their

---

87 A similar behaviour was also reported by Karmiloff-Smith (1979), 3-year-olds used the information provided by the determiner when the noun provided no phonological information.

88 I cannot say that the youngest participants were generally insensitive to the semantic cues because the results of Experiment II (chapter 5) showed clearly how the pattern of responses changed in accordance with the information provided by the natural gender. Even when no syntactic information was provided in Experiment III (chapter 6), neither masculine responses when the words did not end in -a nor feminine responses when the words ended in –a reached ceiling level in this age group.
correlation with noun classes “will be drawn from a mature lexicon” Gagliardi and Lidz (2014, p. 66). Nevertheless, if children acquire noun gender by consulting both agreement dependencies as well as the information provided by the nouns themselves, the discrepancies between children and adults is explained by the fact that children have access first to the phonological form of the words and later to their meaning (Culbertson et al., 2019).

The behaviour of the young participants in the current study suggests that they weigh the available information in accordance with a scale similar to the one given in (60) to decide on the gender of a newly encountered noun. To explain, young children are biased to follow morphophonological information when it is present. When no morphophonological information is there, they give prominence to syntactic agreement over natural gender which is the least salient cue for them.

60) Morphophonological ending >> syntactic agreement >> natural gender

The scale in (60) is supported by the results of both Experiment II and Experiment III. When the items ended in \( -a \) in Experiment II and in the three items used in Experiment III, the youngest children were found to attribute feminine gender out of proportion. When the nouns did not end in \( -a \) in Experiment II, they were found to assign more masculine responses when the agreeing elements were masculine and more feminine responses when the agreeing elements were feminine. They were also found to assign more feminine responses to female-like characters and
more masculine responses to male-like characters. However, when the effect of natural gender was further investigated in Experiment III, the youngest participants behaved differently depending on the familiarity of the items. They assigned the gender predicted by the natural gender of the characters only in the case of familiar items. In the case of the novel items, however, they assigned masculine gender to the majority of the items regardless of the natural gender information. The scale in (60) is also supported by the effect size of each type of information in this age group compared to the other age groups (Chapter 5). Morphophonological ending explains 45%, syntactic agreement explains 23%, and semantic information explains only 11% of the variation in the responses of this age group.

No matter how the young children learn the morphophonological association between grammatical gender and noun endings, the feminine marker -a is a more reliable cue than the absence of -a/the presence of the masculine endings. The feminine marker -a is never misleading. The absence of -a/the presence of the masculine endings however is not always a predictor of the grammatical gender of the noun, making it a less reliable cue. The results of experiment II and III demonstrates clearly how the high reliability of the feminine marker -a makes it a deterministic cue for the youngest children. Whether those children are learning the morphophonological association between masculine gender and masculine endings or whether they are assigning masculine gender by default to everything that lacks the -a ending, it is clear from their responses especially in experiment III that they were not weighting the cue(s) for masculine gender as heavy as the feminine gender
cue. The masculine gender morphophonological cues are not as reliable as the feminine gender marker -a.

Different explanations have been offered for the morphophonological bias among young language learners. Gagliardi (2012), Culbertson and Wilson (2013), Gagliardi and Lidz (2014), and Culbertson et al. (2019) suggest that there is a general active bias among children to rely on intra-linguistic morphophonological cues rather than extra-linguistic semantic cues when learning noun class. Culbertson et al. (2019) also suggest that the early availability of phonological cues compared to semantic cues or meaning might also provide an explanation for the over-reliance on morphophonological cues among children. To explain, children have access to the phonological form of the nouns before they learn their meanings; this access could be as early as when they are new-born infants (Ramus, 2002; Shi et al., 1999). Yet, access to words’ meanings may take up to 6-months of age (Bergelson & Swingley, 2012; Tincoff & Jusczyk, 2012). The child participants in my experiments are much older than 6 months as the youngest participants are two-and-a-half-year old (the child participants in Culbertson et al. are even older, i.e. six-to-seven-year-olds). It is unlikely that accessing semantic information has such a prolonged effect and that it persists up until seven years of age. What is likely however is that phonological information is more salient than semantic information for children when learning language as a problem per se. That is, when children encounter a language-learning problem as in artificial language settings, they attend to the intra-linguistic phonological information to solve this problem but not to the
extra-linguistic semantic information (Gagliardi, 2012; Culbertson and Wilson, 2013; Gagliardi and Lidz, 2014; Culbertson et al., 2019). Another explanation was suggested by Pertsova and Becker (2021). They suggested that one explanation for the phonological bias attested in children, particularly in artificial language learning, is that learning phonological information happens implicitly. Learning semantic information on the other hand involves explicit learning. “The explicit system is hypothesized to be effortful, and involve conscious deliberation and hypothesis testing which can lead to abrupt learning, while the implicit system is hypothesized to be effortless, automatic, and involve unconscious statistical learning leading to gradual improvement over time” Pertsova and Becker (2021: p. 131). They compared the performance of 5-to-11-year-old children to the performance of adult participants in an artificial language learning experiment. Simply put, they reported a kind of phonological bias in both groups of participants, children and adults, though it was more common among children than among adults. What those participants had in common is that they all failed in stating rules that characterize the phonological (number of syllables) and semantic (animacy or shape) patterns according to which the experimental items were classified. On the other hand, no phonological bias was attested among participants of both age groups who succeeded in stating rules that describe the patterns for classifying the experimental items. They took those verbal reports as evidence of the learning

---

89 Note that subjective verbal reports are not always reliable measures of how participants approach experimental tasks.
mechanism with which different learners, no matter what age they are, approach the problem of experimental language learning as well as natural language acquisition. According to them, their findings support a phonological bias in implicit learning only. However, when the task of language learning is approached explicitly, no phonological bias is attested. Pertsova and Becker’s findings can be extended to first language acquisition which happens naturally, effortlessly, and unconsciously. In other words, language acquisition happens implicitly and with no formal instruction that might direct the child’s attention to language learning as a problem that they have to try different hypotheses to solve. Thus, if there exists a phonological bias in implicit learning, children over-reliance on morphophonological cues in gender assignment is not surprising because that is how they approach language learning, i.e. implicitly.

To sum up, this pattern shows that the working system of Arabic grammatical gender in the early stages of language development is driven by the morphophonological similarities among nouns in each class more than by their semantic similarities or by the gender-defining syntactic agreement.

Pattern 2: syntactic predominance or joint effect

This pattern is demonstrated in the pre-schoolers’ behaviour. At this age, syntactic cues were no longer overridden by morphophonological cues. Children were found to attribute more feminine responses when the syntactic agreement was feminine and more masculine responses when the syntactic agreement was
masculine, regardless of the word ending. They also showed slightly greater sensitivity to the semantic cues provided by natural gender than the youngest age group, which was apparent in their responses to the novel items used in Experiment III (chapter 6). They attributed significantly more feminine responses to female humans than to male humans. They also seemed to prefer feminine nouns to be marked morphophonologically. They were adding –a to the novel nouns when they were presented as female humans\textsuperscript{90} or with feminine syntactic agreement.

Pre-schoolers start to internalise all the available types of information as part of their gender system. It has been reported that, cross-linguistically, 4 year olds are biased to rely on morphophonological cues, not on semantic cues, to determine noun gender class (Gagliardi & Lidz, 2014; Correa & Name, 2003; Karmiloff-Smith, 1979; Levy, 1983; Pérez-Pereira, 1991; Seigneuric et al., 2007; Szagun et al., 2007, inter alia).\textsuperscript{91} The data I obtained from Arabic children in this age group is in line with those findings in this regard, i.e. morphophonological cues were weighted heavier than semantic cues by this age group. Yet, the pattern attested in the responses of this age group followed, to a certain extent, a different pattern. The weight those children gave to morphophonological ending was lower than the weight it was given by the youngest age group.\textsuperscript{92} More importantly, this

\begin{itemize}
\item \textsuperscript{90} A similar tendency was reported for Portuguese children by Correa et al. (2011).
\item \textsuperscript{91} Notice that Karmiloff-Smith and Pérez-Pereira also found that French and Spanish children gave more importance to syntactic agreement as they grew older. Yet, that was at an older age than found in the present study.
\item \textsuperscript{92} In the studies I cited above, the bias for morphophonology was compared to the insensitivity to semantics.
\end{itemize}
data shows that children at this age gave more importance to the cues in combination than to the cues in isolation. Only 13% of the variations in their responses was explained by the morphophonological ending (recall that 45% of the variation in the responses of the youngest age group was explained by morphophonological ending), 26% of the variations was explained by the syntactic agreement, and only 7% of the variations was explained by the natural gender information. This effect size suggests that the different types of cues were weighted differently by this age group, with syntactic agreement given the heaviest weight.

A closer look at the pattern of responses of this age group shows that they were assigning the gender suggested by the majority of the cues (Fig. 7.1 below). To explain, when two cues suggest one gender while one cue suggests the other gender, they assigned the gender suggested by the two cues in combination. This suggests the scale in (61) for the weight given to each type of information by this age group, which is supported by the pattern of responses found in Experiment II. Fig. 7.1 (the circled responses) illustrates how important the syntactic agreement is for preschoolers especially when the nouns ended in -a.\footnote{Notice that the results of Experiment II revealed no significant differences in the rate of feminine responses attributed by all the age groups when the nouns did not end in -a.}

\[ (61) \{\text{Syntactic agreement, Morphophonological ending} \} >> \text{Natural gender} \]
They attributed more feminine responses when the agreement in the prompt was feminine and more masculine responses when the agreement was masculine. Their pattern of responses shows clearly how they start to overcome the morphophonological bias and approach an adult-like pattern (pattern III below). Considering natural gender, the pre-schoolers attributed more feminine responses to female-like characters and more masculine responses to male-like characters. That behaviour was attested in Experiment II and confirmed by Experiment III.94 Though the differences between this age group and the youngest age group in using natural gender information did not reach statistical significance in either of the experiments, the difference between this group and the school-age group in using natural gender information was not significant either. This shows how this group is in an intermediate stage; a stage that is different from the previous stage but at the same time a stage that has not yet reached the final stage but approached it very closely.

---

94 This pattern was not observed in the youngest age group as the analysis of the novel items revealed.
In short, all the available information seems to be integrated in the gender system at this age. It is not the morphophonological properties of the nouns that are the most important driver of the system. Rather, the additive effect of the different types of information is taken into account when assigning gender. Pre-schoolers assigned the gender suggested by the majority of the cues. In addition, there is a preference among pre-schoolers for natural gender aspects of the nouns to be overtly-coded. To put it differently, pre-schoolers prefer feminine gender to be overtly-coded in nouns that refer to female humans.

Fig. 7.1. The mean rate of feminine responses by age × noun ending × syntactic agreement × natural gender (circles highlight the pre-schoolers group responses).
Pattern 3: End state

This pattern is manifested in the behaviour of the older three age groups: school children, adolescents, and adults. This behaviour marks the end state or the adult-like state of the gender system. No significant differences were found in the way those three groups responded in all the experiments. The results suggest that it is at school age that children overcome the morphophonological bias and approximate an adult-like weighting of morphophonological, semantic and syntactic cues for grammatical gender classes. The adult-like weighting is suggested in (62).

(62) \{Syntactic agreement, natural gender\} >> morphophonological ending

The scale in (62) is confirmed by the effect size of the different types of information in school-age children, adolescents, and adults. To explain, it is at school-age that children seem to overcome the morphophonological bias because the effect size of morphophonological cues was small in this age group similar to the two older groups (adolescents and adults). Nonetheless, the effect size of semantic cues or natural gender in this age group is still not as big as it is in the two older age groups.

The heavy weight given to syntactic agreement by the three age groups was obvious in the way they responded to non-human stimuli: they attributed more feminine responses when the agreement was feminine and more masculine
responses when the agreement was masculine no matter what ending the nouns had. Overcoming the morphophonological bias in these age groups was also confirmed by the results of Experiment III. The three groups were attributing gender according to the natural gender of the 3D characters even in the three items that ended in -a. Sensitivity to semantic cues and using them in attributing gender to nouns, when both semantic and phonological cues are equally reliable or when they are in conflict, was repeatedly reported for adults but not for school-age children (Brown et al., 2018; Carroll, 1999; Gagliardi & Lidz, 2014; Karmiloff-Smith, 1979; Pérez-Pereira, 1991). The data I obtained from Arabic-speaking school-age children however shows that semantic cues are very salient cues for them, albeit not as salient as they are for the two older age groups. Two potential explanations for the fact that semantic cues are very salient for school-age children (and for the older groups) in Arabic can be offered here. First, it was found that animacy as a semantic cue to noun class is easier to learn than other semantic cues like shape for instance (Culbertson et al., 2017; Pertsova & Becker, 2021). In Arabic, animacy and more precisely humanness is a deterministic cue to the gender of the noun. Yet, this correlation between natural gender and grammatical gender is not something special about Arabic, as it is found, according to Corbett (2014), in three quarters of the languages which have semantic assignment systems. Among those languages, we find French, and Spanish where it has been reported that school-age children are biased by the morphophonological cues to noun gender classes (see the references mentioned above). Second, though grammatical gender tends to
correlate with natural gender in many of the world’s languages, in Arabic, this
correlation is perfect. Natural gender always predicts grammatical gender in Arabic.
This is not the case in all languages that have a gender system in which natural
gender is one of the determining factors. To mention one of the most cited
examples, the words Weib ‘wife/woman’ (derogatory) and Mädchen ‘girl’ in
German are not feminine but neuter (McConnell-Ginet, 2013). Considering
pronominal agreement with such nouns, feminine pronouns are sometimes used95.
Nevertheless, Arabic does not have a single instance where natural gender fails to
predict grammatical gender. Arabic epicene nouns such as fard (individual) and
šaḥṣ (person)96 which could denote males as well as females are not counter
examples for the perfect correlation between natural gender and grammatical
gender in Arabic because almost always those nouns are used for generic reference
where the sex of the referent is not important. Therefore, gender agreement with
such nouns is masculine by default as sentence (63) illustrates.

95 A distinction should be made here between lexical gender which is a property of the noun itself
and referential gender which is determined by the referent of the noun phrase (Dahl, 2000).
96 ʼinsān ‘human’ also has a generic meaning when used in its masculine form but when used in
reference to a female human, -a is suffixed to it. It is worth mentioning that nās ‘people, the plural
of ʼinsān, exhibits different agreement patterns depending on the intended reading. Feminine
singular agreement is used when the collective reading is intended while masculine plural is used
when emphasizing the individuality of the different people (see Benduhais, 2018 for examples and
discussion).
It could be tempting to connect the saliency of natural gender in Arabic to cultural development and the school system. In the Arab world and in Saudi Arabia in particular, pupils are educated in gender-segregated schools. Starting from primary school\(^97\) and up to the university, males and females are segregated from each other (Alhazmi, 2010; Baki, 2004). Yet, Brutsaert (1999) found that girls in co-educational schools have a more enhanced concept of gender identity compared to girls in gender-segregated schools. Therefore, it does not seem that cultural development is the reason behind the saliency of natural gender as a cue to grammatical gender in Arabic.\(^98\)

What is surprising is the saliency of semantic cues/natural gender in the end state of the system. I was expecting to find an overall reliance on syntactic agreement in the end state of the system because it is through agreement one would learn the grammatical gender of newly encountered nouns. Yet, syntactic agreement and semantic properties are of similar importance to Arabic speakers. As illustrated

\(^{97}\) Nowadays however, mixed-gender education has been introduced in the early years (P1 - P3), i.e. 6-to-9-year olds.

\(^{98}\) In principle, natural gender is a very salient concept for school children not only in the Arab world. Being in a multi-cultural society for four years, I can tell how important the differences between boys/males and girls/females are even to non-Arab kids. Their questions, their preferences, and their behaviour all show the importance of such differences.
in Fig. 7.2, it was not always the case that when the noun was presented with feminine agreeing element, feminine gender was assigned to it and when the noun was presented with masculine agreeing elements, masculine gender was assigned to it. Natural gender had a very important effect in shifting the responses even when the syntactic agreement suggested the opposite gender. This could be due to the fact that semantic information is noun-internal and an inherent property of the noun itself whereas syntactic information is noun-external and comes from an element other than the noun itself.

Fig. 7.2. The mean rate of feminine responses by age × noun ending × syntactic agreement × natural gender (circles highlight the three older groups responses).
I am not aware of a study that investigated the competition between natural
gender and syntactic agreement in gender class membership in other languages up
to adulthood. It would be interesting to find out how adolescent and adult speakers
of languages with highly reliable morphophonological cues like Spanish would
behave in a task where semantic and syntactic information compete while
controlling for morphophonological endings. Afonso et al. (2014), for instance,
found that adult native speakers of Spanish only consulted the syntactic information
provided by the definite article when the nouns did not end in -a / -o, a pattern
similar to the one I observed with the youngest subjects. Gollan and Frost (2002)
argued that grammatical gender in Hebrew can be accessed without the need for
consulting gender-form correlating information. Holmes and Segui (2004) found
that native French speakers use the information provided by both the noun ending
and the article when deciding on the gender of a given noun. Nonetheless, the
semantic cues were not among the factors Afonso et al., Gollan and Frost, or
Holmes and Segui tested in their experiments as they all focused on inanimate
nouns. Vigliocco and Franck (1999) found that Italian and French adult speakers
produced more gender agreement errors when the head noun had no gender-
meaning correlation than when the head noun had conceptual gender, i.e. correlates
with the natural gender of the referent. Inconsistent conclusions about the effect of

---

99 Pérez-Pereira (1991) actually investigated this cue combination in items 13 and 14 in his study
where he provided no morphophonological cue to the gender of the novel items. So, the competing
cues were the determiner and the biological sex indications in the Martian-like drawings. He found
that in both items, children between 4 and 11 years predominantly attributed gender according to
the gender of the determiner rather than the natural gender indicated in the drawings.
Animacy and natural gender can also be drawn from studies that investigated the processing of grammatical gender by adult L1 speakers of different languages. Alarcón (2009), for instance, found that adult native Spanish speakers, as well as adult L2 learners of Spanish, processed gender agreement faster when the gender of the nouns was semantically congruent (i.e., indicated by biological sex) than when it was arbitrary. This finding shows the importance of animacy/biological sex as a cue to grammatical gender for adult speakers of Spanish. In contrast, Sagarra and Herschensohn (2011) found that animate nouns were more difficult to process than inanimate nouns by Spanish speakers. Alamry (2019) reported no effect of animacy in the processing of gender agreement by native adult speakers of Arabic. However, syntactic agreement was not among the variables the aforementioned studies investigated. Schiller et al. (2006) found that native German speakers made faster decisions about the gender of a given noun by selecting the correct determiner when the noun was semantically marked in that it belongs to a group of nouns that share a particular meaning (e.g., fruits) which is associated with a particular gender. Furthermore, there is ample evidence that adult speakers make use of the syntactic information provided by the gender of the article in making predictions about the following noun (see Lew-Williams & Fernald, 2007, 2010 for Spanish, Guillelmon & Grosjean, 2001 for French, inter alia). Yet, these studies have not considered a case in which syntactic agreement competes with natural gender information. Two studies that could be related to the different weights given by adult speakers to syntactic and semantic information are Grüter et al. (2020) as well as Tsang and
Chambers (2011). They found that L1 adult native speakers of Mandarin (Grüter et al., 2020) and Cantonese (Tsang & Chambers, 2011) rely on the syntactic information provided by the noun classifier not on the semantic information of the noun itself when selecting the target noun in their experiments. Therefore, there is no sufficient evidence to say that the pattern I found is special for Arabic speakers or if we could observe similar behaviour in other languages.

In a nutshell, Arabic-speaking school-age children overcome the morphophonological bias and weigh the different types of cues for grammatical gender class following a scale similar to adult speakers’ scale. In the end state of the grammatical gender acquisition, syntactic agreement and natural gender information have an additive effect in determining the noun gender. The effect of

---

100 These two studies looked at numeral classifier systems which are different from gender systems. To explain, different nouns require different classifiers which are mainly conditioned by the semantics of the noun. Classifiers are independent items that accompany the noun and they do not trigger agreement. In Japanese for example, *bon* is the numeral classifier used for long objects as in the example below:

- *enpitsu san-bon*
  - pencil three-CL
  - ‘Three pencils’.

In Thai, *tua* is the numeral classifier for animals as in the example below (both examples are cited from Yamamoto, 2011):

- *ma si tua*
  - dog four CL
  - ‘four dogs’

So, simply put, the gender of the noun requires the agreeing elements to take different forms but the numeral classifiers are themselves self-standing items conditioned by the meaning of the noun. Such differences between nominal classifiers and gender could potentially affect the way learners acquire those systems.
syntactic information and natural gender information is never overcome by morphophonological ending.

7.3. Morphophonological and syntactic asymmetry

As stated in the previous section, the results of this investigation revealed an asymmetry in the way Arabic speakers treated the presence of –a and the way they treated its absence. To explain, -a was heavily weighted as a deterministic cue to feminine gender especially by the young participants while the absence of -a/the presence of masculine endings was not used to the same degree as a cue to masculine gender. When the novel nouns ended in –a, the youngest participants ignored all the other cues, whether semantic and/or syntactic. However, when the novel nouns did not end in –a, those same participants consulted the other types of cues and seemed to attribute gender to novel nouns accordingly. This is clearly illustrated in Fig. 7.3 (the circled pattern). The top part of the figure shows the quite high rate of feminine responses attributed by the youngest age group when the nouns ended in –a no matter what other cues were provided. In contrast, the bottom part of Fig. 7.3 shows how the pattern of responses shifted according to the other cues (semantic and syntactic) when the nouns did not end in -a.
I will try to explain this asymmetry in this section by discussing the role of overt-coding, cue uniformity and reliability in section 7.3.1 Section 7.3.2 will discuss the implications that this asymmetry might have for the representation of the morphosyntactic feature [gender] in Arabic. Section 7.3.3 will consult the Tolerance Principle to check the productivity of the assignment patterns participants were found to follow. Finally, section 7.3.4 will place the results within the debate of rule hierarchy/ordering.

Fig. 7.3. The mean rate of feminine responses by age x noun ending x syntactic agreement x natural gender (circles highlight the youngest age group responses).
7.3.1. Cue uniformity and overt coding

Young language learners seem to search for overt and consistent cues (Finley & Wiemers, 2015) to help them in attributing gender to new nouns. The uniformity of -a makes it an appropriate candidate as a cue to the feminine gender class membership. It is never misleading when it is present, which contributes to its validity as a cue (MacWhinney et al., 1984). Additionally, the feminine marker -a is overtly coded, which contributes to its detectability as a cue. This ending uniformly marks more than 90% of the feminine nouns in Arabic. The overt-coding of feminine gender in Arabic is not special for human-referring nouns. It is also manifested in the homogeneity of the morphophonological ending attested in more than 99% of the non-human feminine nouns in Arabic. The case of non-human-referring nouns is different from human-referring nouns in that feminine non-humans are not derived from their masculine counterparts by the affixation of –a in the vast majority of the nouns, such morphophonological homogeneity among the two subclasses seems to enhance the strength of –a as a cue to feminine gender for Arabic learners/speakers. All these characteristics of the feminine ending -a makes learning the morphophonological association between words’ endings and feminine gender a straightforward task. The strength of -a as a cue was clear in the heavy weight attributed to it by the young participants, and indeed all the other participants, throughout the set of experiments reported in this project.

Regarding the morphophonological cues for masculine gender in Arabic, the task of learning the association between these cues and masculine gender is less
straightforward. Two plausible scenarios were considered in this project. The first one, which I will call the default masculine strategy, represents an imbalanced situation for the learner. In Arabic, nouns either end in –a or not; there is no particular morphophonological ending with which masculine gender tends to correlate. Thus, the issue that Arabic learners seem to encounter is that feminine gender is characterised by a positive correlation with noun ending, namely the presence of -a. In contrast, masculine gender is characterized by a negative correlation with noun ending, namely the lack of a gender marker. If Arabic is compared to Spanish for example, we find that in Spanish a fair proportion of the masculine nouns end in -o compared to a fair proportion of the feminine nouns that end in –a. Both genders positively correlate with noun endings. A similar and more transparent situation is found in Italian where the majority of the feminine nouns end in –a while the majority of the masculine nouns end in –o. The situation in Arabic is not as balanced as the situation in Spanish or Italian, with one gender having an overt (positive) correlation with noun ending and the other having a lack of noun ending cue or negative correlation between gender and noun ending. As a result, when the nouns ended in –a, young children attributed feminine gender out of proportion. Yet, when -a was not there, young participants started to consult other types of cues such as syntactic agreement and natural gender in Experiment II.

The lack of the predictive ending -a could in principle in itself be a cue, just as its presence is. Yet, these two types of cues were not given similar weight by the learners. In Arabic, there is a group of nouns that have –a/no –a pairs (the nouns in
(64), repeated from chapter 3, give some examples). Arabic learners could learn through these nouns that the lack of –a is a cue to masculine gender in the same way that –a is a cue to feminine gender. Though this possibility seems to receive support from the patterns of responses found in Experiment III, it does not rule out the possibility that masculine gender is assigned by default to everything that lacks the feminine marker -a. In fact, the rate of masculine responses attributed to the nouns that lacked the -a ending in Experiment III was never as high as the rate of feminine responses attributed to the nouns that ended in -a in the same experiment. Two possible explanations can be offered here. First, there could be a preference among the learners for the cues to be overtly coded. For that reason –a is treated as a cue but the lack of it is not treated as a cue to the same degree. Frigo and McDonald (1998) found that in order for the gender-like subclass learning to take place, the cues need to be noticeable. Second, the grammatical distinction that the -al/ no -a pairs express correlates with the semantic distinction (i.e. male vs. female) which itself is not a salient cue for the young learners. 

64) 

a. ṭabīb (doctor)  ṭabīb-a (doctor-F)  
b. ṭālib (student)  ṭālib-a (student-F)  
c. kātib (writer)  kātib-a (writer-F)  

Yet, an alternative scenario for learning the morphophonological association between masculine gender and words’ endings would be what I will call
the diverse masculine endings strategy. It is possible that L1 Arabic learners are not simply learning the morphophonological association between nouns’ endings and grammatical gender as a two-way distinction, i.e. -a vs. lack of -a. Rather, they could be learning the association between -a and feminine gender and the association between the other endings, mainly consonantal endings, and masculine gender. Contrary to the uniform highly reliable feminine marker, the diverse masculine phonological endings are expected to differ in their degree of reliability. As I discussed in chapter 5, the corpora analysis in chapter 3 revealed that some masculine endings (i.e. stops) are more common in Arabic than other endings like approximants and long vowels. Therefore, the degree of reliability of a particular ending is predicted to be influenced by its frequency. When this strategy was considered in light of the results of Experiment II and Experiment III, mixed patterns were found. It is clear that the long vowels were not treated differently from the -a ending by all the age groups which might suggest that L1 Arabic learners may make an association between vowel endings and feminine gender. For the consonantal endings, with all the caveats mentioned in chapter 5 in mind, some masculine endings were weighted heavier than others as cues for masculine gender. The heavy weight given to some masculine endings in Experiment II was not always in accordance with the frequency of those endings. However, all the endings were weighted similarly in Experiment III where no syntactic cues were provided. Though there is no sufficient evidence to rule out the second strategy in favour of
the first strategy, the asymmetry in the marking of the two gender classes in Arabic is there and its effect in learning the gender paradigm is evident.

Similarly, the results demonstrated an asymmetry in the way the syntactic agreement was treated. Though feminine agreement was strongly associated with feminine gender with no differences between all the age groups, the association between masculine agreement and masculine gender followed a developmental trajectory. The association between masculine agreement and masculine gender seemed to gain more strength as participants grew older. This again could be due to the imbalance situation in Arabic. To explain, feminine agreement is overtly coded in Arabic whereas masculine agreement is a lack of agreement. For example, if the nouns in (64) were to be described by adjectives as in (65), it is always the case that feminine gender is overtly coded by the suffixation of -a compared to the masculine gender which is zero-coded.

65)

a. ṭabīb  muṣliṣ   ṭabīb-a  muṣliṣ-a
   doctor(M) dedicated(M)  doctor-F  dedicated-F
   ‘A dedicated doctor.’  ‘A dedicated (female) doctor.’

b. ṭālib  muʿaddab   ṭālib-a  muʿaddab -a
   student(M) polite(M)  student-F  polite-F
   ‘A polite student.’  ‘A polite (female) student.’
The same pattern is also attested in subject-verb gender agreement (the examples in 66). Feminine agreement is realized as a suffix in the verb whereas masculine agreement is zero-coded.

66)

a. 'il-ṭālib  daxal  'il-faṣil

The student entered the-classroom.

‘The student entered the classroom.’

b. 'il-ṭaليل-ا  daḥala-t  'l-faṣl

The student entered-F the-classroom.

‘The (female) student entered the classroom.’

Feminine nouns, determiners, and adjectives are not only morphophonologically overtly coded but also semantically more specific forms in Arabic.\textsuperscript{101} One case that shows the semantic specificity of feminine gender in

\textsuperscript{101}The semantic specificity of gender in Arabic is apparent only in human-denoting nouns because it is only in this subclass of nouns that grammatical gender correlates with the semantic properties of the nouns. Non-human-denoting feminine nouns are not in any way more specific than non-human-denoting masculine nouns because the gender of this subclass of nouns is semantically arbitrary. The latter subclass of nouns in Arabic interacts with number in a special way, which is out of the scope of this thesis (see Benduhais, 2018 for an extended discussion).
Arabic is that a group of females can only be described using a feminine plural adjective (see Bobaljik & Zocca, 2011 for examples from other languages). On the other hand, coordination of different gender groups can only be described with a masculine plural adjective similar to a group of males as the examples in (67), (68), and (69) illustrate.

67) ʾal-ʾsātiḍa wa ʾal-ʾustāḏ-āt muhtamm-īn

The-professor(MPI) and the-professor-FPl interested-MPl

bi-lbaḥṭ.
in-research

‘The professors (males and females) are interested in research.’

68) * ʾal-ʾsātiḍa wa ʾal-ʾustāḏ-āt muhtamm-āt

The-professor(MPI) and the-professor-FPl interested-FPl

bi-l-baḥṭ.
in-the-research

‘The professors (males and females) are interested in research.’

69) ʾl-ʾustāḏāt *muhtammān / muhtammāt bi-l-baḥṭ.

The-professor-FPl *interested-MPl / interested-FPl in-the-research

‘The (female) professors are interested in research.’
Another case that illustrates the semantic specificity of feminine gender is that in Arabic masculine titles can refer to males or can have a generic meaning whereas feminine titles can only refer to females; such as, ‘uṣṭāḍ vs. ‘uṣṭāḍ-ā (professor(M) vs. professor-F).

By and large, the morphophonological ending -ā and the feminine syntactically agreeing demonstratives and adjectives seem to have more strength and validity as cues to feminine gender compared to the lack of -ā/the presence of the masculine endings and the zero-coded masculine agreement as cues to masculine gender. According to MacWhinney et al., 1984, the strength of a cue in a language is a reflection of the relative validity of this cue in that language. One of the factors that contribute to the validity of a cue is being highly reliable, “never misleading or ambiguous” (p. 130). The -ā ending in Arabic and the feminine agreeing elements are more reliable when present compared to the lack of -ā/the presence of the masculine endings and the masculine agreeing elements as reflected in the discussion above.

This multidimensional asymmetry between masculine and feminine gender might suggest something about the representation of gender in general and in

---

102 It is worth mentioning that the strength of feminine agreement compared to masculine agreement cannot be simply due to the feminine agreement being more common in Arabic than masculine agreement. As I mentioned in chapter 3, though singular feminine agreement is used with plural non-human nouns in standard Arabic, this is not always the case in Arabic dialects like Hasawi Arabic. Additionally, it was found that singular nouns are more common in Arabic CDS than plural nouns (Al-Akeel, 1998; Alqaitani, 2016).
Arabic in particular. Is gender in Arabic single-valued or multi-valued? I will discuss this in the next subsection.

7.3.2. Masculine or lack of gender

The asymmetry in treating –a vs. no –a ending I found in the results and discussed above, as well as the asymmetry in the contexts in which masculine agreement is used vs. the contexts in which feminine agreement is used, as discussed in the previous subsection, makes it enticing at this point to suggest that masculine in Arabic is nothing but the absence of gender specification. This interpretation is congenial with the proposal that φ-features are privative not binary (see Ackema & Neeleman, 2013; Gazdar & Pullum, 1982; Harley & Ritter, 2002). However, before spelling out this conclusion explicitly, I need to have some pieces of evidence to argue that Arabic gender can work appropriately without the gender feature specification [masculine].

The generalizations in (70) and (71) below seem to be sufficient for describing the Arabic gender system from a linguistic point of view.\textsuperscript{103} Generalization (70) states that nouns denoting females in Arabic are grammatically feminine. Similarly, generalization (71) states that nouns ending in -a are grammatically feminine in Arabic. Elsewhere, nouns are masculine by default.

\textsuperscript{103} These two generalizations correctly assign gender to more than 90% of the 5,000 most frequently used words in Arabic. For now, it suffices to assume that the remaining 10% are exceptions that need to be stored in the lexicon, which is very common cross-linguistically.
Harris (1991) argued that this is the situation in Spanish. According to Harris, masculine in Spanish is actually the absence of [feminine] or the lack of gender and it is there by default. He also argued that -o and -a are not gender markers, rather they are declension class markers; both markers are found in masculine and feminine nouns and also in genderless items like adverbs. “They are unrestricted with respect to gender” (p. 59). According to Harris, the -o declension class is the default class unless the membership in this class is blocked by some other mechanism. The empirical evidence for this argument is the absence of feminine nouns and adjectives with the declensional marker -o derived from -a/-e-class masculine counterparts. Harris’ view was also supported experimentally by Fuchs et al. (2015), who found that gender in Spanish is single-valued\textsuperscript{104}. They investigated the feature content of both number and gender relying on agreement attraction: “a situation in which ungrammatical sequences are perceived as grammatical when one of the NPs is erroneously identified as determining agreement” Fuchs et al. (2015, p. 703). They argued that if the feature in question is multi-valued, i.e. both of its values are specified, both values should have a

\textsuperscript{104} See also Adamson and Sereikaite (2019) for the same line of argument for the representation of gender in Lithuanian.
grammaticality effect. However, if the feature is single-valued, i.e. only one of its values is specified, one value but not both should have a grammaticality effect. They found that feminine but not masculine had an effect on grammaticality in their experiment, which supports Harris proposal that masculine is represented as the absence of gender specifications, at least in Spanish.

The representation of gender features as single-valued has been suggested for many languages. Adamson and Sereikaite (2019) argued that in Lithuanian, neuter is the absence of gender, masculine is the unmarked gender [GEND] whereas feminine is the marked gender [FEM]. This proposal seems also to work for the representation of some Dutch dialects where there is still a distinction between feminine, masculine, and neuter, as well as for the standard dialect where the distinction is between common and neuter (Audring, 2009, 2014).

The examples in (67), (68), and (69) mentioned in section 7.3.1 provide the first evidence for the non-specification and default status of the masculine agreement in Arabic. As those examples illustrate, coordinations of different genders (example 67) require masculine agreeing adjectives by default just as a group of males does. Only the group of females require feminine agreeing adjectives (example 69) which is in line with generalization (70). Another piece of evidence comes from mixed-gender groups. When a noun is used in the masculine form, for example, ʾalʾṣātiḍa (professors), it can refer to both male and female

\[\text{\footnotesize{\textsuperscript{105} Boloh and Ibernon (2010 & 2013) argued that masculine is used as a default gender by French-learning children as I discussed in chapter 2.}}\]
professors. Whereas when the noun is used in the feminine form, for example, ‘al’ustāḏīt (professors-F), it can only refer to female professors. I cannot think of a situation that requires the specification of the gender feature as [masculine] in Arabic. Even when gender interacts with number in a special way in Arabic, which is out of the scope of my thesis, this interaction does not require the feature specification [masculine]. In Arabic, non-human plurals are assigned feminine gender no matter what gender they have in their singular forms. I will go back to this interaction when I introduce ‘animacy’ in section 7.3.4.

The results of my experiments provide further evidence in support of the representation of gender as single-valued in Arabic. The generalization in (71) is manifested in the behaviour of Arabic speakers in all age groups. The generalization in (70) though present as early as 2 years and 6 months, seems to gain strength as Arabic learners progress in age. Thus, my data brings additional support from Arabic gender acquisition to the representation of gender as a single-valued feature. In the semi-spontaneous production task used as a screening criterion (chapter 4), masculine gender attributed to grammatically feminine nouns was more common than feminine gender attributed to grammatically masculine nouns among the young participants. Notwithstanding that the children’s performance in this task could reflect rote learning, overgeneralization of masculine gender reflects that this gender was used as a default gender where the gender of a noun has not yet been

---

106 See footnote 102 for reference.
learned. Furthermore, the asymmetry attested in experiment II strongly suggests that feminine gender is specified as [feminine] while masculine gender is the absence of gender specification.

Related to the discussion here is the debate on rule/pattern learning. It is interesting to see if my data supports representing the generalizations in (70) and (71) as abstract rules or not. I will discuss this in the following section by consulting Yang’s Tolerance Principle.

7.3.3. Productive vs. unproductive: the Tolerance Principle

Yang (2005, 2016) proposed the Tolerance Principle [TP] (72) that distinguishes between productive and unproductive rules in language learning. According to Yang, the distinction between productive and unproductive rules\textsuperscript{107} is categorical, i.e. a rule is either productive or not. This categorical distinction, according to Yang, is supported by the patterns attested in morphology acquisition in different languages such as the acquisition of English past tense and German participle. The Tolerance Principle predicts not only when to generalize but also when not to generalize.

\begin{align*}
\text{72) The Tolerance Principle} \\
\text{Let } R \text{ be a rule that is applicable to } N \text{ items, of which } e \text{ are exceptions. } R \text{ is productive if and only if } e \leq 0_N = N/\ln (N).
\end{align*}

\textsuperscript{107} As I mentioned in the introduction (chapter 1), in this section, I am using ‘rule’ as an abstract algebraic representation.
According to Schuler et al. (2016), “It is only more efficient to form a productive rule when the number of exceptions is less than the number of items divided by the natural log of the number of items.” To put it simply, if the exceptions to the rule are so many, it would be computationally more economical and more efficient to store all the items as exceptions rather than as a rule and exceptions because in order for the rule to work, it has to check all the exceptions first. Otherwise, the rule is productive. “Computational efficiency is computed by calculating the time complexity required for forming a rule with the time complexity required for accessing individual lexical forms.” (Björnsdóttir, 2021, p. 1212). Note that the tolerance of a rule depends on the size of the lexicon at a particular stage of development and it might change as the lexicon expands and the number of exceptions changes. That is, a particular rule could be productive at a particular stage but become unproductive when some more items are added to the lexicon and the number of exceptions changes. Similarly, a rule could be unproductive because the number of exceptions is above the threshold predicted by the Tolerance Principle, but it becomes productive when more items that comply to it are added to the lexicon making the number of exceptions below its threshold.

As stated in Yang (2016, p. 79), “the execution of the Tolerance Principle follows a tedious mechanical routine:

a. Obtain a rule $R$ along with its structural description and structural change.
b. Count $N$, the number of lexical items that meet the structural description of $R$.

c. Count $e$, the subset of $N$ that are exceptions to $R$.

d. Compare $e$ and the critical threshold $\theta_N = N / \ln N$ to determine productivity.”

Björnsdóttir (2021) showed how the TP correctly predicts the productivity of $-ø$ as a cue to masculine and $-a$ as a cue to feminine and also the productivity of masculine as a default in the absence of $-o/-a$ in Spanish gender. The main focus of her paper was evaluating the predictions made by the TP in accounting for the correlation between noun suffixes ($-r$, $-i$, $-a$, $-ø$) and grammatical gender in Icelandic. Those suffixes are nominative singular case suffixes which exhibit a high degree of correlation with gender in Icelandic. That is, it is typical for a noun that takes the nominative singular suffix $-r$ or $-i$ to be masculine. Similarly, it is typical for a noun that takes the nominative singular suffix $-a$ to be feminine. Those tendencies have some exceptions, as is the case in almost all gender systems, as exemplified in Table 7.1 (Björnsdóttir, 2021, p. 1215). There is no productive nominative singular suffix for neuter gender in Icelandic and the absence of a suffix is not always predictive of neuter as the examples in (73) illustrate.

73) a. *Borg-ø / torg-ø.*

    city-F / square-N

    ‘A city / a square.’
Table 7.1. Mappings between gender and nominative singular suffixes in Icelandic.

<table>
<thead>
<tr>
<th></th>
<th>-r</th>
<th>-i</th>
<th>-a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Masculine</td>
<td>sokku-r (a sock)</td>
<td>burst-i (a brush)</td>
<td>NA</td>
</tr>
<tr>
<td>Feminine</td>
<td>brúðu-r (a bride)</td>
<td>NA</td>
<td>fat-a (a bucket)</td>
</tr>
<tr>
<td>Neuter</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

Björnsdóttir (2021) consulted three types of corpora (child-directed speech, child spontaneous speech and adult speech) to evaluate the predictions made by the Tolerance Principle for the productivity of those patterns. In particular, the TP predicts -r and -i to be productive of masculine gender and -a to be productive of feminine gender because the exceptions to those suffixes are below the exception threshold ($\theta_N$). However, in the absence of those suffixes, no gender is predicted to be productive. The predictions of the TP were born out by the data obtained from the three corpora. The TP predictions were further tested using an elicited production task in which 26 Icelandic children between the ages of 2;09 and 6;03 in addition to 18 adult controls took part. Novel words with and without productive gender-correlated noun suffixes were used in the experiment. The results of the experiment confirmed the results obtained from the corpora analysis. No significant
difference was found between children’s and adults’ performance. According to Björnsdóttir, the results of this investigation do not support a gradient view of productivity. Rather, productivity seems to be categorical, i.e. rules were either productive or unproductive as predicted by the Tolerance Principle.

In the same manner, I will try to consult the TP to evaluate the productivity of the generalizations in (70) and (71). Because the TP depends on the size of the lexicon at a particular stage, I will consult Salama’s (2016) corpus which was obtained from 10 Egyptian children (1;7-3;8 years) interacting with their Egyptian mothers or with the instructor. This corpus constitutes a good estimate of the size of the Arabic-learning child’s lexicon.

The children in this corpus produced 297 singular noun types. There were 157 (52.9%) masculine nouns compared to 140 (47.1%) feminine nouns. All the masculine nouns produced by the children had no –a ending except 2 nouns. On the other hand, 15 feminine nouns had no -a ending compared to 125 nouns which ended in -a. The numerical distributions of those nouns and the predictions of the TP are presented in table 7.2. ¹⁰⁸

¹⁰⁸ Other sounds are found at the end of the words in this corpus (z, d, s, d, b, g, ħ, ḥ, q, t) but they are not presented in this table for two reasons. First, those sounds are found at the end of three or fewer than three words produced by the children. Second, those sounds were not selected as endings of the no -a words in my experiments, i.e. the predictions made by the TP regarding these words cannot be compared to the children’s behaviour in the experiments.
Table 7.2. Numerical distribution of the singular nouns produced by the Egyptian children in Salama Corpus (Salama & Alansary, 2016) by ending and gender.

<table>
<thead>
<tr>
<th>Ending</th>
<th>M</th>
<th>F</th>
<th>$0_N$</th>
<th>Productive?</th>
</tr>
</thead>
<tbody>
<tr>
<td>-a</td>
<td>2</td>
<td>125</td>
<td>127/ln127 = 27</td>
<td>Yes ($2 &lt; 27$)</td>
</tr>
<tr>
<td>b</td>
<td>16</td>
<td>0</td>
<td>16/ln16 = 6</td>
<td>Yes ($0 &lt; 6$)</td>
</tr>
<tr>
<td>k</td>
<td>5</td>
<td>0</td>
<td>5/ln5 = 3</td>
<td>Yes ($0 &lt; 3$)</td>
</tr>
<tr>
<td>m</td>
<td>12</td>
<td>0</td>
<td>12/ln12 = 5</td>
<td>Yes ($0 &lt; 5$)</td>
</tr>
<tr>
<td>n</td>
<td>16</td>
<td>3</td>
<td>19/ln19 = 7</td>
<td>Yes ($3 &lt; 7$)</td>
</tr>
<tr>
<td>r</td>
<td>19</td>
<td>1</td>
<td>20/ln20 = 7</td>
<td>Yes ($1 &lt; 7$)</td>
</tr>
<tr>
<td>l</td>
<td>16</td>
<td>1</td>
<td>17/ln17 = 6</td>
<td>Yes ($1 &lt; 6$)</td>
</tr>
<tr>
<td>s</td>
<td>10</td>
<td>2</td>
<td>12/ln12 = 5</td>
<td>Yes ($2 &lt; 5$)</td>
</tr>
<tr>
<td>t</td>
<td>8</td>
<td>3</td>
<td>11/ln11 = 5</td>
<td>Yes ($3 &lt; 5$)</td>
</tr>
<tr>
<td>d</td>
<td>8</td>
<td>1</td>
<td>9/ln9 = 4</td>
<td>Yes ($1 &lt; 4$)</td>
</tr>
<tr>
<td>š</td>
<td>9</td>
<td>0</td>
<td>9/ln9 = 4</td>
<td>Yes ($0 &lt; 4$)</td>
</tr>
<tr>
<td>Long vowels</td>
<td>12</td>
<td>1</td>
<td>13/ln13 = 5</td>
<td>Yes ($1 &lt; 5$)</td>
</tr>
<tr>
<td>(ā/ī)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>female</td>
<td>0</td>
<td>9</td>
<td>9/ln9 = 4</td>
<td>Yes ($0 &lt; 4$)</td>
</tr>
</tbody>
</table>

As illustrated in table 7.2, the TP predicts that it is productive to learn not only the generalizations in (70) and (71) (repeated here as 74 & 75) as rules but also to learn the association between masculine gender and the endings presented here as rules (76 – 85) because the number of exceptions to all these rules is below the threshold suggested by this principle.

74) Female PERSON $\rightarrow$ [feminine]  
75) -a $\rightarrow$ [feminine]  
76) b $\rightarrow$ [masculine]  
77) k $\rightarrow$ [masculine]  
78) m $\rightarrow$ [masculine]  
79) n $\rightarrow$ [masculine]  
80) r $\rightarrow$ [masculine]  
81) l $\rightarrow$ [masculine]
The results of Experiment II give the opportunity for evaluating the TP predictions shown in table 7.2. If the youngest children, who were found biased to follow the morphophonological ending cue for grammatical gender classes, had learned the rules (76 – 85), the expectation is that the rate of feminine responses attributed by them to the novel items that ended in each of those sounds was significantly lower than the rate of feminine responses they attributed to the novel nouns that ended in -a. To test this, I ran a one-way within-subjects ANOVA with the sound as the independent variable. The results of this test revealed no significant effect of sound, \( F(10, 49) = 1.79, p = 0.087 \). This result (Fig.7.4) might suggest that the young learners were not learning the rules predicted to be productive by the Tolerance Principle. However, this insignificant result could well be due to the lack of statistical power. The number of participants in this age group is very small (N = 7). So, it could be that the learners were actually learning the rules predicted to be productive by the Tolerance Principle but my data is not enough to reveal that. Still a third possibility is that the learners were not learning all the rules suggested in (75 – 85), rather they were only learning the two rules (74) and (75) as rules and the others as “everything else”. That is, it could be that the learners were assigning feminine gender by rules while assigning masculine gender as the default or the elsewhere gender. Learning the morphophonological generalization (75) as a rule

\[
\begin{align*}
82) s & \rightarrow [\text{masculine}] \\
83) t & \rightarrow [\text{masculine}] \\
84) d & \rightarrow [\text{masculine}] \\
85) \& & \rightarrow [\text{masculine}] 
\end{align*}
\]
receives support from the current investigation. When the rate of feminine responses attributed by the youngest participants to the nouns that ended in -a was compared to the rate of feminine responses attributed to the nouns that did not end in -a in general (i.e., not as separate endings as in table 7.3) in Experiment II, the result was significant, $F(1, 12) = 23.01, p = 0.0004$. This behaviour is in support of learning rule (75) ($-a \rightarrow \text{[Feminine]}$) and assigning masculine gender by default to the nouns that lack this ending. Thus, all learners need to learn is the association between -a and feminine gender and elsewhere, masculine or as I argued in the previous section, the lack of gender is assigned by default. This is correctly predicted by the TP as presented in Table 7.3.

Fig. 7.4. Mean rate of feminine response attributed by the youngest group to the different endings. Error bars represent standard errors of the mean (SEM).
Table 7.3. Numerical distribution of the singular nouns produced by the children in Salama Corpus (Salama & Alansary, 2016) by ending and gender.

<table>
<thead>
<tr>
<th>Ending</th>
<th>M</th>
<th>F</th>
<th>$\theta_N$</th>
<th>Productive?</th>
</tr>
</thead>
<tbody>
<tr>
<td>-a</td>
<td>2</td>
<td>125</td>
<td>$127/\ln 127 = 27$</td>
<td>Yes ($2 &lt; 27$)</td>
</tr>
<tr>
<td>Ø</td>
<td>155</td>
<td>15</td>
<td>$170/\ln 170 = 34$</td>
<td>Yes ($15 &lt; 34$)</td>
</tr>
<tr>
<td>total</td>
<td>157</td>
<td>140</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Though the TP also predicts the semantic rule (74) to be productive, the results of Experiments II and III showed that the youngest participants were not sensitive to it. As I discussed in section 7.2, the youngest children are biased to learn generalization (75) and they show emergent sensitivity to generalization (74). Generalization (74) requires time to be integrated as part of the gender system of the Arabic speakers. The behaviour of the older subjects showed that this generalization takes precedence over the morphophonological generalization (74). This brings me to the debate about rule hierarchy/ordering, which I will discuss in the next section. In sum, it is hard to decide on how the youngest participants were approaching the task of gender learning due to their small number. In particular, whether they were assigning feminine gender by rule while attributing masculine gender to everything else by default, or whether they were constructing a number of productive rules as in (75 - 85) cannot be determined by the current investigation.
Yet, what seems to be uncontroversial from the current investigation is that whatever representation of grammatical gender young children are constructing, there is no doubt that this representation has the morphophonological properties of the nouns but not their semantic properties as part of it. Nonetheless, the two possibilities, i.e. learning rules for everything such as the rules in (74 -85) vs. learning some rules and a default as in (70 and 71), suggest different scenarios. To explain, if the child is constructing/learning the rule that assigns feminine gender to nouns that end in -a while attributing masculine gender by default to everything else, then when the child encounters a noun that ends in -a, they need to check the exceptions first and if the noun is not among the list of exceptions it is assigned feminine gender. This would prevent the assignment of feminine gender to masculine nouns ending in -a which should be listed as exceptions to the rule. It is clear from the current investigation that the young children are not constructing the semantic rule (70). This predicts that feminine nouns that do not end in -a would be wrongly assigned masculine by default. This prediction is partly supported by the behaviour of the youngest participants in Experiment III. Young participants attributed masculine gender to novel nouns that referred to female humans but did not end in -a. However, attributing masculine gender to familiar nouns referring to female humans was very rare in the same experiment. This indicates that not everything is learned as rules and exceptions as Yang claims. The young children were correctly attributing feminine gender to familiar nouns that referred to female humans but did not end in -a despite that they did not seem to be sensitive to the
correlation between natural gender and feminine gender in their language for which they could have formulated a rule like the one in (70). An alternative explanation would be that the young children learn two rules (86 and 87) each of which has its exceptions. In this case, rule (86) takes precedence over rule (87) because it is more specific. So, when the child is encountered with a noun that does not end in -a, whether it refers to a female person or not, they need to check the exceptions of rule (87) among which will be words like bint (girl) and šams (sun). Learning those two rules and their exceptions would perfectly assign the correct gender to Arabic human-referring and non-human-referring nouns. Learning those two rules also receives support from the youngest participants behaviour in Experiment III.

\[
\begin{align*}
86) & \quad -a \rightarrow \text{[feminine]} \\
87) & \quad \emptyset \rightarrow \text{[masculine]}
\end{align*}
\]

On the other hand, if children were to learn all the rules (75 - 85), but not the semantic rule (74), the learner has to learn the exceptions to each separate rule. So, for example, when the child encounters the noun bint (girl), rule (83) (t \rightarrow [masculine]) is activated but before assigning masculine gender to bint according to the rule, the child has to check the exceptions to this particular rule among which bint is found. Similarly, when the child encounters the noun bāša (pasha), rule (52) (-a \rightarrow [feminine]) is activated but before assigning feminine gender to bāša, the child has to check the exceptions to this rule among which bāša is found. So basically, if a noun is among the list of exceptions to a rule that assigns masculine
gender, it will be assigned feminine gender. If the noun is among the list of exceptions to a rule that assigns feminine gender, it will be assigned masculine gender. As I mentioned above, it is hard to rule out this possibility because of the small number of the young participants in this project and because almost all the sounds were represented by one experimental item. Nonetheless, learning all the morphophonological rules (75 – 85) and their exceptions would assign the correct gender to Arabic human-referring and non-human referring nouns similar to learning only the two rules (86) and (87).

To sum up this section, Yang’s TP tries to generate ‘everything’ by rules. Those rules vary on their degree of specification. The most general rule applies by the elsewhere principle unless there is a specific rule that blocks its application. Therefore, a default rule is an essential component for an account that imports the mathematical assumptions of the TP. Thus, rules (86) and (87) would be appropriate candidates for the TP because rule (87) is more general than rule (86). Though rules (75 - 85) are all specified in some way, there is no evidence from the current investigation to rule out the possibility that L1 Arabic learners are learning such rules.

7.3.4. Rule hierarchy/ ranking

I argued in section 7.3.2 that the generalizations in (70) and (71) are sufficient to predict the correct grammatical gender of more than 90% of the 5,000 most frequently used words in Arabic (listed in Buckwalter & Parkinson, 2011). I also presented evidence that the feature specification [masculine] is not needed for the
Arabic gender system to function properly and that masculine gender can be assigned by default in cases other than the specific cases accounted for by the two generalizations (70) and (71). I also tested if these generalizations are predicted to be productive according to Yang’s Tolerance Principle in the previous section. I have not yet said anything about the ordering of these two generalizations in order not to make wrong predictions about the gender of nouns with conflicting cues, i.e. nouns whose meanings correlate with one gender whereas their forms correlate with the other gender. For example, ʾuxt (sister) in Arabic is feminine though it has a form typical of masculine nouns, namely it does not in -a.

This brings up the hot debate about the hierarchy of the rules in gender assignment for which I presented different accounts in chapter 2. In this section I will reconsider these accounts in light of the results of this investigation. Notice that, as I mentioned in chapters 1 and 2, the reviewed accounts were proposed to account for the gender distribution across nouns in a language. But I am trying to consider the potential implications of these accounts for the end state of the acquisition process of the grammatical gender system as represented in the generalizations made by the adult language users when attributing gender to nouns in order for those nouns to function as controllers of gender agreement with other elements.

In the Arabic gender system, the generalizations in (70) and (71) must be ranked such that the semantic one (70) takes precedence over the formal one (71). This ranking is necessary to prevent the assignment of feminine gender to nouns
referring to male humans in Arabic but ending in –a109 or the assignment of the default masculine gender to nouns referring to female humans but lacking the -a ending such as bint (girl) and ’umm (mother). Rice’s (2006)110 Optimal gender assignment theory which assumes no “group wise” ranking, i.e. the ranking of the rules is not determined by them being semantic vs. morphophonological in nature, but works on the basis of rule counting and gender markedness, would predict the correct gender in the second case but not in the first case. Nouns referring to male humans but ending in -a such as ħalīfa (caliph) would be wrongly assigned feminine gender because the only possible generalization in this case is (71) which states that when a noun ends in -a, it is assigned feminine gender. bint and ’umm would be assigned the correct feminine gender by generalization (70).

Likewise, Nesset’s (2006) proposal would assign feminine gender to bint, and similarly to ’umm, as illustrated in Fig. 7.5 To explain, schema (a) includes information related to biological sex111 which according to the Core Semantic Override Principle (88) takes precedence over other schemas. Therefore, bint is assigned feminine gender because schema (a) takes precedence over schema (b), the activation of which might result in assigning masculine gender by default to bint

109 This combination is very rare in Arabic. In the frequency dictionary of Arabic consulted in this study, there were only three nouns that refer to male humans and end in –a.

110 When referring to Rice’s and Nesset’s accounts, I’m not excluding other accounts proposed for explaining the interaction between gender assignment rules, but I’m taking Rice’s account as a representative of the accounts that assume no type-wise ranking and Nesset’s account as a representative of the accounts that assume the precedence of semantic and particularly biological sex precedence.

111 Biological sex information is indicated by this schema being shaded following Nesset’s style.
because it does not end in –a. *walad* (boy) would be assigned masculine gender by default because neither of the schemas will be activated as in Fig. 7.6.

88) **Core Semantic Override Principle**

“Rules referring to biological sex take precedence in gender assignment.” (Nesset, 2006)

Fig. 7.5. Gender assignment to *bint*. 
Fig. 7.7 shows that while duktūr-a (female GP) ends in –a, schema (a) is already sufficient to assign it feminine gender because it refers to a female person. However, Fig. 7.8 shows that similar to Rice’s proposal, in Nesset’s proposal, xalīf-a (caliph) could be wrongly assigned feminine gender because it does not refer to a female person but it ends in -a which activates schema (b). At this point, it seems that a third generalization is needed in order to assign the default masculine to male persons that end in -a.\textsuperscript{112} This generalization is suggested in (89) in addition to generalizations (70) and (71) (repeated here as 90 and 91 respectively).

\textsuperscript{112} Notice that adding this generalization would also predict the assignment of the correct masculine gender to xalīf-a in Rice’s account, albeit with no reference to the content of this generalization. In Rice’s account generalization (64) would compete with generalization (66) presenting what Rice dubs a balanced conflict. In this case, xalīf-a would be assigned masculine gender because masculine is the least marked gender in Arabic.
According to generalization (89), a noun that refers to a person is assigned masculine gender by default which is represented as the lack of gender specification as I argued in section (7.3.2). Generalization (90) which has more information than generalization (89) indicates that if the person is female, it is assigned feminine gender. Finally, generalization (91) assigns feminine gender to nouns that end in -a. Thus, *xalīfa* (caliph) would be assigned the default masculine as in Fig. 7.9 because schema (a) takes precedence over schema (c) according to the Core Semantic Override Principle. Schema (b) is not competing in this case because *xalīfa* does not refer to a female person.

89) PERSON → [..]

90) PERSON/Female → [feminine]

91) -a → [feminine]

---

Fig. 7.7. Gender assignment to *duktūr-a*. 

---
Fig. 7.8. Gender assignment to *xalīfa* (wrong assignment).

Fig. 7.9. Gender assignment to *xalīfa* (correct assignment).
Fig. 7.10 and 7.11 also illustrate the assignment of feminine gender to the non-human \(\tau\text{āwila}\) (table) and the default masculine gender to \(kurs\text{i}\) (chair).

---

**Legend:**
- — masc.
- — fem.
- 😏 winning cand.

---

**Fig. 7.10. Gender assignment to \(\tau\text{āwila}\).**

---

**Fig. 7.11. Gender assignment to \(kurs\text{i}\).**
Nonetheless, such ranking would not prevent the assignment of masculine gender to feminine nouns that refer to non-humans but do not end in –a such as *yad* (hand) and *nār* (fire). Though nouns with the latter properties represent less than 10% of the feminine nouns in the Arabic lexicon, they are among the most frequently used nouns by both child and adult speakers. Therefore, learning the gender of those nouns does not seem to be a problem for the learner as their gender can be learned on an item base and get entrenched over time. Item-based learning is very common in the early stages of development and the results of the familiar items in experiment III (chapter 6) give support to this pattern of learning. Therefore, the assignment of feminine gender to *šams* (sun) for example is suggested in Fig. 7.12. The gender of the feminine, non-human, no –a ending nouns in Arabic can be learned by rote which is suggested by having these nouns and their gender (as represented by the dashed boxes) as part of the grammar. It is plausible that the gender of some nouns might be learned before others depending on their frequency. It is plausible as well that gender assignment errors to those nouns could outnumber gender assignment errors to nouns that are more ‘regular’ due to the gender-form correlation which is the most predominant cue in the early stages of language learning. Indeed, the word *šams* (sun) which was one of the nouns used in experiment I, received the lowest rate of feminine responses compared to the other feminine items.
In general, the presented examples demonstrate that the gender system of the adult Arabic speaker having the three proposed generalizations (89, 90 & 91) could assign the correct gender to human as well as non-human referring nouns. I have shown that both Rice’s assignment proposal which assumes no ranking but a markedness hierarchy as well as Nesset’s proposal which assumes the precedence of the rules referring to biological sex would correctly assign gender to more than 90% of the Arabic lexicon. This also shows that the interaction between the different types of rules is very much related to what rules we accept in the first place. Before introducing rule (89) that assigns masculine gender to humans by default, both accounts could not assign the correct gender to nouns ending in -a but
refer to male humans. Going back to the behaviour attested in the experiments, the adult subjects, as well as adolescents and school children, were weighting the semantic properties of the nouns significantly heavier than their morphophonological properties. This behaviour could potentially imply that the generalizations suggested in (89 – 91) were ranked with the ones referring to biological sex taking precedence or that the least marked masculine gender was assigned in the case of a balanced conflict. Both possibilities could lead to the assignment of the correct masculine gender to the novel nouns that ended in -a but were represented as male humans as in the -a ending items used in Experiment III.

Deciding which of the aforementioned procedures the participants were following seems to be unfeasible. Despite this, for both of the accounts to predict the correct gender, they need to be sensitive to the semantic feature [Person/human] and the biological sex distinctions within this feature. I would therefore suggest the general Core Assignment Principle in (92). This principle is supported by the typological correlation between grammatical gender and biological gender which is found in three quarters of the languages which have semantic assignment systems (Corbett, 2013b). The Core Assignment Principle does not make claims about the ordering of the rules/generalizations which is indeed language specific and cannot be decided on unless one have access to the whole system of the language. What this principle dictates is that grammatical gender assignment rules in a language must include the feature [human/person] as one of their components.
92) Core Assignment Principle

Gender assignment principles must be sensitive to the feature [human] and, sometimes, the biological-sex distinctions within this feature.

This principle introduces the notion of animacy to the discussion. It is well known that animacy bears on a wide range of linguistic phenomenon such as case marking and agreement in addition to gender (Comrie, 1999; Dahl, 2000, 2008; De Swart & De Hoop, 2018). De swart & De Hoop distinguished between conceptual animacy which could be gradient or scalar and grammatical animacy which is discrete. According to them, what seems to matter in most linguistic phenomena, where animacy is of relevance, is whether something is animate/human or inanimate/nonhuman. In Arabic for instance the feature human/person bears on subject-verb and noun-adjective number and gender agreement. Briefly speaking, agreement with plural nonhuman nouns in Arabic has to be feminine singular (see Benduhaish, 2018 for a discussion). That being the case, to find an effect of animacy on gender is not surprising. According to Enger’s (2009, P. 1292) criteria for what could be a plausible semantic assignment rule, “if the rule invokes a particular ‘semantic feature’, it is an advantage if we recognize this feature from other parts of the grammar of the language in question.”

In my view, Nesset’s account fares better than Rice’s account because, as also pointed by Enger (2009), it is not clear how speakers of a particular language
would arrive at the markedness hierarchy in the first place\textsuperscript{113}. The mission could be more accomplishable for speakers of two-gender languages than for speakers of three-or-more-gender languages such as German and Norwegian. This is in part because it is not clear what a default gender is in different languages. It could be no more than “a dustbin category of no explanatory value” (Kilarski, 2001). So, unless we make it clear what a default gender is in a particular language, proposals based on markedness hierarchies seem to have many problems. Furthermore, it is not clear how the non-human feminine nouns that do not end in -a would be assigned the ‘more marked’ gender not the least marked gender according to the hierarchy. In Nesset’s usage-based account, such nouns do not seem to be at all problematic as they can be learned on an item based and get entrenched in the lexicon.

These accounts raise the question of whether a child acquiring their mother tongue needs to make generalizations about how the gender system is organised and eventually order these generalizations over time (Nesset’s account), or whether they needs to learn in addition to the plausible generalizations, the markedness hierarchy of the different genders in their language. It is clear from the current investigation, and from the investigations of other languages reviewed in chapter 2, that what the learner seems to do (after discovering that their language classifies nouns into classes) is making generalizations and adjusting their order over the course of development. Initially, morphological generalizations take precedence over

\textsuperscript{113} Enger (2009) also pointed out that such accounts are unfalsifiable because many counter examples to the markedness hierarchies can be saved by simply adding new rules (just as happened when I added generalization 89).
semantic (i.e. biological sex) generalizations while at the end, biological sex generalizations predominate over morphological generalizations.
Chapter 8: Conclusion and recommendations

“Whether cues are phonological, morphological, or semantic, fastest acquisition of subclass knowledge should occur when subclasses are systematically marked with highly noticeable cues...” Frigo and McDonald (1998, p. 240)

8.1. Conclusion and main contribution

This thesis was motivated by the lack of a comprehensive picture of the developmental trajectory of grammatical gender. I traced the developmental trajectory that Arabic grammatical gender undertakes until it reaches its end state. As setting the scene in the destination makes tracing the milestones along the way more attainable, I started by describing the gender system of the adult speakers/users of the language as reflected in their performance in the experimental tasks. Setting this representation of the gender system as the target towards which the younger participants’ representations were moving, I was able to find the discrepancies as well as the similarities between the representations of the grammatical gender system along the course of development. Looking at the acquisition of gender from this developmental perspective illuminates the gradual steps it takes until it arrives at its destination. Though syntactic agreement gains its role as the defining property of gender at a very early age (around 4 years), the competition between the noun-internal morphophonological and semantic intrinsic
properties moves in a more gradual fashion. As revealed by this investigation, preschoolers were not biased by the morphophonological properties of the nouns and they attributed gender to the nouns according to the class indicators, i.e. agreeing demonstratives and adjectives. Overcoming the predominance of morphophonological properties over semantic properties was documented in school-age children. Yet, the effect size of the semantic information provided by natural gender in this age group was smaller than the effect size it had in the two older age groups.

Grammatical gender is a morphosyntactic feature the mastery of which is manifested in assigning the correct gender value to the noun which controls the behaviour of other elements syntactically related to this noun. Thus, agreement is the ultimate goal of gender acquisition. When the adult participants were presented with a new noun whose gender was defined by an agreeing demonstrative and an agreeing adjective, they relied on this defining information as a cue to the gender of the noun (Experiment II). When those participants were presented with a new noun without a defining syntactic context, they relied on the intrinsic properties of the noun itself to determine its gender (Experiment III). Yet, they were selective on the type of noun-internal properties they used as a cue to its gender; the semantic properties of the noun were more important for them than its morphophonological properties. Nonetheless, when the noun was presented in a syntactic context suggesting a gender class different from the gender suggested by its intrinsic semantic properties, the participants weighted the two types of information
relatively equally (Experiment II). Both types of information are central to the adult users of the language as reflected in the large effect size that both types of information had in this age group.

The important role of the noun-internal properties as cues to grammatical gender is obvious in the early stages of development. The youngest participants in this investigation were relying on the intrinsic morphophonological properties of the noun to produce the agreeing forms. It is the noun marker that decides on the form of the gender indicator. This behaviour could support the serial model suggested by Frigo and McDonald (1998) where learners start by linking noun markers (here the ending) to class indicators (demonstratives and adjectives). Then, the learner links the class indicators to each other. However, the asymmetry attested when -a was present vs. when -a was absent as a noun marker, as well as the asymmetry in treating the masculine and the feminine class indicators, suggest that both processes may start simultaneously. What seems to matter for the young children is the nature of the information that could be used as a cue to the gender class. As Frigo and McDonald, (1998: 239) state, “[c]hildren appear to base their original grammatical gender classification structure on cues that are frequently present and systematically mark the subclass”. The presence of the -a at the end of the majority of feminine nouns in Arabic increases the detectability and the reliability of it as a cue to feminine gender. The same is done by the presence of -a in the class indicators, in particular adjectives. Contrary to this, the lack of a systematic marker in the masculine nouns in Arabic weakens the association
between the lack of -a and masculine gender and lowers the possibility of detecting the lack of it as a cue to masculine gender in the first place.

In languages like Spanish and Italian, both masculine and feminine nouns correlate with reliable noun markers which in turn correlate with obligatory class indicators. In other words, both genders are overtly coded. This overt-coding contributes to the early detectability of these endings as cues to gender class membership. Unlike in Arabic where the association between feminine gender and -a ending is stronger than the association between the lack of -a and masculine gender, in those languages both genders are associated with overtly-coded markers. It is not surprising therefore, that those endings retain their role as cues to noun gender up to adulthood (Afonso et al., 2014). Because the gender markers are present for both classes, they do not undergo the same sort of devaluation over the course of development as happens for the role of the morphophonological ending in Arabic. In addition, the obligatoriness of the class indicators in languages like Spanish strengthens the linkage between the markers and the indicators and also establishes the link between the class indicators more easily. The presence of one class indicator triggers the other in experimental contexts as with definite and indefinite articles (Pérez-Pereira, 1991).

The overt coding of gender cues is interconnected with the frequency of those cues in a language. In Arabic, as seen in this investigation, the -a ending is a very frequent cue as it marks more than 99% of the feminine inanimate nouns in Arabic. In Spanish and Italian, -a and -o mark a fair proportion of the feminine and
the masculine nouns respectively in those languages. On the other hand, languages like French and German\textsuperscript{114} have a number of class markers that vary in their frequencies (see Zubin & Köpcke, 1981 for German; Boloh & Ibernon, 2010 for French). It has been found that the association between grammatical gender and high frequency markers was easier to detect by children learning those languages than the association between grammatical gender and low frequency markers. Thus, in order for a marker to be used as a cue to gender class membership, it has to be overtly coded and noticeable, systematic and reliable, and frequent enough to be associated with the gender with which it correlates.

Even if L1 Arabic-learning children are learning the association between masculine gender and the different endings found in Arabic masculine words, the asymmetry between those endings and the feminine marker -\textit{a} is evident in the patterns attested in this investigation. The heterogeneity of the masculine endings makes them less reliable cues than the uniform feminine marker -\textit{a}. Additionally, none of the masculine endings is suffixal in nature, i.e. none of them is used to derive the opposite gender in the same way as the feminine marker -\textit{a} is used.

Moving to the semantic properties with which the gender of the nouns correlates in Arabic, the correlation between natural gender and grammatical gender in Arabic is not only systematic but deterministic and perfect. It is with no exceptions.\textsuperscript{115} This consistency of natural gender as a cue to grammatical gender in

---

\textsuperscript{114} Gender in German is also complicated by the fusion it has with case.

\textsuperscript{115} In the singular nouns which are the focus of this thesis.
Arabic is a potentially important contributor to the early sensitivity attested to it in this investigation. The correlation between grammatical gender and natural gender in languages such as Arabic, French, Spanish, Italian, and Russian among others, has to be learned and sometimes has to take precedence over the morphophonological properties in order for the grammatical system to work properly. To put it differently, an Arabic learner has to learn that *bint* (girl) is assigned feminine gender not masculine gender because it refers to a female human though it does not end in -*a*. Although *bint* and similar words can be learned as exceptions to the morphophonological assignment principles, the empirical evidence obtained from the current investigation does not support that. It is clear from the behaviour of the three older age groups that the association between natural gender and grammatical gender is part of their representation of the grammatical gender system. It has been reported that the association between natural gender and grammatical gender takes up to adolescence to be part of the representation of grammatical gender systems of languages like Spanish and French. It would be interesting to find how gender is assigned by children to novel nouns in languages with purely semantic assignment systems like Tamil and other Dravidian languages. The late sensitivity to semantic information in general and natural gender in particular as cues to grammatical gender suggests that the gender system of such languages might be learned at a relatively late age than the gender system of a language with formal or morphophonological gender system.
All in all, in their attempt to acquire gender, learners seem to look for systematically reliable cues. No matter whether morphophonological or semantic, the winners are the more deterministic ones.

8.2. Limitations and future directions

One of the major limitations of this research is the small number of the participants in the youngest age group. This was due to a number of reasons. The first reason was the critical time in which the data was collected. The initial intention was to recruit participants from nurseries and schools. However, due to the nurseries’ and schools’ closure because of the Covid-19 pandemic, participants had to be recruited through social media and personal communication with friends and relatives. Though many people were happy and willing to participate, most were hesitant because of the uncertainty of the situation. The second reason was the criteria used to filter the participants towards experiments II and III. More than 50 parents gave consent for their children to take part in the study. Yet, only 26 participants passed the MGIT test successfully. Of those 26 only 10 participants passed the semi-spontaneous production task (chapter 4, screening criterion II). Three out of those 10 participants were excluded from the final analyses of Experiments II and III due to failure to understand or to complete the tasks. The last reason is simply that it was not easy to elicit informative responses from young participants or to get them engaged in the tasks. Many children refused to say anything. Children’s unwillingness to respond cannot be due to the experimental design or setting. If that
was the reason, I would expect an effect of the experimental design or setting on the other age groups as well not only on the youngest children. The children were too young to say anything. They even refused to respond to their mothers not only to the experimenter. Despite the small number of the participants in the youngest age group, the pattern found in their responses especially regarding the use of the semantic properties of the nouns as cues to the gender class membership, is very informative because it is based on solid grounds. All those 7 participants passed the MGIT test successfully meaning that they have acquired the conceptual meaning of natural gender or biological sex, but they were not integrating it as part of their representation of the grammatical gender system. It is worth mentioning that the seminal work by Karmiloff-Smith (1979) which influenced many later gender acquisition studies had only 24 subjects between the ages of 3;02 and 3;11. Those 24 subjects participated in five different experiments, making the number of the participants in each experiment even smaller than the number of the young participants in this thesis. Thus, the minimum age recommended for an elicited production task that uses agreement as a test is three years. Gender acquisition in ages younger than three may require a simpler experimental task.

As mentioned at the beginning of this thesis, the acquisition of gender is intertwined with syntactic agreement. Gender agreement is what learners ultimately need to master because agreement is the way by which gender is expressed. Therefore, a recommended step for future research is to set cross-linguistic criteria for gender acquisition. How can we conclude that gender has been acquired by a
language learner? Different tests are being used to investigate gender acquisition. In languages with obligatory determiners such as French, Spanish, and Portuguese, gender-marked determiners and sometimes adjectives are used as tests for the acquisition of gender (Correa and Name, 2003; Correa et al., 2011; Karmiloff-Smith, 1979; Pérez-Pereira, 1991; Seigneuric et al., 2007, among others). In my investigation, I used adjectives and demonstratives as the gender defining elements in the prompt sentences that I provided the participants with in Experiment II. Whatever gender agreeing targets the participants produced in their responses were included in the analyses. It is still not clear however on the basis of what criteria one can conclude that gender has been acquired by the language learner. Languages vary with regard to gender agreement targets. Gender agreement with some targets in a particular language may also vary according to the syntactic configuration. Adjectives in German for instance agree in gender with their head nouns while predicative adjectives are gender invariable (Corbett, 1991). Furthermore, acquiring gender agreement with some targets could be easier than with others. Acquiring gender agreement that is realized via a closed-class of determiners could be approached in a way different from acquiring gender agreement with an open-class of verbs or adjectives. By and large, a criterion for the acquisition of the morphosyntactic feature gender is essential as a baseline measurement. Studying the acquisition of grammatical gender apart from agreement seems to be an insoluble dilemma. Defining a criterion for measuring the acquisition of gender is a good starting point for solving this dilemma. Should one take all the gender
agreement targets in a particular language or should one focus on specific agreement targets when studying the acquisition of gender in this language, needs to be determined by future research. To put it differently, should the researcher who is investigating the acquisition of a gender system in which determiners, adjectives, and verbs are targets for gender agreement conclude from the correct use of the determiners only that gender has been acquired? Or should the researcher consider all the gender agreement targets in the language in question? This might seem a minor practical issue, but the non-canonicity of gender as a morphosyntactic feature requires additional consideration.

Finally, the design of the set of experiments reported in this thesis did not allow for the evaluation of the two suggested strategies that L1 Arabic learners may follow for learning the morphophonological association between words’ endings and masculine gender. This was not one of the aims of this project but it was one of the issues arose from the investigation. This would be an area for a future investigation that takes into account the different frequencies of the masculine endings.
References


Ferry, A., Nespor, M., & Mehler, J. (2020). Twelve to 24-month-olds can understand the meaning of morphological regularities in their language. Developmental Psychology, 56(1), 40.


Appendix A: The Arabic script for Experiment I with some example answers

الباحثة: قولي لي ياحلوة شنو عندي؟
الطفلة: (سمكة).
الباحثة: شو في عندي بعد سماكة. سماكة و سماكة. انتي أي سماكة تحبين؟
الطفلة: هذي.
الطفلة: هذي! طيب ليش؟
الطفلة: لأن هذي حبيتها.
الطفلة: طيب و شنو عندي بعد؟
الطفلة: (حصان).
الطفلة: و عندي بعد حصان. حصان و حصان. انتي أي حصان تحبين؟
الطفلة: الصغير.
الطفلة: وايش عندي بعد؟
الطفلة: (شمس).
الطفلة: أوكي شمس و شمس. طيب انتي أي شمس تحبين؟
الطفلة: هذي.
الطفلة: ليش؟
الطفلة: لأن ذي أحلى من ذي.
الباحثة: طيب ايش عندي بعد؟
الطفلة: (بقرة).
الطفلة: و عندي بعد بقرة. انتي أي بقرة تحبين؟
الطفلة: هذي.
الطفلة: ليش؟
الطارئة: بس.
Appendix B: Example Arabic script for Experiment II with example answers

الباحثة: الحين بنلعب لعبة بوريك فيها شخصيات و سأسألك عليهم أسئلة. نبدأ؟

الباحثة: هذي (كموش). مين هذي؟

المشارك: كموش.

الباحثة: و هذي كموش ثانية.

الباحثة (مسكة بالدميتين): هذي كموش و هذي كموش ثانية.

الباحثة (بعد أن أخفت إحدى الدميتين خلف ظهرها): الحين قول لي أي كموش خشيت؟

المشارك: كموش الصغير / الصغيرة

الباحثة: هذا (فديره). مين هذا؟

المشاركة: فديره

المشارك: فيديره

الباحثة: و هذا فيديره ثاني.

الباحثة (مسكة بالدميتين): هذا فيديره و هذا فيديره ثاني.

الباحثة (بعد أن أخفت إحدى الدميتين خلف ظهرها): الحين قولني لي أي فيديره خشيت؟

المشاركة: الأصفر ؟ الصفرة
Appendix C: Example Arabic script for Experiment III with example answers

الباحثة: هنا عني (بنت). مين عندي؟
المشارك: بنت.

الباحثة: و هنا بعد عندي بنت.

المشارك (ممسكة بالدميتين): عندي بنت و بنت.

الباحثة (بعد أن أخفت إحدى الدميتين خلف ظهرها): قول لي أي بنت خشيت؟
المشارك: البنت الوردية.

المشارك: دغفل.

المشارك (ممسكة بالدميتين): عندي دغفل و دغفل.

الباحثة (بعد أن أخفت إحدى الدميتين خلف ظهرها): قوللي لي أي دغفل خشيت؟
المشارك: دغفل الأخضر/ دغفل الخضراء
Appendix D: Pictures used for introducing the words ‘mother’ and ‘father’ in Experiment I
Appendix E: pictures of the toys used in Experiment I
Appendix F: Pictures of the 3D characters used in Experiment II
Appendix G: Pictures of the 3D characters used in Experiment III
Appendix H: the mixed effect model

➢ Model 1
Attempting at making a model with full interactions and random intercepts. This results in singularity
```
Model1 = glmer(GG ~ AGE*ENDING*NG*Agree + (1|participant) + (1|item), data = mydata, family="binomial", control = glmerControl(optimizer = 'bobyqa'))
```

➢ Model 2
Keeping the random intercepts but removing the four-way interactions. This results in singularity.
```
Model2 = glmer(GG ~ AGE+ENDING+NG+Agree+AGE:ENDING+AGE:NG+AGE:Agree+ENDING:NG+ENDING:Agree+AGE:ENDING:Agree + (1|participant) + (1|item), data = mydata, family="binomial", control = glmerControl(optimizer = 'bobyqa'))
```

➢ Model 3
Removing three-way interactions. This also results in singularity.
```
Model3 = glmer(GG ~ AGE+ENDING+NG+Agree+AGE:ENDING+AGE:NG+AGE:Agree+ENDING:NG+ENDING:Agree+AGE:ENDING:Agree + (1|participant) + (1|item), data = mydata, family="binomial", control = glmerControl(optimizer = 'bobyqa'))
```

➢ Model 4
Removing random intercept for items. Successful model.
```
```
(1|participant), data = mydata, family="binomial", control = glmerControl(optimizer = 'bobyqa', optCtr = list(maxfun = 1e9)))

```
summary(Model4)
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) [ glmerMod ]
Family: binomial ( logit )
Data: mydata
Control: glmerControl(optimizer = "bobyqa", optCtr = list(maxfun = 1e+09))

AIC      BIC   logLik d eviance df.resid
4104.5   4447.6  -1999.3   3998.5     4734

Scaled residuals:
   Min      1Q  Median      3Q     Max
-9.3264 -0.4113 -0.0330  0.4544 11.1567

Random effects:
  Groups   Name        Variance Std.Dev.
          participant (Intercept) 1.113     1.055
Number of obs: 4787, groups: participant, 202

Fixed effects:
       Estimate   Std. Error   z value  Pr(>|z|)
(Intercept) -6.06578    0.88377   -6.864    6.72e-12 ***

289
<table>
<thead>
<tr>
<th>Variable</th>
<th>Value1</th>
<th>Value2</th>
<th>Value3</th>
<th>Value4</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGEA</td>
<td>3.04163</td>
<td>1.32429</td>
<td>2.297</td>
<td>0.02163</td>
</tr>
<tr>
<td>AGEB</td>
<td>1.45046</td>
<td>0.94501</td>
<td>1.535</td>
<td>0.12482</td>
</tr>
<tr>
<td>AGEC</td>
<td>2.69422</td>
<td>0.90221</td>
<td>2.986</td>
<td>0.00282**</td>
</tr>
<tr>
<td>AGED</td>
<td>0.68036</td>
<td>1.04276</td>
<td>0.652</td>
<td>0.51410</td>
</tr>
<tr>
<td>ENDINGa</td>
<td>2.44636</td>
<td>0.76490</td>
<td>3.198</td>
<td>0.00138**</td>
</tr>
<tr>
<td>NGfemale</td>
<td>5.42070</td>
<td>0.87474</td>
<td>6.197</td>
<td>5.76e-10***</td>
</tr>
<tr>
<td>NGnon</td>
<td>1.98820</td>
<td>0.86733</td>
<td>2.292</td>
<td>0.02189</td>
</tr>
<tr>
<td>AgreeF</td>
<td>5.55403</td>
<td>0.85663</td>
<td>6.484</td>
<td>8.96e-11***</td>
</tr>
<tr>
<td>AGEA:ENDINGa</td>
<td>1.31646</td>
<td>1.24331</td>
<td>1.059</td>
<td>0.28967</td>
</tr>
<tr>
<td>AGEB:ENDINGa</td>
<td>0.92900</td>
<td>0.78423</td>
<td>1.185</td>
<td>0.23617</td>
</tr>
<tr>
<td>AGEC:ENDINGa</td>
<td>-0.95229</td>
<td>0.74331</td>
<td>-1.281</td>
<td>0.20014</td>
</tr>
<tr>
<td>AGED:ENDINGa</td>
<td>0.00578</td>
<td>0.85271</td>
<td>0.007</td>
<td>0.99459</td>
</tr>
<tr>
<td>AGEA:NGfemale</td>
<td>-3.83797</td>
<td>1.34417</td>
<td>-2.855</td>
<td>0.00430**</td>
</tr>
<tr>
<td>AGEB:NGfemale</td>
<td>-2.48269</td>
<td>0.92073</td>
<td>-2.696</td>
<td>0.00701**</td>
</tr>
<tr>
<td>AGEC:NGfemale</td>
<td>-2.70721</td>
<td>0.88864</td>
<td>-3.046</td>
<td>0.00232**</td>
</tr>
<tr>
<td>AGED:NGfemale</td>
<td>-0.36479</td>
<td>1.02531</td>
<td>-0.356</td>
<td>0.72200</td>
</tr>
<tr>
<td>AGEA:NGnon</td>
<td>-2.50571</td>
<td>1.56568</td>
<td>-1.600</td>
<td>0.10951</td>
</tr>
<tr>
<td>AGEB:NGnon</td>
<td>-1.41060</td>
<td>0.89962</td>
<td>-1.568</td>
<td>0.11688</td>
</tr>
<tr>
<td>AGEC:NGnon</td>
<td>-1.54369</td>
<td>0.86383</td>
<td>-1.787</td>
<td>0.07393</td>
</tr>
<tr>
<td>AGED:NGnon</td>
<td>-0.70404</td>
<td>0.96663</td>
<td>-0.728</td>
<td>0.46640</td>
</tr>
<tr>
<td>AGEA:AgreeF</td>
<td>-2.87992</td>
<td>1.29266</td>
<td>-2.228</td>
<td>0.02589*</td>
</tr>
<tr>
<td>AGEB:AgreeF</td>
<td>-1.22923</td>
<td>0.89236</td>
<td>-1.378</td>
<td>0.16836</td>
</tr>
<tr>
<td>AGEC:AgreeF</td>
<td>-2.37689</td>
<td>0.85588</td>
<td>-2.777</td>
<td>0.00548**</td>
</tr>
<tr>
<td>AGED:AgreeF</td>
<td>-0.62922</td>
<td>0.99055</td>
<td>-0.635</td>
<td>0.52529</td>
</tr>
<tr>
<td>ENDINGa:NGfemale</td>
<td>-0.85877</td>
<td>0.76666</td>
<td>-1.120</td>
<td>0.26265</td>
</tr>
<tr>
<td>ENDINGa:NGnon</td>
<td>0.10563</td>
<td>0.72613</td>
<td>0.145</td>
<td>0.88434</td>
</tr>
<tr>
<td>ENDINGa:AgreeF</td>
<td>-1.03482</td>
<td>0.71879</td>
<td>-1.440</td>
<td>0.14996</td>
</tr>
<tr>
<td>Factor</td>
<td>Estimate 1</td>
<td>Estimate 2</td>
<td>Estimate 3</td>
<td>Estimate 4</td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td>NGfemale:AgreeF</td>
<td>-1.67731</td>
<td>0.85431</td>
<td>-1.963</td>
<td>0.04961</td>
</tr>
<tr>
<td>NGnon:AgreeF</td>
<td>-0.44636</td>
<td>0.83843</td>
<td>-0.532</td>
<td>0.59447</td>
</tr>
<tr>
<td>AGEA:ENDINGa:NGfemale</td>
<td>0.50367</td>
<td>1.34196</td>
<td>0.375</td>
<td>0.70742</td>
</tr>
<tr>
<td>AGEB:ENDINGa:NGfemale</td>
<td>-0.05261</td>
<td>0.76540</td>
<td>-0.069</td>
<td>0.94520</td>
</tr>
<tr>
<td>AGEC:ENDINGa:NGfemale</td>
<td>0.93229</td>
<td>0.74764</td>
<td>1.247</td>
<td>0.21240</td>
</tr>
<tr>
<td>AGED:ENDINGa:NGfemale</td>
<td>-0.27943</td>
<td>0.84908</td>
<td>-0.329</td>
<td>0.74209</td>
</tr>
<tr>
<td>AGEA:ENDINGa:NGnon</td>
<td>2.07073</td>
<td>1.52923</td>
<td>1.354</td>
<td>0.17570</td>
</tr>
<tr>
<td>AGEB:ENDINGa:NGnon</td>
<td>0.13086</td>
<td>0.68319</td>
<td>0.192</td>
<td>0.84810</td>
</tr>
<tr>
<td>AGEC:ENDINGa:NGnon</td>
<td>0.02546</td>
<td>0.65328</td>
<td>0.039</td>
<td>0.96891</td>
</tr>
<tr>
<td>AGED:ENDINGa:NGnon</td>
<td>-0.11054</td>
<td>0.68851</td>
<td>-0.161</td>
<td>0.87245</td>
</tr>
<tr>
<td>ENDINGa:NGfemale:AgreeF</td>
<td>-0.15684</td>
<td>0.55893</td>
<td>-0.281</td>
<td>0.77901</td>
</tr>
<tr>
<td>ENDINGa:NGnon:AgreeF</td>
<td>-0.11850</td>
<td>0.58859</td>
<td>-0.201</td>
<td>0.84045</td>
</tr>
<tr>
<td>AGEA:ENDINGa:AgreeF</td>
<td>-0.76458</td>
<td>1.21863</td>
<td>-0.627</td>
<td>0.53039</td>
</tr>
<tr>
<td>AGEB:ENDINGa:AgreeF</td>
<td>-0.82154</td>
<td>0.67789</td>
<td>-1.212</td>
<td>0.22555</td>
</tr>
<tr>
<td>AGEC:ENDINGa:AgreeF</td>
<td>0.43908</td>
<td>0.65154</td>
<td>0.674</td>
<td>0.50037</td>
</tr>
<tr>
<td>AGED:ENDINGa:AgreeF</td>
<td>-0.50057</td>
<td>0.75396</td>
<td>-0.664</td>
<td>0.50674</td>
</tr>
<tr>
<td>AGEA:NGfemale:AgreeF</td>
<td>2.41082</td>
<td>1.39450</td>
<td>1.729</td>
<td>0.08384</td>
</tr>
<tr>
<td>AGEB:NGfemale:AgreeF</td>
<td>0.38722</td>
<td>0.85791</td>
<td>0.451</td>
<td>0.65173</td>
</tr>
<tr>
<td>AGEC:NGfemale:AgreeF</td>
<td>1.11073</td>
<td>0.84855</td>
<td>1.309</td>
<td>0.19054</td>
</tr>
<tr>
<td>AGED:NGfemale:AgreeF</td>
<td>0.55219</td>
<td>0.95578</td>
<td>0.578</td>
<td>0.56344</td>
</tr>
<tr>
<td>AGEA:NGnon:AgreeF</td>
<td>1.45906</td>
<td>1.57872</td>
<td>0.924</td>
<td>0.35538</td>
</tr>
<tr>
<td>AGEB:NGnon:AgreeF</td>
<td>0.55276</td>
<td>0.82511</td>
<td>0.670</td>
<td>0.50291</td>
</tr>
<tr>
<td>AGEC:NGnon:AgreeF</td>
<td>1.46772</td>
<td>0.80426</td>
<td>1.825</td>
<td>0.06801</td>
</tr>
<tr>
<td>AGED:NGnon:AgreeF</td>
<td>0.61894</td>
<td>0.88997</td>
<td>0.695</td>
<td>0.48677</td>
</tr>
</tbody>
</table>

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1